The Network Structure of International Trade*

Thomas Chaney†
University of Chicago, NBER and CEPR

July, 2010

Abstract

I build a simple dynamic model of the formation of an international social network of importers and exporters. Firms can only export into markets in which they have a contact. They acquire new contacts both at random, and via their network of existing contacts. This model explains (*i*) the cross-sectional distribution of the number of foreign markets accessed by individual exporters, (*ii*) the cross-sectional geographic distribution of foreign contacts, and (*iii*) the dynamics of firm level exports. I show that the firm level dynamics of trade can precisely explain the observed cross section of firm level exports. All theoretical predictions have a very tight connection with the data.

Introduction

Individual firms differ hugely in their exposure to international trade. Whereas most firms do not export abroad, and a large fraction of exporters export to a single foreign market, few firms export to a large number of countries. This heterogeneity in the access to foreign markets of individual firms has dramatic implications for the patterns of international trade. Melitz (2003) shows that in the presence of heterogeneity in the ability of individual firms to access foreign markets, a reduction in trade barriers can induce aggregate productivity gains. Bernard, Eaton, Jensen and Kortum (2003) and Chaney (2008) show that in the presence of firm heterogeneity, firm level exports aggregate up to the well established gravity equations in international trade, but that the sensitivity of trade flows with respect to trade barriers is magnified. The source of this heterogeneity in the ability of individual firms to access foreign markets however remains largely

*I want to thank Xavier Gabaix with whom I started thinking about the ideas developed in this paper. I am grateful to Sylvain Chassang, Ali Hortaçsu, Pierre-Louis Lions, Bob Lucas, Marc Melitz, Roger Myerson, David Sraer, and seminar participants in Chicago (Math and Econ), Sciences Po (Paris), and the Toulouse School of Economics for helpful discussions. I am indebted to Ferdinando Monte and Enghin Atalay for their superb research assistance.

†Contact: Department of Economics, The University of Chicago, Chicago, IL 60637. Tel: 773-702-5403. Email: tchaney@uchicago.edu.

unexplained. Whereas Bernard, Eaton, Jensen and Kortum (2003) or Melitz (2003) assume that this heterogeneity is entirely driven by productivity differences, Armenter and Koren (2009) point out that productivity differences can only account for a fraction of the exposure to international markets.

In this paper, I develop a geographic extension of the model developed in Jackson and Rogers (2007), and offer a simple explanation for the heterogeneous ability of individual firms to access foreign markets based on the formation of an international social network. I show that a dynamic network where exporters meet foreign importers both at random and through their network of existing foreign contacts matches remarkably well the cross-section and the time-series patterns of firm level entry into foreign markets. The predictions of the model on both the cross-sectional distribution of the number of foreign contacts, and on the cross-sectional geographic distribution of foreign contacts are supported by the data. Furthermore, this model generates novel predictions for the dynamic evolution of trade flows. I show how the entry of individual exporters into a given country is influenced by changes in aggregate trade flows between third countries, in a way that is consistent with the model and with the cross-sectional evidence on the distribution of foreign contacts. I also show new evidence on the different behavior of large versus small exporters that is consistent with the theory.

Following Jackson and Rogers (2007), I assume that potential exporters meet foreign contacts in two distinct ways. First, they can meet foreign contacts at random, which is a reduced form for the active search for foreign trading partners. Second, once a firm has acquired some foreign contacts, it can meet the contacts of those contacts. This process generates predictions for the steady state distribution of the number of foreign contacts across exporters, and for the geographic distribution of these contacts across exporters.

The possibility to use existing contacts to find new contacts gives an advantage to firms that already have many contacts. This generates a fat tailed distribution of the number of contacts across firms. The key parameter that determines the shape of the cross-sectional distribution of the number of contacts is the relative importance of random versus network-based meetings. The empirical distribution of the number of foreign contacts is remarkably close to the theoretical predictions. It allows for a precise estimation of the relative importance of random versus network-based meetings. Moreover, direct evidence on the time-series evolution of firm level trade flows confirms the assumed mechanism. I find that the more contacts a firm has, the more likely it is to acquire additional contacts. The time-series evidence is both qualitatively and quantitatively in line with the cross-sectional evidence.

The more novel contribution of this paper is that the network formation is embedded into geographic space. Network-based meetings allow say a French exporter that has a acquired a contact in Japan to radiate from Japan as if it were a Japanese firm itself. It does so by using its Japanese contacts as a remote hub from which it can expand out of Japan. The theory therefore predicts that as firms acquire more foreign contacts, they can expand into more remote countries, so that their exports become geographically more dispersed. Note that the speed at which the geographic dispersion increases depends on the relative importance of random versus network-based meetings. I find strong empirical evidence that geographic dispersion increases with the number of foreign contacts, in a way that is quantitatively in line with the theory and the cross-sectional distribution of the number of foreign contacts.

This is a theory of a network. Therefore, a shock that happens anywhere will be transmitted to all the components in the network, with an intensity that depends on the structure of the network. I find empirical support for these novel predictions on the dynamics of firm level trade flows. For instance, I show that an increase in the volume of trade between country a and b will have a positive impact on the probability that a French firm that already exports to a starts exporting to b, but not on firms that do not export to country a yet. The magnitude of this effect is in line qualitatively and quantitatively with the theory and the cross-sectional distribution of the number of foreign contacts.

Existing models of international trade at the firm level fail to match any of the predictions of the proposed model, and are therefore at odds with the novel empirical regularities I uncover. Melitz (2003) and its extension in Chaney (2008), or Bernard, Eaton, Jensen and Kortum (2003) do not make any systematic prediction regarding the cross-sectional distribution of the number of foreign markets reached by individual exporters. Even under the assumption of Pareto distributed productivity shocks in Chaney (2008), or Fréchet distributed productivities in Bernard, Eaton, Jensen and Kortum (2003), the distribution of the number of foreign markets reached depends on the distribution of exogenous parameters such as country size, productivities, and bilateral export costs. Not only do these models have nothing to say about the distribution of those exogenous parameters, but there is no empirical support for the ad hoc distribution that would be needed to generate the observed distribution of the number of foreign markets reached.

The proposed model offers a novel theory for the entry of individual exporters into foreign markets. It generates a series of novel predictions that find strong support in the data.

The remainder of the paper is organized as follows. In section 1, I present a simple theoretical model of the formation of an international network of importers and exporters. In section 2, I test empirically the main theoretical predictions of the model. I relegate to the Appendix all mathematical proofs (Appendix A), some additional economic assumptions (Appendix B), the description of the data and robustness checks (Appendix C).

1 A simple dynamic model of network formation

In this section, I develop a simple model of the formation of an international network of importers and exporters. This model is an extension of Jackson and Rogers (2007), where I embed the formation of links into geographic space.

The purpose of this model is to explain the extensive margin of international trade, that is the patterns of entry of individual exporters into different foreign markets. I assume that individual firms enter a foreign market if and only if they have acquired a contact in that market.[1] The proposed model formalizes one particular way through which exporters enter foreign markets: firms may either meet foreign contacts at random, or alternatively, once they have acquired some foreign contacts, they can meet some of the contacts of their contacts.

This model delivers a series of predictions that are tightly supported by data on firm level trade. First, the model replicates precisely the distribution of entry of individual firms into different foreign markets (see Proposition 2 and the empirical test in Section 2.2). Second, the model correctly predicts that the geographic dispersion of export destinations increases with the number of foreign markets a firm serves (see Proposition 3 and the empirical test in Section 2.3). Third, the main assumption of the model are supported by the data, and the model's predictions for the dynamics of firm level and aggregate trade flows are confirmed by the data (see the empirical tests in Section 2.4).

[1] The actual trade that occurs once two firms are linked can be derived in a simple extension of the Krugman (1980) model, as shown in Appendix B.1.

1.1 Set-up

The formal set-up is as follows. Firms are distributed over a one-dimensional infinite space, represented by $\mathbb{R}$.[2] Time is discrete and starts at $t = 0$, when an initial cohort of firms is uniformly distributed over space with a density of 1 in each location. From then on, population grows in each location at the same constant growth rate γ. So at time t, the total density of firms in each location is equal to $N_t = (1+\gamma)^t$. The density of new firms born between time t and $t+1$ is simply γN_t.[3] At a point in time t, a firm is designated by the pair $(x, i) \in \mathbb{R} \times \mathbb{N}$, where $x \in \mathbb{R}$ is the firm's location, and $i \leq t \in \mathbb{N}$ is the firm's birth date. Once born, a firm never changes location, and it never dies. Firms form directed links with one another.

Before describing the formation and the evolution of this network, it will be useful to introduce some notations. Each firm has both an *out-degree distribution* (the set of firms with which it has initiated a contact), and an *in-degree distribution* (the set of firms that have initiated a contact with it). I will mostly focus on the in-degree distribution of firms. The in-degree distribution of firm (x, i) at time t is described by a continuous distribution $f_{x,i,t}$,

$$f_{x,i,t} : \mathbb{R} \to \mathbb{R}^+ \text{ with } \int_{\mathbb{R}} f_{x,i,t}\,(y-x)\,dy \equiv M_{x,i,t}$$

so that the mass of firms located in $[a, b] \subset \mathbb{R}$ that know firm (x, i) at time t is $\int_a^b f_{x,i,t}\,(y-x)\,dy$. I will systematically use the expression "i knows j" in the sense that i initiated a contact with j, and "i is known by j" in the sense that j has initiated a contact with i. The *in-degree* of firm (x, i) at time t, defined as the total mass of firms that know (x, i) at time t is then simply $M_{x,i,t}$.[4]

Three clarifying remarks are in order.

First, the model assumes that links are directed (outward links evolve differently from inward links), even though the data on individual exporters does not contain any information on which side initiated a link (the importer or the exporter). I will later assume that when a link exists between two firms, irrespective of which firm initiated the link, they exchange one unit of output. The directedness of the network is therefore irrelevant for trade. However, the specific directed link formation I model greatly simplifies the analysis.

[2]Note that most results hold in a space of higher dimensionality. I will consider in the empirical applications the two-dimensional geographic space. This model can be applied to non physical spaces, such as product spaces, or preference spaces, as long as the symmetry assumptions made below are satisfied.

[3]Note that the total mass of firms in the system is infinite, but the density of firms in each location remains well defined.

[4]Note that $f_{x,i,t}$ is not a probability density function, since it sums up to $M_{i,t}$ which will differ from 1.

Second, this is a model of firms meeting firms, with no a priori notion of national boundaries. As I will describe when I bring this model to the data in Section 2, countries can be thought as arbitrary segments partitioning the one-dimensional space of the model.

Third, I use a continuous representation of space and of the number of firms. This continuous representation simplifies the analysis greatly. However, I will use the language of a discrete model to describe the set up and the intuitions of the model. For instance, I will say that firm (x, i) is known by exactly a number $f_{x,i,t}(y-x)\,dy$ of different firms location y (strictly speaking, in a small neighborhood dy around y), it is known by a total number of $\int_a^b f_{x,i,t}(y-x)\,dy$ different firms in an entire interval $[a, b]$, and by a total number of $M_{x,i,t}$ different firms worldwide. This language is both formally rigorous, and intuitively accessible.

1.2 Network formation

The process through which firms acquire both an out-degree and an in-degree distribution will be purposefully simple. I assume that firms acquire their out-degree distribution in the first period of their life, and never alter it subsequently. So the out-degree distribution of each firm will be trivial and of no interest. All the action will take place on the in-degree distribution, with existing firms being met by some of the newly born firms of each new cohort. The evolution of the network is described next. Each period, newly born firms meet existing firms in two distinct ways.

First, each newly born firm in any location randomly samples a mass m_r out of the existing firms (where m_r stands for *random meetings*). Geographic distance however affects the link formation in the following way. Firm x forms $m_r f_0(y-x)\,dy$ successful links with firms in location y, where f_0 is a well defined symmetric probability density function.[5] All successive waves of random meetings are assumed independent from one another.

Second, a newly born firm will meet a mass m_n out of the union of the out-degree distributions of all m_r firms met at random (where m_n stands for *network-based meetings*). For simplicity, I assume that geographic distance plays no role in impeding the transmission of information between firms that are linked. In other words, once a newly born firm has met a set of existing firms at random, it is equally likely to meet any of the firms in union of their out-degree distribution, no matter where they are located. This strong assumption is meant to capture the existence of

[5]Note that geographic distance will actually *hinder* the formation of random links only if $\partial f_0(|x|)/\partial|x| < 0$. I do not need to make that assumption for all $|x|$'s. It is for instance possible that geographic distance facilitates link formation over some range of distances.

informational frictions.[6] Once formed, a contact is never lost. Note that trivially, all firms have the same out-degree equal to $(m_r + m_n)$.[7]

In order to ensure that the initial conditions are well defined for all configurations of the parameters, I assume that each firm is born with an in-degree $M_0 \geq 0$.[8] This initial in-degree, if present, is distributed over space according to the density f_0.

To further simplify the model, I will use the following mean-field approximation. I assume that the number of links any firm receives is exactly the average number of links it is supposed to receive. In other words, I assume away the intrinsic randomness of the network formation, so that I do not have to keep track of the probability distribution around the mean number of contacts received by each firm. Jackson and Rogers (2007) show numerically that this mean-field approximation is innocuous.[9]

With this simple process for link formation, I can now describe how the in-degree distribution of a firm born at time i evolves over time. Given the geographic symmetry of the network, I will consider a firm located at the origin ($x = 0$) without loss of generality. To ease notations, I will drop the $x = 0$ subscript from now on.[10]

Assume that firm i has an in-degree distribution $f_{i,t}(\cdot)$ at time t. At time $t + 1$, some newly born firms will meet i at random. There are γN_t newly born firms in each location. Newly born firms located in x meet $m_r f_0(0 - x)\, dx = m_r f_0(x)\, dx$ firms in the origin. Since there are N_t firms at the origin, the number of new random meetings received by firm i originating from location x is given by $\frac{\gamma N_t m_r}{N_t} f_0(x)\, dx = \gamma m_r f_0(x)\, dx$.

In addition, some of the newly born firms will meet firms in the in-degree distribution of firm i and form a link with i through a network-based meeting. For instance, a newly born firm located in x may meet at random a firm located in y that knows firm i and form a link with i via this firm in y. A newly born firm located in x meets a total of m_r firms at random. Each of these m_r firms has an out-degree of $(m_r + m_n)$. Firm x will form a network-based link by picking at random m_n out of those $m_r(m_r + m_n)$ firms. Conditional on being in the union of out-

[6]This process can be rationalized in a simple model with information asymmetry, as shown in Appendix B.2.

[7]See Proposition 4 on page 39 in Appendix A for a formal derivation of the out-degree distribution. I am grateful to Enghin Atalay for solving for this distribution.

[8]The condition $M_0 > 0$ has to hold only in the case of purely preferential networks ($m_r = 0$), or else a firm would never be contacted. For simplicity, I will consider mostly cases where $M_0 = 0$. All the results hold in a slightly more general form for $M_0 > 0$, as shown in Appendix A.

[9]Jackson and Rogers (2007) are only interested in the total number of links a firm receives, whereas I care about both the total number of links and their geographic locations. I conjecture that the mean-field approximation is innocuous for the geographic distribution of contacts as well.

[10]Trivially, $f_{0,i,t}(y) = f_{x,i,t}(y + x)$ for any $(x, y) \in \mathbb{R}^2$.

degrees of the firms randomly met by firm x, any given firm has a probability $\frac{m_n}{m_r(m_r+m_n)}$ of being chosen through one of the network-based meetings. I can now add all the pieces of network-based meetings together. Each firm in location y is met by $\frac{\gamma N_t m_r}{N_t} f_0(y-x)\,dy = \gamma m_r f_0(x-y)\,dx$ firms from location x. The number of firms in location y that already know i is given by $f_{i,t}(y)\,dy$, inherited from the previous period. The number of firms located in x that form network-based meetings with a firm born at date i and located in the origin via a firm in y is then equal to $\gamma m_r f_0(x-y)\,dx \times f_{i,t}(y)\,dy \times \frac{m_n}{m_r(m_r+m_n)}$. A network-based meeting can potentially be intermediated by firms located in any location $y \in \mathbb{R}$. So the in-degree distribution of a firm born in i evolves recursively according to,

$$f_{i,t+1}(x)\,dx = f_{i,t}(x)\,dx + \gamma m_r f_0(x)\,dx + \gamma m_r \frac{m_n}{m_r(m_r+m_n)} \int_{y\in\mathbb{R}} f_0(x-y)\,f_{i,t}(y)\,dydx$$

$$\text{or } f_{i,t+1} = f_{i,t} + \gamma m_r f_0 + \frac{\gamma m_n}{(m_r+m_n)} f_0 * f_{i,t} \tag{1}$$

where $*$ stands for the convolution product. This recursive structure allows to derive a simple solution for the in-degree distribution of all firms. The following proposition describes this distribution.

Proposition 1 *The in-degree distribution of firm i at time t, $f_{i,t}$, is given by,*

$$f_{i,t} = \underbrace{\left(\delta + \gamma \frac{m_n}{m_r+m_n} f_0\right) * \ldots * \left(\delta + \gamma \frac{m_n}{m_r+m_n} f_0\right)}_{t-i \text{ times}} * \frac{m_r}{m_n}(m_r+m_n)\,\delta - \frac{m_r}{m_n}(m_r+m_n)\,\delta$$

where δ is the Dirac delta function, $*$ is the convolution product, γ is the growth rate of the population, f_0 is the geographic distribution of contacts for newborn firms, and m_r and m_n are respectively the number of random and network-based meetings of newly born firms.

Proof. See Appendix A page 33. ■

This solution for the in-degree distribution of firms allows to easily describe several moments of this distribution. The next two section analyze two of these moments. Section 1.3 describes the distribution of the total number of contacts across firms, whereas Section 1.4 describes the geographic dispersion of contacts across firms.

1.3 The distribution of foreign contacts

As firms age, they acquire more contacts by forming links with a fraction of each newly born cohort. Given the exponential population growth, the total number of contacts of a firm evolves approximately exponentially. The rate of growth however depends on the relative importance of random versus network-based meetings. Intuitively, the growth rate of the number of contacts of a firm with many existing contacts is highest when most links are acquired through network-based contacts, and is lowest when most links are acquired at random. This is due to the fact that network-based meetings give a larger advantage for acquiring new contacts to firms that already have many contacts, since a firm can use each and every one of those existing contacts as a bait to "fish" for new contacts. The following lemma describes the in-degree of any given firm.

Lemma 1 *The total mass of contacts at time t of a firm born in i, i.e. the in-degree of this firm, $M_{i,t}$, is given by,*

$$M_{i,t} = \left(1 + \gamma \frac{m_n}{m_r + m_n}\right)^{t-i} \times \frac{m_r}{m_n}(m_r + m_n) - \frac{m_r}{m_n}(m_r + m_n)$$

where γ is the growth rate of the population, and m_r and m_n are respectively the number of random and network-based meetings of newly born firms.

Asymptotically, the in-degree grows at a constant rate. This growth rate is highest (equal to $1+\gamma$) when most links are network-based ($\frac{m_n}{m_r+m_n} \approx 1$), and lowest (equal to 1) when most link are random ($\frac{m_n}{m_r+m_n} \approx 0$).

Proof. See Appendix A page 34. ■

This model encompasses the space-less model of Jackson and Rogers (2007). The total number of contacts at time t of any firm born in i is exactly as in the Jackson and Rogers model. Moreover, if distance were to represent an insurmountable barrier to the acquisition of information, so that newly born firms could only meet at random other firms located in the very same location, then each location would behave like an isolated island that replicates the Jackson and Rogers model.[11] From this observation, it is easy to see that the distribution of in-degree across firms is the same in any arbitrary subset of locations. The following proposition gives a precise characterization of the distribution of in-degree across heterogeneous firms.

[11] Formally, this would correspond to the case where $f_0 = \delta$, the Dirac delta function. A formal proof of this result can be found in the proof of Proposition 1 in Appendix A page 33.

Proposition 2 *For a population growth rate γ small, the distribution of the number of contacts, M, across individual firms in any arbitrary set of locations is given by the cumulative distribution function,*

$$F(M) = 1 - \left(\frac{r \times m}{M + r \times m}\right)^{1+r}$$

where $r = m_r/m_n$ is the ratio of random versus network-based meetings, and $m = (m_r + m_n)$ is the total number of contacts made by newly born firms.

Proof. See Appendix A page 35. ■

I will show in section 2.2 that proposition 2 is remarkably well supported by data on firm level exports, with random meetings accounting for roughly 60% of all new meetings.[12]

Let me briefly describe the properties of the cross sectional in-degree distribution, and provide some intuition for those properties. The upper tail of the in-degree distribution asymptotes to a scale-free Pareto distribution, whereas the lower tail is close to an exponential distribution.[13] Firms that already have acquired many contacts will almost only meet new firms through network-based meetings. Random meetings become a negligible fraction of their new contacts. Hence, the growth of the number of contacts for well connected firms is roughly proportional to the number of contacts they already have. This explains why the upper tail of the in-degree distribution converges to a scale-free Pareto distribution. On the other hand, young firms, or firms with few existing contacts, meet newly born firms mostly at random. The distribution in the lower tail is therefore described by a discrete binomial distribution, which corresponds to a continuous exponential distribution.

As the relative importance of random versus network-based meetings changes, the range over which the in-degree distribution is Pareto versus exponential changes. In the polar case where almost all meetings are random ($r \to +\infty$), the whole in-degree distribution is exponential, whereas in the opposite polar case where almost all meetings are network-based ($r \to 0$), the in-degree distribution follows a Zipf law.[14] In cases in-between, the in-degree is given by some combination of these polar cases.

[12] I estimate using firm level trade data that $m_r/m_n \approx 1.58$, so that $m_r/(m_r + m_n) \approx .6$.

[13] Note that $1 - F(M) \approx \left(\frac{M}{r \times m}\right)^{-(1+r)}$ for M large, a Pareto distribution; whereas $1 - F(M) \approx \exp\left(-(1+r)\frac{M}{r \times m}\right)$ for M small, an exponential distribution.

[14] See the proof of Proposition 2 in Appendix A page 35 for a formal proof of this statement.

1.4 The geography of exports

The characterization of the in-degree distribution in Proposition 1 not only allows me to study the behavior of the total mass of contacts of a given firm, but it also provides a precise description of how the geographic dispersion of a firm's contacts evolves over time. If all interactions between firms were to only take place locally, space would never play any role, and each location would behave exactly as in Jackson and Rogers (2007).[15] If firms do interact beyond the boundaries of their location, geographic distance starts playing a novel role. Because geographic distance impedes the formation of random meetings, newly born firms are constrained to interact mostly over short distances. As firms age, their ability to use their existing network of contacts allows them to reach deeper into geographic space.

A tractable way to describe the geographic dispersion of a firm's network is to follow the second moment of the distance from a firm's contacts over time. This second moment is given by the variance $\sigma_{i,t}^2$ of the in-degree distribution $f_{i,t}$, defined as,

$$\sigma_{i,t}^2 \equiv \int_{\mathbb{R}} x^2 \frac{f_{i,t}}{M_{i,t}}(x)\, dx$$

where $f_{i,t}/M_{i,t}$ is the well defined probability density function of the location a firm's contacts. The variance $\sigma_{i,t}^2$ will change over time because of two distinct forces. First, new waves of firms will meet firm i at random. Since all successive cohorts of newly born firms face the same geographic hurdle for forming random contacts, this first effect will not affect the geographic shape of a firm's contacts. Second, new waves of firms will meet firm i through its existing network of contacts. One can think of the existing network of contacts of a firm as tentacles that allow this firm to reach deeper into geographic space. In other words, as a firm acquires faraway contacts, it can use those contacts as a platform to reach contacts even further away. This second effect will expand the geographic dispersion of a firm's contacts.

The following lemma and proposition describe this result formally. Lemma 2 shows how the dispersion of the distance from a firm's contacts evolves as a firm ages, and Proposition 3 describes the relationship between a firm's in-degree and the variance of the distance from a firm's contacts.

Lemma 2 *The geographic dispersion of a firm's contacts increases as a firm ages. Formally,* $\sigma_{i,t}^2$,

[15] As noted earlier, and as proven formally in the proof of Proposition 1 in Appendix A page 33, this would be the case if $f_0 = \delta$, the Dirac delta function.

the variance of the distance from a firm's contacts is asymptotically proportional to a firm's age,

$$\sigma_{i,t}^2 \underset{t-i\to\infty}{\approx} (t-i)\,\frac{\gamma}{1+r+\gamma}\sigma_0^2$$

at time t, for a firm born at date i,where γ is the growth rate of the population, $r = m_r/m_n$ is the ratio of random versus network-based meetings, and σ_0^2 is the variance of the distance from a firm's random contacts.

Proof. See Appendix A page 37. ■

From this Lemma, I can derive a prediction for the relationship between a firm's in-degree and the variance of the distance from a firm's contacts. The following proposition states this result formally.

Proposition 3 *The geographic dispersion of a firm's contacts increases with the firm's in-degree. For a population growth rate γ small, the variance of the distance from a firm's contacts for a firm with M contacts, $\sigma^2(M)$, is given by,*

$$\sigma^2(M) = \left(1+\frac{r\times m}{M}\right)\ln\left(1+\frac{M}{r\times m}\right)\times\sigma_0^2$$

where σ_0^2 is the variance of the distance from a newly born firm's contacts, $r = m_r/m_n$ is the ratio of random versus network based meetings, and $m = (m_r + m_n)$ is the total number of contacts made by newly born firms.

Proof. See Appendix A page 38. ■

Note that all the results derived hold for any arbitrary probability density f_0 with a finite variance.[16] Note also that I only need information about the geographic dispersion of random contacts, σ_0^2, to explain the dispersion of the contacts of all firms.

The reason why the geographic dispersion of a firm's contacts increases with this firm's in-degree comes entirely from the network based meetings. It is easy to see analytically from Proposition 3 that if all new contacts are made at random ($r \times m \to +\infty$), the variance $\sigma^2(M)$ is constant and equal to σ_0^2. As the relative share of network-based meetings increases ($r \times m$ shrinks), not only does $\sigma^2(M)$ become larger for all M's, but $\sigma^2(M)$ increases faster with M.

[16]The only condition on f_0 is that it is symmetric, and admits a finite second moment. As noted earlier, the special case of zero variance ($f_0 = \delta$) collapses exactly to the space-less model of Jackson and Rogers.

The intuition for this result is the following. If all contacts are made at random, older firms have more contacts than younger ones because they have been contacted by a larger number of waves of entrants since their birth. However, since each new wave of entrants is independently and identically distributed over space, the geographic distribution of a firm's contacts remains unchanged. It is simply given by the distribution of each wave of entrants (f_0 with variance σ_0^2). Network-based meetings follow a different spatial dynamic. Firms use their existing contacts as local hubs to acquire new contacts. This allows them to gradually expand deeper into space. Each new wave of network-based meetings therefore allows firms to increase the geographic dispersion of their contacts.

I will show in Section 2.3 that Proposition 3 is remarkably well supported by data on firm level exports. Using the estimate for the relative importance of random versus network-based meetings estimated from the cross section of entry into different markets,[17] and simply calibrating the units of measurement,[18] I can describe very precisely how the geographic dispersion of contacts increases as firms get more contacts.

1.5 Discussion

There are several alternative interpretations of the proposed model. The most literal one is that individual firms meet other individual firms, some of them located in foreign countries, in the way described by the model. I follow this literal interpretation when bringing the model to the data, and I explain in details how to circumvent the fact that I only have data on which countries a firm exports to, and not directly on how many contacts it has in each country.

A less literal interpretation is the following. Firms try to enter foreign contacts. The entry into the very first foreign market can be described as random, which is a reduced form for all the idiosyncrasies of different firms (type of product the firm is producing and the taste for that product in various foreign markets, specific comparative advantage of that particular firm, actual information that workers in that firm have about various foreign markets...). However, upon successfully entering a given country, this firm can acquire information locally, and has the option of expanding into other foreign markets from this given country. The quality of this information is simply discounted compared to information acquired from the home market. The model assumes

[17] $r \times m$ is estimated from fitting Proposition 2 to the data in Section 2.2.

[18] σ_0^2 is not a unit-free parameter.

that the discount in the quality of information is the same for all foreign markets.[19] In other words, a French exporter that has successfully entered the Japanese market can subsequently fan out into Asia as easily as a French exporter that has successfully entered the Argentine market can fan out into Latin America. So entry into each new market increases the chances of that exporter to enter yet another market. The geographic dispersion of a firm's exports increases as it enters more markets. The model makes the extreme assumption that for a French firm exporting to Japan and Germany, the difference in the information of that firm compared to that of a German or a Japanese firm is the same. However, one must bear in mind that it is arguably much easier for a French exporter to enter the German market than the Japanese market in the first place, which the model does take into account. As I show in the next empirical section, this simplification describes several dimensions of the data very well.

It is important to note that international trade models with heterogeneous firms, such as Bernard, Eaton, Jensen and Kortum (2003) or Melitz (2003), have nothing to say a priori about the distribution of the number of countries reached by different firms. In both models, the number of firms able to export to a given foreign market depends on the underlying distribution of productivities across firms, and on a productivity threshold for entering that specific market.[20] These productivity thresholds are functions only of exogenous parameters. So even for a simple distribution of productivities across firms, by changing those exogenous parameters, one can generate any arbitrary shape for the cross-sectional distribution of the number of foreign markets accessed. This distribution does not even have to be downward sloping. I develop this argument in more details in the next Section, and formally in the Appendix.[21] If trade barriers increase with distance, and if there is no systematic correlation between country size and distance from France, both models would correctly predict that the geographic dispersion of foreign markets increases with the number of markets a firm enters. However, neither model offers any specific prediction for the shape of this relationship. Again, by changing the exogenous parameters of the model, on can generate any arbitrary relationship between the geographic dispersion and the number of foreign markets entered. Finally, since those models are static in essence, they do not offer any

[19]Formally, the information gathered from foreign markets is discounted by the same factor $\frac{m_n}{m_r(m_r+m_n)} < 1$ in all foreign markets.

[20]There is a one-to-one mapping between the productivity of a firm and which market this firm enters in Melitz (2003), whereas this mapping holds on average in the stochastic model of Bernard, Eaton, Jensen and Kortum (2003). However, in both models, provided there is a large number of firms, the fraction of firms exporting to a given market is not stochastic, and only depends on the size and relative labor productivity of this market, on the the variable cost of exporting there, and on the fixed cost of exporting there in Melitz (2003).

[21]See Section 2.5 on page 28 and Appendix B.3 on page 45.

guidance regarding the time-series of entry into foreign markets.

Given the simplicity of the structure of the network that emerges above, I can describe the welfare implications of this model in a variety of economic settings.[22] For instance, if consumers have access to differentiated goods according to the process described above, and if they value the diversity of the goods they consume, then aggregate welfare will increase if m increases. Trivially, a larger m implies that all consumers have access to more goods, which unambiguously increases welfare. On the other hand, an increase in r will reduce the inequality in the access to goods variety across consumers. If goods are sufficiently substitutable, then an increase in r will increase aggregate welfare, whereas if goods are less substitutable, the large welfare gains of those consumers that have access to many goods dominate the welfare losses of those consumers that have access to few goods. A similar argument can be made regarding aggregate productivity if firms, not consumers, have access to differentiated intermediate inputs according to the process described above, and if a firm's productivity increases with the number of differentiated inputs it has access to.

I have developed above a simple model of the formation of an international network of importers and exporters. I use this model to describe the patterns of entry of exporters into different foreign markets. This model delivers a series of empirically testable predictions. First, the model predicts that a stable cross sectional distribution of the the number of countries a firm exports to should arise. The shape of this distribution only depends on the relative importance of random versus network-based meetings. Second, the model predicts that as firms enter more foreign markets, the geographic dispersion of their exports should increase. Third, the model delivers predictions for the dynamics of both firm level and aggregate trade. I empirically test those predictions in the next section.

2 Empirical evidence

In this section, I bring several key testable predictions from the theoretical model to the data. In Section 2.1, I describe the data on firm level exports for French firms, as well as aggregate bilateral trade flows for the rest of the world. In Section 2.2, I test the first main prediction of

[22] See Appendix B.1 for a formal derivation of a model where firms access consumers, and consumers access goods through the process described above, as well as the predictions of this model regarding welfare and the size distribution of firms.

the model regarding the cross-sectional distribution of entry into different foreign markets, derived from Proposition 2. In Section 2.3, I test the second main prediction of the model regarding the geographic dispersion of exports across firms, derived from Proposition 3. In Section 2.4, I test some of the assumptions of the model on the dynamics of exports at the firm and aggregate level. I doing so, I link formally the time-series and the cross-section of firm level exports. Finally, in Section 2.5, I compare the predictions of my model to existing trade theories.

2.1 Data

To bring the model to the data, I use two sources of data.[23] First, I use firm level export data for French exporters, over the period 1986-1992. The data used comes from the same source as the data used by Eaton, Kortum and Kramarz (2010). For the purpose of this paper, I will only use information on French exporters in the years 1986 to1992, not information on domestic sales within France. For each firm, I know the total value (in French Francs) of its exports over a given year, to a given country. There are between 119,000 exporters (in 1988) and 130,000 exporters (in 1987) in my sample. Those firms export to a total of 210 different foreign countries. French exporters export on average to between 3.8 (in 1991) and 4.2 (in 1986) different foreign markets.

In addition to these data on firm level exports for France, I use information on the size of countries, their distance from France and from one another, and aggregate bilateral trade between country pairs. The size of a country is measured as nominal GDP. The data are collected from the Penn World Tables.[24] The distance between France and a particular foreign country is the population weighted geodesic distances between the main cities in both countries. The data come from the CEPII.[25] Finally, I use data on aggregate bilateral trade flows between countries other than France. The data are collected from the NBER.[26]

2.2 Matching the distribution of export destinations

In this section, I test the first main prediction of the model, Proposition 2. The model predicts that the out-degree is the same for all firms, and that the in-degree distribution of a given firm can be described by a mixture of an exponential and a Pareto distribution, where the only two

[23]See Appendix C.1 for a detailed description of the data.
[24]See the description of the data in `http://pwt.econ.upenn.edu/`.
[25]See the description of the data in `http://www.cepii.fr/distance/noticedist_en.pdf`.
[26]See the description of the data in Feenstra et al. (2004).

parameters governing this distribution are r, the ratio of random to network-based links initiated by new firms, and m, the total number of links initiated by new firms.

There are two complications that arise when bringing this prediction to the data.

The first complication stems from the fact that the data on firm level exports do not distinguish between an outward and an inward link. The simplest approach is to assume that trade occurs whenever a link exists, irrespective of whether this link has been originally initiated as an inward or an outward link.[27] Since in the model, the out-degree is the same for all firms, the distribution of the number of contacts will simply be the distribution of in-degrees, shifted to the right by the (constant) number of outward links. The distribution will therefore have the same shape as the distribution in Proposition 2.

The second complication arises from the fact that the data on firm level exports only provide information on the value of exports to a given country, not the number of links between an exporter and firms in that country. The model I developed is meant primarily to explain the extensive margin of international trade, i.e. the patterns of entry of firms into different foreign markets, not the intensive margin of trade, i.e. that value of sales per firm. I will therefore only use information about the number of foreign markets reached by each exporter, not the value of its exports to different countries. Furthermore, I assume that the number of foreign contacts of a French exporter is simply equal to the number of foreign countries it has entered.

The justification for this assumption is the following. First, direct evidence from other countries support this assumption. Eaton, Eslava, Krizan, Kugler and Tybout (2010) are able to match each Colombian exporter to the U.S. with everyone of their contacts in the U.S. They show that 80% of Colombian firms that export to the U.S. have a single contact (buyer) there. Colombia is a smaller market than France, so that Colombian firms may be smaller than French firms on average, but the U.S. is the largest market. If even for this largest market, exporters typically have a single contact, assuming that firms have a single contact per market for other markets is a plausible approximation. Second, the total number of contacts that firms have is small relative to the total number of countries in the world. Using detailed information on the entire input-output linkages between individual US firms, Atalay, Hortaçsu, Roberts and Syverson (2010) find that the average number of suppliers is only marginally above 1, and that even the firm with most suppliers (Walmart in 2005) has only 130 suppliers. In comparison, I have a sample of 210

[27] See Appendix B.1 for a simple model that would support this approach.

different foreign countries. The argument that the number of contacts is small relative to the total number of countries, or that the number of contacts is relatively scarce, is reminiscent of the argument developed in Armenter and Koren (2010). Third, I provide indirect evidence in the Appendix that most exporters seem to have no more than a single contact per market.[28] Even the largest exporters do not seem to have more than a few contacts per market. Fourth, the geographic dispersion of the foreign destinations reached is large, even among firms exporting to a single foreign market ($f_{i,t}$ is dispersed even for $i = t$). Finally, there is no systematic correlation between country size and the geographic distance from France, so that there is no systematic tendency for large exporters to be more or less likely to have several contact per market.[29]

From all the observations above, I can simply think about the world as a relatively fine discrete grid of the continuous theoretical model, and have a probabilistic interpretation of this continuous theoretical model. Since the total number of contacts is small relative to the number of segments in the grid, even for well connected firms, the total number of different segments (foreign countries) reached by exporters is a very good proxy for the total number of foreign contacts.

I discuss this assumption in details in the Appendix, and derive a formal correction for the fact that firms may be have more than one contacts in the markets where they export.[30] The proposed correction does not change the results in any significant way.

To empirically test Proposition 2, I estimate r and $r \times m$ using a non linear least squares regression of the following equation,[31]

$$\ln(\text{fraction of firms exporting to } M \text{ countries}) = \alpha - (2 + r)\ln(M + r \times m) + \epsilon \qquad (2)$$

This equation is directly derived from the p.d.f. associated with the c.d.f. in Proposition 2. I use data on all French exporters and all countries for the year 1992.

Table 1 shows the result from the estimation. The empirical cross sectional distribution of entry into different foreign markets by French exporters suggests that among French exporters, approximately 60% of their foreign contacts are met at random, while 40% are met through

[28]See Appendix C.4 on page 64.

[29]Note that bundling contacts together does not change any of the predictions of the model. If exporters have on average α contacts per country, so that I observe only $\tilde{M} = M/\alpha$, the distribution of the number of countries reached, $\tilde{M}$, is the same as the distribution of the number of contacts, M: $1 - F\left(\tilde{M}\right) = \left(\frac{r \times \tilde{m}}{\tilde{M} + r \times \tilde{m}}\right)^{1+r}$ with $\tilde{m} = m/\alpha$. Therefore, I only need that the average number of contacts per country is approximately independent of the number of countries reached.

[30]See Appendix C.2 on page 54 for this correction, and Appendix C.3 on page 57 for a series of robustness checks.

[31]See Appendix C.3 for a series of robustness checks using different years, different samples of firms, different estimation procedures, and different corrections for the fact that the number of contacts is unobserved.

Table 1: Empirical fit of Proposition2

Dependent Variable: $\ln(f(M))$	
$\widehat{\alpha}$	6.0175***
	(.6489)
$\widehat{r}$	1.5847***
	(.1356)
$\widehat{r \times m}$	7.6768***
	(1.2639)
N. Obs.	124
R^2	.9759

Notes: This table shows the results of the Non Linear Least Square estimation of Equation (2) derived from Proposition 2 for French exporters in 1992. The dependent variable is the log of the fraction of firms that export to M markets. Standard errors are in parentheses. *, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.

network-based meetings. Newborn firms form approximately $m_r = 3$ contacts at random versus $m_n = 1.88$ network-based contacts.[32]

Figure 1 plots the empirical density of the number of foreign markets served by French exporters and the theoretical prediction. The empirical fit of this prediction is remarkable. The fact that both random and network-based meetings coexist explains the curvature of the empirical density in a log-log scale. In unreported regressions, I confirm that this density exhibits significant curvature (in a log-log scale) all along, so that a Pareto distribution would be an imprecise description of the data.

As I will discuss in details in Section 2.5, no existing firm level trade model can generate such a relationship. For instance, it is not enough to introduce Pareto distributed productivity shocks into a Melitz (2003) model, as done in Chaney (2008), to match even approximately the data presented in Figure 1, nor is it enough to assume the convenient Fréchet distribution in the Bernard, Eaton, Jensen and Kortum (2003) model. In addition to the assumption of Pareto or Fréchet distributed productivity shocks, one would need to assume that some specific combination of the (fixed and variable) trade barriers and the sizes and labor productivities of all foreign countries are themselves Pareto distributed. There is no a priori justification for such an ad hoc

[32] $m = \frac{r \times m}{r} \approx \frac{7.68}{1.58} \approx 4.86$; $m_n = \frac{m}{1+r} \approx \frac{4.86}{2.58} \approx 1.88$; $m_r = m - m_n \approx 4.86 - 1.88 \approx 3$.

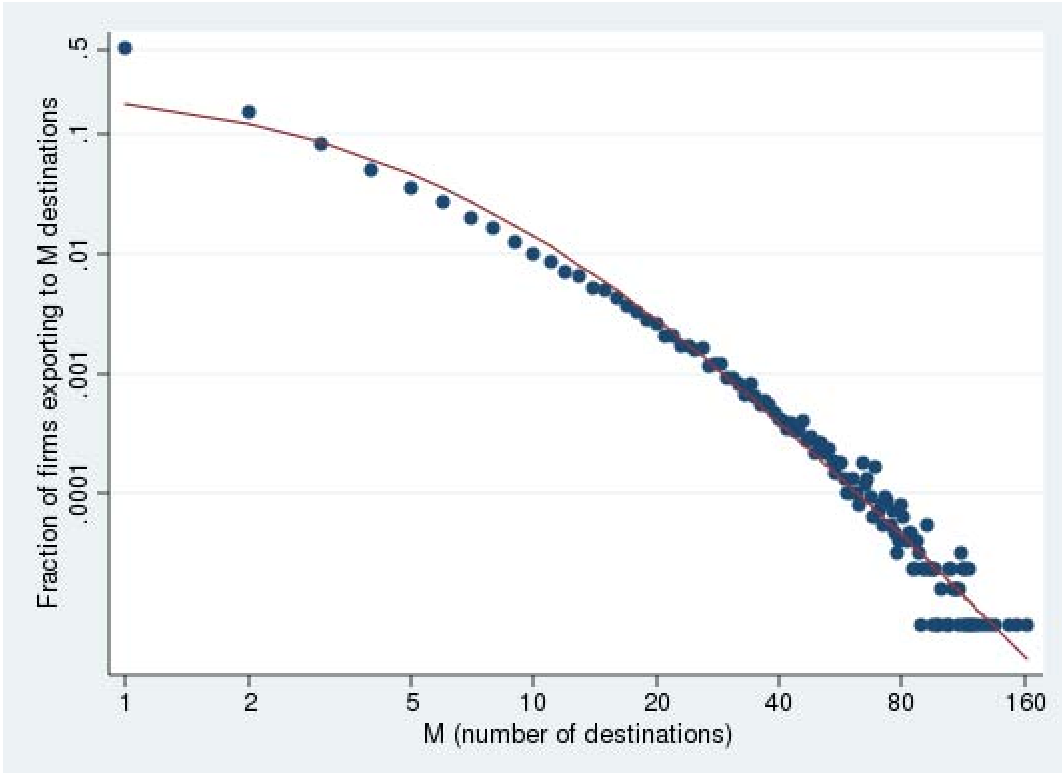

Figure 1: Empirical fit of Proposition 2, $f(M)$ versus M.

Notes: $f(M)$ is the fraction of firms exporting to M destinations; dots: data, all French exporters in 1992; line: theory. $r = 1.58$ (.13) and $r \times m = 7.68$ (1.26) are estimated through non linear least squares.

assumption, nor would such an assumption be empirically plausible.

Armed with an estimate for the relative importance of random versus network-based meetings, I study the geographic dispersion of exports across firms in the next section.

2.3 Matching the geographic dispersion of exports

In this section, I test the second main prediction of the model, Proposition 3. The model predicts how the geographic dispersion of exports increases as firms enter more foreign markets. This relationship only depends on the relative importance of random versus network-based meetings.

Using data on the geographic distribution of exports among firms exporting to exactly M foreign markets, I construct an empirical measure of the geographic dispersion of exports, $\sigma^2(M)$.[33] As a reminder, the theory predicts the following relationship between the geographic dispersion of

[33] I describe and discuss in great details this empirical measure in Appendix C.2 page 56.

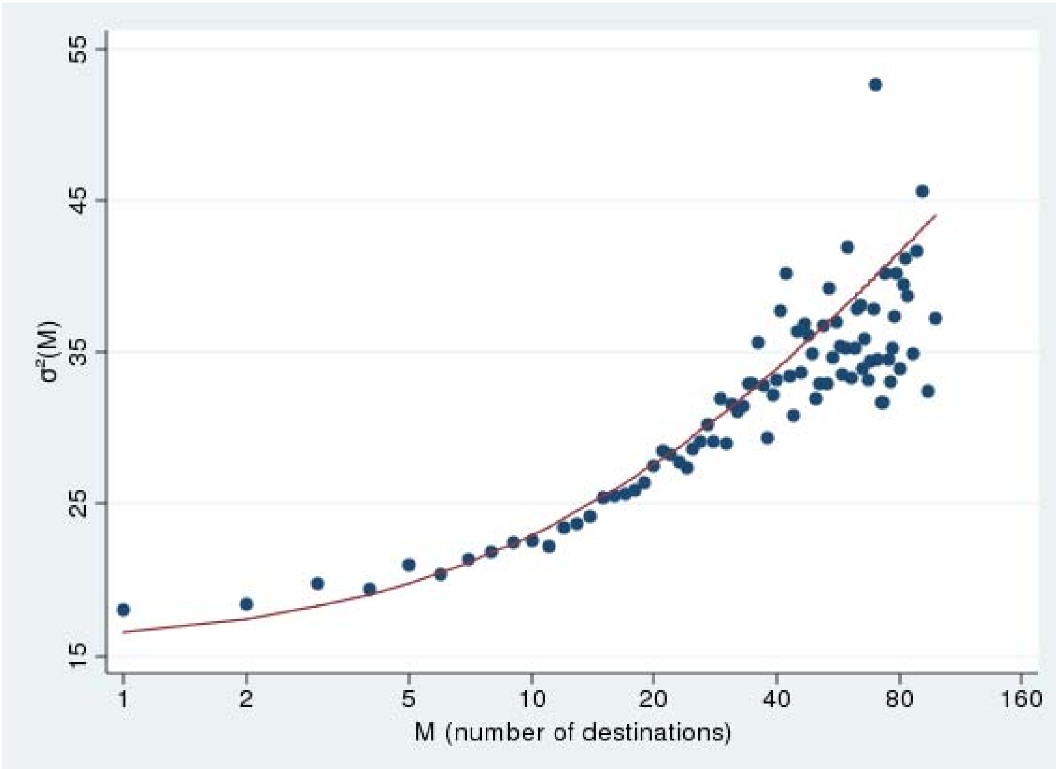

Figure 2: Empirical fit of Proposition 3, $\sigma^2(M)$ versus M.

Notes: $\sigma^2(M)$ is the second moment of the distance from a firm's export destinations, among firms exporting to M destinations; dots: data, all French exporters in 1992; line: theory. $r \times m = 7.68$ is taken from the estimation of Proposition 2, and σ_0^2=15.58 (.10) is estimated through non linear least squares, each point weighted by the square root of the number of observations used to compute $\sigma^2(M)$.

exports and the number of markets a firm is able to enter,

$$\sigma^2(M) = \left(1 + \frac{r \times m}{M}\right) \ln\left(1 + \frac{M}{r \times m}\right) \times \sigma_0^2$$

Using the cross sectional distribution of the number of export destinations across firms, I estimated in the previous section that $r \times m \approx 7.68$. I only need to calibrate σ_0^2, which is not a unit-free measure, to bring the theoretical prediction to the data.[34] I use a non linear least square estimation of the previous equation, and recover $\sigma_0^2 \approx 15.58$ (.10), with an R^2 of 89%.[35]

Figure 2 plots the geographic dispersion of exports, $\sigma^2(M)$, as a function of the total number

[34]As discussed in Section 1.5, I do not need to rely on any specific assumption on the shape of the distribution f_0, except for symmetry and finite variance. The evolution of the geographic dispersion of exports, $\sigma^2(M)$, as a function of Monly depends on the single moment σ_0^2, whichever the shape of distribution f_0 is. The multiplication by σ_0^2 is only needed to match the initial conditions and to scale the units, as $\sigma^2(M)$ is not a unit-free measure.

[35]Each observation is weighted by the precision of its estimation. This precision is given by the square root of the number of observations used to estimate each second moment. See Appendix C.3 on page 57 for a series of robustness checks using different years and different empirical measures of $\sigma^2(M)$.

of foreign countries entered. The empirical fit of the theoretical prediction from Proposition 3 is remarkable. Note that I have no degree of freedom that would allow me to calibrate the shape of this relationship to the data. The shape of this distribution is entirely governed by the theoretical prediction, and by the value for $r \times m$, estimated in the previous section on the cross-sectional distribution of the number of foreign destinations.

The theory connects two distinct empirical observations. First, few firms are able to export to many markets. The proposed explanation is that few firms are able to acquire a large network of contacts. The exact shape of the distribution of the number of foreign contacts is governed by the process of network formation. Second, as firms enter more foreign markets, the geographic dispersion of their exports increases. Here again, the proposed explanation is that through network-based contacts, exporters are able to reach further and further into geographic space. They use their faraway contacts as remote hubs to access even more distant markets. Again, the exact shape of the relation between geographic dispersion and number of export destinations is governed by the process of network formation. The evidence presented in this and the previous section connects these two observations very tightly, showing strong support for the theory.

The next section directly tests some of the underlying assumptions of the model regarding the process of network formation.

2.4 Matching trade dynamics

In this section, I directly test in a reduced form some of the main assumptions of the theoretical model.

By introducing network-based meetings, the model assumes that the more existing foreign contacts a firm already has, the more likely it is to enter an additional market. Note that existing trade models do not provide any clear prior on this correlation. To start with, most existing trade models are static in essence. But even trying to extrapolate the intuitions from existing models does not offer any clear guidance. On the one hand, one may think that export growth is persistent, so that the more existing contacts a firm already has, the more likely it is to acquire new foreign contacts. On the other hand, one may think that small exporters are more likely to expand than large exporters, so that the more contacts a firm already has, the less likely it is to acquire new foreign contacts. By comparison, the model I develop with random and network-based meetings offers a clear prediction that the more contacts a firm has, the more likely it is to acquire new ones.

The next assumption that I test empirically is that exporters benefit from the contacts of their contacts via network-based meetings. The model makes two predictions in that regard.

First, a firm benefits from the contacts of its contacts. In other words, if a firm i has a contact in country c' which itself has a contact in country c, then firm i is more likely to enter country c, everything else being equal. I do not have any direct information on the contacts of the contacts of French exporters. I will instead use data on aggregate trade flows between third countries. The prediction that I test is that if firm i exports to country c' at time $t-1$, and if aggregate exports from country c' to c increases from $t-1$ to t, then firm i is more likely to enter country c at time t, everything else being equal. I use implicitly the fact that if aggregate exports from c' to c increase, some individual firms in c' must acquire new contacts in c.

Second, a firm benefits from the location of its existing contacts. In other words, despite the fact that distance hinders the acquisition of foreign contacts, a firm can use its network of contacts to acquire new contacts in the vicinity of its existing contacts. In that sense, the relevant distance that hampers the acquisition of contacts in country c is not only the distance between France and country c, but also the distance from all the countries c' where a firm has existing contacts and country c.

I test all those three predictions using a Probit regression of different specifications of the following equation,

$$\begin{aligned}\mathbb{I}\{export_{i,c,t}>0\} &= \alpha\times\{\text{N. contacts}_{i,t-1}\} \\ &+\beta_1\times\sum_{c'\in C_{i,t-1}}\frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}}+\beta_2\times\sum_{c'\neq Fr}\frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}} \\ &+\gamma_1\times\frac{1}{|C_{i,t-1}|}\sum_{c'\in C_{i,t-1}}\ln Dist_{c',c}+\gamma_2\times\frac{1}{|\{c':\, c'\neq Fr\}|}\sum_{c'\neq Fr}\ln Dist_{c',c} \\ &+\delta\times\mathbb{I}\{export_{i,c,t-1}>0\}+\text{Controls}_{c,t}+\epsilon_{i,c,t}\end{aligned}\quad (3)$$

where $\mathbb{I}\{export_{i,c,t}>0\}$ is an indicator function equal to 1 if firm i exports to country c in year t, $\{\text{N. contacts}_{i,t-1}\}$ is the total number of foreign markets firm i exports to in year $t-1$, $\frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}}$ is the growth of aggregate exports from c' to c between year $t-1$ and t, $\ln Dist_{c,c'}$ is the natural logarithm of the geographic distance between country c and c', and $C_{i,t-1}$ is the set of countries where firm i exports in year $t-1$. $\sum_{c'\in C_{i,t-1}}\frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}}$ is therefore the growth of aggregate imports by country c from all countries where firm i exports at time $t-1$, and $\frac{1}{|C_{i,t-1}|}\sum_{c'\in C_{i,t-1}}\ln Dist_{c',c}$ the average distance between c and those countries, whereas

$\sum_{c' \neq Fr} \frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}}$ is the growth of aggregate imports by country c from all countries in the world (excluding France), and $\frac{1}{|\{c' : c' \neq Fr\}|} \sum_{c' \neq Fr} \ln Dist_{c',c}$ is the average distance between c and all countries in the world (excluding France). $\epsilon_{i,c,t}$ is a normally distributed error term.

The model assumes that the more existing contacts a firm has, the more likely it is to acquire new contacts. Therefore, I expect, $\alpha > 0$.

The model assumes that firms benefit from the contacts of their contacts. Therefore, I expect that $\beta_1 > 0$. Note that it is possible that some country c may see an increase in its exports from all the world, including France. This would increase the likelihood that any firm enters country c, irrespective of its network of existing contacts. I control for such a direct effect by using information on aggregate imports from country c, and expect $\beta_2 > 0$.

The model assumes that firms benefit from the location of their contacts. Therefore, I expect $\gamma_1 < 0$. Note that it is possible that if country c is more isolated from the rest of the world, in the sense that it is more distant from all other countries, competition in c will be relatively mild, and it will therefore be easier to access c. I control for such a direct effect by using information on the location of country c, and expect $\gamma_2 > 0$.

Table 2 shows the results of the Probit estimation of different specifications of Equation (3), and Table 3 shows the marginal effects of these regressions. In every specification, all coefficients are statistically significant (at the 1% confidence level), and of the expected signs.

More interestingly, the estimates from this panel regression are very close to the predictions of the model calibrated on the cross-sectional distribution of the number of contacts only. From the results in column (1), the estimated increment in the probability of exporting to a given country due to adding an extra contact is .46%. Using the estimate for $r \times m$ from the estimation of the distribution of the number of foreign contacts across firms in Section 2.2, I would predict that this increment is equal to .34%.[36] This is surprisingly close to the actual .46% estimated in the data.

The interpretation of the coefficient on the growth of imports from countries where a firm was already exporting is less obvious. Given that I do not have any direct data on the foreign contacts of French exporters, I can only infer that if aggregate trade increases between two countries, new

[36] From Equation (1) in Section 1.2, I derive that $M_{i.t} - M_{i,t-1} = \gamma m_r \left(1 + \frac{M_{i,t-1}}{r \times m}\right)$. So that adding one contact increases the growth in the number of contacts by $\gamma m_r \left(1 + \frac{M_{i,t-1}+1}{r \times m} - 1 - \frac{M_{i,t-1}}{r \times m}\right) = \frac{\gamma m_r}{r \times m}$. Given that the average probability of entering a new country in the sample is 4.7%, given that the average number of contacts in the sample is 6, and given the estimate $r \times m \approx 7.68$ from estimating Equation (2) on the cross-sectional distribution of the number of foreign contacts, I predict that the increment in the probability of entering a new country stemming from moving from 6 to 7 contacts is given by $\frac{\gamma m_r}{r \times m} \approx \frac{4.7\%}{7.68+6} \approx .34\%$. This is close to the actual .44% in the data.

Table 2: Existing number of contacts and trade between third countries predict entry (PROBIT)

	Dependent Variable: $\mathbb{I}\{export_{i,c,t} > 0\}$					
	(1)	(2)	(3)	(4)	(5)	(6)
N. contacts$_{i,t-1}$	0.0439***	0.0564***	0.0327***	0.0334***	0.0351***	0.0352***
	(0.00025)	(0.00036)	(0.00027)	(0.00028)	(0.00031)	(0.00031)
$\sum_{c' \in C_{i,t-1}} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t-1}}$				0.0376***		0.0357***
				(0.00239)		(0.00246)
$\sum_{c' \neq Fr} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t-1}}$				0.0803***		0.0545***
				(0.00404)		(0.00409)
$\frac{1}{\lvert C_{i,t-1}\rvert} \sum_{c' \in C_{i,t-1}} \ln Dist_{c',c}$					-0.1743***	-0.1381***
					(0.00264)	(0.00276)
$\frac{1}{\lvert\{c' : c' \neq Fr\}\rvert} \sum_{c' \neq Fr} \ln Dist_{c',c}$					0.4893***	0.4883***
					(0.00664)	(0.00685)
$\ln GDP_{c,t}$		0.2400***	0.1484***	0.1537***	0.1286***	0.1316***
		(0.00142)	(0.00089)	(0.00089)	(0.00070)	(0.00071)
$\ln Dist_{Fr,c}$		-0.6056***	-0.3802***	-0.3812***	-0.3689***	-0.3965***
		(0.00158)	(0.00109)	(0.00111)	(0.00274)	(0.00277)
$\mathbb{I}\{export_{i,c,t-1} > 0\}$			2.0798***	2.0539***	1.9963***	1.9951***
			(0.00342)	(0.00340)	(0.00349)	(0.00351)
Constant	-1.9239***	0.0968***	-0.8799***	-0.9545***	-3.5624***	-3.6768***
	(0.00219)	(0.01915)	(0.01258)	(0.01267)	(0.04907)	(0.05014)
N. obs		24,964,110			24,664,895	
N. firms		40,395			40,330	
N. years		6			6	
N. destinations		103			103	
Pseudo-R^2	0.1253	0.3598	0.5612	0.5588	0.564	0.5604

Notes: This table shows the results of the PROBIT estimation of Equation (3) for a panel of all French exporters between 1986 and 1992. The dependent variable is an indicator function that takes the value 1 if firm i is exporting to country c at time t. The description of the explanatory variables is given along with Equation (3) on page 23. Standards errors are clustered at the firm level. *, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.

Table 3: Existing number of contacts and trade between third countries predict entry (PROBIT: marginal effects)

Dep. Var.: $\mathbb{I}\{export_{i,c,t} > 0\}$	dy/dx					
	(1)	(2)	(3)	(4)	(5)	(6)
N. contacts$_{i,t-1}$	0.0046***	0.0027***	0.0013***	0.0013***	0.0014***	0.0014***
	(0.00002)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
$\sum_{c' \in C_{i,t-1}} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t-1}}$				0.0015***		0.0014***
				(0.00010)		(0.00010)
$\sum_{c' \neq Fr} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t-1}}$				0.0032***		0.0022***
				(0.00016)		(0.00016)
$\frac{1}{\lvert C_{i,t-1}\rvert} \sum_{c' \in C_{i,t-1}} \ln Dist_{c',c}$					-0.0069***	-0.0055***
					(0.00011)	(0.00011)
$\frac{1}{\lvert\{c' : c' \neq Fr\}\rvert} \sum_{c' \neq Fr} \ln Dist_{c',c}$					0.0193***	0.0193***
					(0.00025)	(0.00026)
$\ln GDP_{c,t}$		0.0114***	0.0059***	0.0062***	0.0051***	0.0052***
		(0.00005)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
$\ln Dist_{Fr,c}$		-0.0287***	-0.0152***	-0.0153***	-0.0145***	-0.0157***
		(0.00013)	(0.00006)	(0.00006)	(0.00010)	(0.00011)
$\mathbb{I}\{export_{i,c,t-1} > 0\}$			0.4094***	0.4008***	0.3762***	0.3769***
			(0.00134)	(0.00134)	(0.00135)	(0.00137)

Notes: This table shows the marginal effects for the PROBIT estimation of Equation (3) presented in Table 2. The marginal effect is calculated as dy/dx at the average value of each x in the sample. dy/dx is for a discrete change from 0 to 1 when x is a dummy variable. Standards errors are clustered at the firm level. *, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.

contacts must have been created between those countries. However, even with this caveat in mind, the results are qualitatively and quantitatively in line with the theory. Using the results in column (6), the estimated increment in the probability of entering a given foreign country following an increase in this country's imports is positive, and roughly 57% larger than the increment coming from an increase in the imports from the countries where a firm is actually exporting. A rough interpretation of this result suggests that random meetings are approximately 57% larger than network-based meetings, or that $r = \frac{m_r}{m_n} \approx \frac{.0022}{.0014} \approx 1.57$. This is surprisingly close to $r \approx 1.58$ estimated from the cross-sectional distribution of the number of foreign contacts across firms in Section 2.2.

Taking the specific functional form of the model even more seriously, I can structurally estimate the law of motion for the number of contacts implied by Equation (1). Integrating the in-degree distribution $f_{i,t}$ over $\mathbb{R}$, I get the following law of motion for $M_{i,t}$, the number of foreign contacts of firm i at time t,

$$M_{i,t+1} - M_{i,t} = \frac{\gamma}{1+r} M_{i,t} + \gamma m_r$$

Adding a series of controls, including on the growth rate of domestic sales of those firms to control for the growth trajectory a firm follows, does not affect those results substantially. A simple OLS estimation of the relationship above gives $\widehat{\frac{\gamma}{1+r}} = .165$ (.00040) and $\widehat{\gamma m_r} = .876$ (.0048). This implies $r \times m \approx 5.5$, which is surprisingly close to $r \times m \approx 7.68$ estimated from the cross-sectional distribution of the number of foreign contacts across firms in Section 2.2.[37]

I have presented in this section direct evidence in support of the assumptions of the proposed theoretical model based on the panel dimension of firm level French exports. The proposed theory not only delivers predictions for the cross section of firm level exports that fit the data remarkably well, but I offer direct time-series evidence in support of the proposed mechanism that generates these predictions. The empirical evidence in support of the predictions of the model are in line both qualitatively and quantitatively with the empirical evidence in support of the assumptions of the model. Both sets of evidence come from two distinct sources of data. The prediction of the model are tested using cross-sectional evidence, whereas the assumptions are tested using time-series evidence.

[37]Depending on the specifications, the time-series estimate of $r \times m$ ranges between 2 (without any controls) and 5.68. Note that the cross-sectional maximum likelihood estimate of $r \times m \approx 6.32$ is somewhat lower than the non linear least squares estimate, 7.68, presented above. See Appendix C.3 for a series of robustness checks on the cross-sectional and the time-series estimations.

In the next section, I compare the empirical success of the predictions of my model to those of existing trade models.

2.5 Comparison with existing trade theories

It is important to note that none of the existing firm level trade models can match the empirical regularities described above. In this section, I discuss the predictions of the two main existing firm level trade models, Melitz (2003) and Bernard, Eaton, Jensen and Kortum (2003). I derive those predictions formally in the Appendix.[38]

In the original Melitz (2003) model where all trade barriers are symmetric, any firm that exports export to all foreign markets. This is obviously an artifact of the counter-factual assumption that all trade barriers are perfectly symmetric. Chaney (2008) offers a simple extension of Melitz (2003) with asymmetric country sizes and fixed and variable trade barriers. In this model, from the point of view of a given exporting country, say France, there is a strict hierarchy of foreign markets. This means that markets can be strictly ordered in a decreasing level of accessibility, so that if a French firm exports to market m, it will necessarily export to all markets $n \leq m$. Therefore, the fraction of firms that export to exactly M markets is simply the fraction of firms that have a productivity between the productivity threshold for exporting to market M and the threshold for exporting to market $M+1$. Even if productivities are distributed Pareto, the fraction of firms that export to exactly M markets can take any value, depending on the thresholds for exporting to country M and $M+1$. Even if country sizes are themselves Pareto distributed, and if fixed export costs are log-proportional to country size, there is no reason to assume that variable trade barriers are themselves log-proportional to country size. The fraction of firms that export to exactly M markets does not even have to be decreasing in M. Note that this result does not depend on the strict hierarchy of foreign markets, but on the fact that the distances between productivity thresholds do not have to follow any systematic patterns. Adding noise to the fixed or variable export costs that different firms face will not improve the ability of the Melitz model to match the observed cross-sectional distribution of the number of foreign markets reached by French exporters.

In the stochastic model of Bernard, Eaton, Jensen and Kortum (2003), there is no strict hierarchy in the accessibility of foreign markets. A given exporter, even if it is has a low productivity, may still export to many foreign countries, if this exporter is lucky enough that foreign competitors

[38] See Appendix B.3 on page 45.

happen to have an even lower productivity. However, the structure of country sizes, relative productivities and labor costs across countries, and bilateral trade barriers between countries imposes a severe restriction on the cross-sectional distribution of the number of foreign markets entered. For a large number of firms, or for the continuous limit developed in the model, there is no uncertainty neither in the fraction of firms entering any given market, or the distribution of the number of markets entered. This distribution depends on the specific trade barriers and country characteristics. Even with the assumed ad hoc and convenient Fréchet distribution of productivities, there is no reason why any particular distribution should arise. As in the Melitz model, the fraction of firms that export to exactly M markets does not even have to be decreasing in M. The following argument makes this point clear. In the limit of infinitely large trade barriers, all firms only sell in their domestic market, so that no firm sells to any $M > 0$ foreign markets. In the other extreme of perfectly free trade, all firms that sell domestically also export to all countries in the world. So whereas the fraction of firms that export to all foreign countries in the world is monotonically decreasing from 1 to 0 with the level of trade barriers, the fraction of firms exporting to any other number of foreign countries is non monotone. The fraction of firms exporting to exactly M markets can be made arbitrarily small or close to 1 by simply varying bilateral trade barriers.

Both models would correctly predict that the geographic dispersion of exports increases with the number of countries a firm is able to enter. The reason is the following. In both models, there is a hierarchy of the different foreign markets potentially accessible to exporters. Foreign markets can be ranked as a function of how easily accessible they are to exporters. The ranking is strict in Melitz (2003), and holds on average in the stochastic model of Bernard, Eaton, Jensen and Kortum (2003). Therefore, mechanically, as firms enter more markets, they also enter less accessible markets. Since less accessible markets tend to be geographically more remote, in both models, the geographic dispersion of the markets entered by a firm mechanically increases with the number of markets this firm enters. However, neither model makes any robust prediction regarding the specific shape of relationship between the number of markets accessed and the geographic dispersion of those markets. This relationship is shaped by the distribution of the exogenous export barriers, for which no restriction exists a priori. Moreover, the cross-sectional distribution of the number of foreign markets accessed does not provide any relevant information on this relationship, unless some additional assumptions are made regarding how geographic distance from France affects both trade barriers and country sizes.

Finally, both models being static in essence, they do not generate any predictions regarding the dynamic evolution of firm level trade flows. The proposed model on the other hand links theoretically the cross-sectional distribution of the number of foreign markets accessed, the relationship between the number of foreign markets accessed and the geographic dispersion of those markets, and the time series evolution of the entry of individual firms into foreign markets. As I have shown above, all three predictions are supported by the data in a consistent way.

It is also important to note that the proposed model does not rely on any particular functional form assumption. Most importantly, I do not need to impose any particular restriction on the geographic structure of the export costs. This freedom from any functional form assumption allows me to match some key empirical regularities by calibrating only three parameters.[39] Given the intrinsic complexity of a large scale network, I have to introduce a series of simplifying assumptions for the formation of this international network of importers and exporters (mean-field approximation, symmetry, randomness of the link formation, growth of the in-degree and not of the out-degree). However, beyond those simplifying assumptions, I do not require any functional form restriction.

Conclusion

I have developed a theoretical model of the dynamic formation of an international network of importers and exporters. Firms can only export in countries where they have a contact. I assume that firms acquire contacts either at random, or via their existing network of contacts. This dynamic model generates a stable network structure. The model makes precise predictions about the cross-sectional distribution of the number of foreign contacts, the cross-sectional distribution of the geographic dispersion of foreign contacts, and the dynamics of entry of individual firms into foreign markets. All theoretical predictions are tightly supported by the data on firm level exports from France. Firms acquire 60% more contacts at random than they do via network-based meetings.

This model and the empirical findings that support it suggest several extensions and generalizations. First, the emergence of a stable distribution of entrants into different foreign markets, and the fact that firms that export to more countries are less affected by geographic distance, may generate aggregate trade flows that follow the so called gravity equations. This may provide an

[39] I only estimate r, the ratio of random to network-based meetings; m, the number of contacts initiated by new exporters; and σ_0^2, the geographic dispersion of foreign contacts for new exporters

explanation for the stable role that geographic distance plays in explaining bilateral trade flows. Second, I have only studied a simple symmetric case, and described its steady state properties. Large shock to this dynamic system would generate non trivial transitional dynamics. For example, a large disruption of trade linkages may have a long lasting impact on the world geography of trade, since rebuilding contacts is a lengthy process. Third, whereas I have only sketched the welfare implications of a simple economic model that would support the proposed dynamics, the structure of the network lends itself to further analysis of the welfare gains from trade. Jackson and Rogers (2007) propose tools to analyze the welfare implications of different network structures, and the model developed in this paper adds a geographic dimension to their space-less model. The robust predictions of the model regarding the geographic distribution of exports may allow for precise statements on the welfare gains from trade. I leave these questions for future research.

References

ARMENTER, Roc and Miklós KOREN. 2009. "Economies of Scale and the Size of Exporters," Central European University, *mimeo.*

ARMENTER, Roc and Miklós KOREN. 2010. "A Balls-and-Bins Model of Trade," Central European University, *mimeo.*

ATALAY, Enghin, Ali HORTAÇSU, James W. ROBERTS, and Chad SYVERSON. 2010. "The Structure and Dynamics of the U.S. Inter-firm Network," University of Chicago, *mimeo.*

BERNARD, Andrew B., Jonathan EATON, J. Bradford JENSEN, and Samuel KORTUM. 2003. "Plants and Productivity in International Trade," *American Economic Review*, 93(4): 1268-1290.

CHANEY, Thomas. 2008. "Distorted Gravity: the Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4): 1707-21.

EATON, Jonathan, Marcela ESLAVA, C.J. KRIZAN, Maurice KUGLER, and James TYBOUT. 2010. "A Search and Learning Model of Export Dynamics," Penn State University, *mimeo.*

EATON, Jonathan, Samuel KORTUM, and Francis KRAMARZ. 2010. "An Anatomy of International Trade: Evidence from French Firms," University of Chicago, *mimeo.*

FEENSTRA, Robert C., Robert E. LIPSEY, Haiyan DENG, Alyson C. MA, and Hengyong MO. 2004. "World Trade Flows: 1962-2000," NBER WP 11040.

JACKSON, Matthew O., and Brian W. ROGERS. 2007. "Meeting Strangers and Friends of Friends: How Random Are Social Networks?" *American Economic Review*, 97(3): 890-915.

KRUGMAN, Paul. 1980. "Scale Economies, Product Differentiation, and the Patterns of Trade," *American Economic Review*, 70(5): 950-59.

MELITZ, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocation and Aggregate Industry Productivity," *Econometrica*, 71(6): 1695-1725.

Appendix

A Mathematical proofs

In this section, I give the detailed mathematical proofs of the various propositions, lemmas and claims I have presented in Section 1.

Proposition 1 (reminded) *The in-degree distribution of firm i at time t, $f_{i,t}$, is given by,*

$$f_{i,t} = \underbrace{\left(\delta + \gamma \frac{m_n}{m_r + m_n} f_0\right) * \ldots * \left(\delta + \gamma \frac{m_n}{m_r + m_n} f_0\right)}_{t-i \text{ times}} * \frac{m_r}{m_n}\left(m_r + m_n\right)\delta - \frac{m_r}{m_n}\left(m_r + m_n\right)\delta$$

where δ is the Dirac delta function, $*$ is the convolution product, γ is the growth rate of the population, f_0 is the geographic distribution of contacts for newborn firms, and m_r and m_n are respectively the number of random and network-based meetings of newly born firms.

Proof. Taking a Fourier transform of Eq. (1), and using the convolution theorem which states that the Fourier transform of the convolution of distributions is the product of the Fourier transform of those distributions, I get,

$$\mathcal{F}\left[f_{i,t+1}\right] = \gamma m_r \mathcal{F}\left[f_0\right] + \gamma \frac{m_n}{m_r + m_n} \mathcal{F}\left[f_0\right] \times \mathcal{F}\left[f_{i,t}\right] + \mathcal{F}\left[f_{i,t}\right]$$

Rearranging, this gives a simple first order linear recursive equation for $\mathcal{F}\left[f_{i,t}\right]$,

$$\mathcal{F}\left[f_{i,t+1}\right] = \gamma m_r \mathcal{F}\left[f_0\right] + \left(1 + \gamma \frac{m_n}{m_r + m_n} \mathcal{F}\left[f_0\right]\right) \times \mathcal{F}\left[f_{i,t}\right]$$

for $t > i$, with the initial condition at $t = i$,

$$\mathcal{F}\left[f_{i,i}\right] = M_0 \mathcal{F}\left[f_0\right]$$

This recurrence admits a simple solution,

$$\mathcal{F}\left[f_{i,t}\right] = \left(1 + \gamma \frac{m_n}{m_r + m_n} \mathcal{F}\left[f_0\right]\right)^{t-i} \times \left(\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0 \mathcal{F}\left[f_0\right]\right) - \frac{m_r}{m_n}\left(m_r + m_n\right)$$

Taking the inverse Fourier transform of this equation, and noting that $\mathcal{F}^{-1}\left[1\right] = \delta$, where δ is the Dirac delta function, I solve for the in-degree distribution of firm i at time t,

$$f_{i,t} = \underbrace{\left(\delta + \gamma \frac{m_n}{m_r + m_n} f_0\right) * \ldots * \left(\delta + \gamma \frac{m_n}{m_r + m_n} f_0\right)}_{t-i \text{ times}} * \left(\frac{m_r}{m_n}\left(m_r + m_n\right)\delta + M_0 f_0\right) - \frac{m_r}{m_n}\left(m_r + m_n\right)\delta$$

Taking $M_0 = 0$, I get the proposed expression for $f_{i,t}$.

Note that in the special case where distance represents an insurmountable barrier to the formation of trade linkages, which corresponds formally the case $f_0 = \delta$, the model collapses to the space-less model of Jackson and Rogers (2007) where each location is an isolated island that behaves exactly like in Jackson and Rogers,

$$f_0 = \delta \Rightarrow f_{i,t} = \left\{ \left(1 + \gamma \frac{m_n}{m_r + m_n}\right)^{t-i} \times \left(\frac{m_r}{m_n}(m_r + m_n) + M_0\right) - \frac{m_r}{m_n}(m_r + m_n) \right\} \times \delta$$

■

Lemma 1 (reminded) *The total mass of contacts at time t of a firm born in i, i.e. the in-degree of this firm, $M_{i,t}$, is given by,*

$$M_{i,t} = \left(1 + \gamma \frac{m_n}{m_r + m_n}\right)^{t-i} \times \frac{m_r}{m_n}(m_r + m_n) - \frac{m_r}{m_n}(m_r + m_n)$$

where γ is the growth rate of the population, and m_r and m_n are respectively the number of random and network-based meetings of newly born firms.

Asymptotically, the in-degree grows at a constant rate. This growth rate is highest (equal to $1 + \gamma$) when most links are network-based ($\frac{m_n}{m_r + m_n} \approx 1$), and lowest (equal to 1) when most link are random ($\frac{m_n}{m_r + m_n} \approx 0$).

Proof. The in-degree of firm i at time t, $M_{i,t}$, is defined as the integral over $\mathbb{R}$ of its in-degree distribution, $f_{i,t}$,

$$M_{i,t} \equiv \int_{\mathbb{R}} f_{i,t}(x)\, dx$$

Using the expression for the in-degree distribution from Proposition 1, and integrating, I get,

$$M_{i,t} = \int_{\mathbb{R}} \underbrace{\left\{ \left(\delta(x) + \gamma \frac{m_n}{m_r + m_n} f_0(x)\right) * \ldots * \left(\delta(x) + \gamma \frac{m_n}{m_r + m_n} f_0(x)\right)\right.}_{t-i \text{ times}}$$
$$\left. * \left(\frac{m_r}{m_n}(m_r + m_n)\delta(x) + M_0 f_0(x)\right) - \frac{m_r}{m_n}(m_r + m_n)\delta(x) \right\} dx$$

Using the known result that the integral of a convolution of distributions is the product of the integral of those distributions, I get,

$$M_{i,t} = \left(\int_{\mathbb{R}} \delta(x)\, dx + \gamma \frac{m_n}{m_r + m_n} \int_{\mathbb{R}} f_0(x)\, dx\right)^{t-i}$$
$$\times \left(\frac{m_r}{m_n}(m_r + m_n) \int_{\mathbb{R}} \delta(x)\, dx + M_0 \int_{\mathbb{R}} f_0(x)\, dx\right) - \frac{m_r}{m_n}(m_r + m_n) \int_{\mathbb{R}} \delta(x)\, dx$$

Since both the Dirac δ function and f_0 are well defined probability density functions, so that $\int_{\mathbb{R}} \delta(x)\,dx = \int_{\mathbb{R}} f_0(x)\,dx = 1$, I get,

$$M_{i,t} = \left(1 + \gamma \frac{m_n}{m_r + m_n}\right)^{t-i} \times \left(\frac{m_r}{m_n}(m_r + m_n) + M_0\right) - \frac{m_r}{m_n}(m_r + m_n)$$

Taking $M_0 = 0$, I get the proposed expression for $M_{i,t}$.

Asymptotically, the in-degree expands at a constant growth rate equal to,

$$\lim_{t-i \to +\infty} \frac{M_{i,t+1} - M_{i,t}}{M_{i,t}} = \left(1 + \gamma \frac{m_n}{m_r + m_n}\right)$$

The growth rate of the in-degree is highest when most links are network-based, and lowest when most links are random,

$$\begin{cases} \lim\limits_{t-i \to +\infty} \frac{M_{i,t+1} - M_{i,t}}{M_{i,t}} = (1+\gamma) \text{ if } \frac{m_n}{m_r+m_n} = 1 \\ \lim\limits_{t-i \to +\infty} \frac{M_{i,t+1} - M_{i,t}}{M_{i,t}} = 1 \text{ if } \frac{m_n}{m_r+m_n} = 0 \end{cases}$$

■

Proposition 2 (reminded) *For a population growth rate γ small, the distribution of the number of contacts, M, across individual firms in any arbitrary set of locations is given by the cumulative distribution function,*

$$F(M) = 1 - \left(\frac{r \times m}{M + r \times m}\right)^{1+r}$$

where $r = m_r/m_n$ is the ratio of random versus network based meetings, and $m = (m_r + m_n)$is the total number of contacts made by newly born firms.

Proof. At time t, in any give location, and therefore in the union of any given location, the fraction of firms with in-degree above M is simply the fraction of firms born before $i(M,t)$, defined as $M_{i(M,t),t} = M$. Using the expression for the in-degree at time t of a firm born at date i from Lemma 1, and assuming away integer constraints (i should in principle be an integer[40]), I can solve for $i(M,t)$,

$$\left(1 + \gamma \frac{m_n}{m_r + m_n}\right)^{t-i(M,t)} \times \left(\frac{m_r}{m_n}(m_r + m_n) + M_0\right) - \frac{m_r}{m_n}(m_r + m_n) = M$$

$$\Rightarrow i(M,t) - t = \ln\left(\frac{M_0 + \frac{m_r}{m_n}(m_r + m_n)}{M + \frac{m_r}{m_n}(m_r + m_n)}\right) \Big/ \ln\left(1 + \gamma \frac{m_n}{m_r + m_n}\right)$$

[40]Note that I do not make any continuous approximation of the discrete model. The proposed formulas are exactly correct when i is an integer. Those formulas simply extrapolate to non integer values for i.

Given the exponential growth of the population, the fraction of firms born before date $i(M,t)$ is given by,

$$\frac{N_{i(M,t)}}{N_t} = (1+\gamma)^{i(M,t)-t}$$

Using the previous expression for $i(M,t)-t$, I get the fraction of firms born before date $i(M,t)$,

$$\frac{N_{i(M,t)}}{N_t} = \left(\frac{M_0 + \frac{m_r}{m_n}(m_r+m_n)}{M + \frac{m_r}{m_n}(m_r+m_n)}\right)^{\ln(1+\gamma)/\ln\left(1+\gamma\frac{m_n}{m_r+m_n}\right)}$$

Using the following approximation for γ small,

$$\lim_{\gamma\to 0}\frac{\ln(1+\gamma)}{\ln\left(1+\gamma\frac{m_n}{m_r+m_n}\right)} = (1+m_r/m_n)$$

and given that the fraction of firms born before date $i(M,t)$ is the counter-cumulative distribution of in-degrees, $\frac{N_{i(M,t)}}{N_t} = 1 - F_t(M)$, I get the proposed cumulative distribution function for the in-degree of firms,

$$F(M) = 1 - \left(\frac{M_0 + \frac{m_r}{m_n}(m_r+m_n)}{M + \frac{m_r}{m_n}(m_r+m_n)}\right)^{1+m_r/m_n}$$

Taking $M_0 = 0$, I get the proposed expression for $F(M)$.[41]

Note the following polar cases, when $m_r/m_n \to +\infty$ or $m_r/m_n \to 0$. Given the above expression for $F(M)$, for any M,

$$\lim_{m_r/m_n\to+\infty} F(M) = 1 - \exp\left(\frac{M_0 - M}{m_r+m_n}\right)$$

so that the in-degree distribution converges to an exponential distribution when almost all contacts are made at random. Similarly, for any M,

$$\lim_{m_r/m_n\to 0} F(M) = 1 - \frac{M_0}{M}$$

so that the in-degree distribution converges to a Zipf law when almost all contacts are network-based. Alternatively, it would have been easy to solve for $F(M)$ in both polar cases using Proposition 1 directly. ■

[41] Note that formally, the c.d.f. of the in-degree distribution within the population is a step function that corresponds to the true discrete distribution. However, the values of $F(M)$ for the M's corresponding to the discrete ages among the population at any point in time are given by the exact formula for $F(M)$ above. Whereas the function $F(M)$ is time invariant, the location of the steps for the M's evolves through time. All the formulas above hold exactly at any point in time, without any continuous approximation.

Lemma 2 (reminded) *The geographic dispersion of a firm's contacts increases as a firm ages. Formally,* $\sigma_{i,t}^2$*, the variance of the distance from a firm's contacts is asymptotically proportional to a firm's age,*

$$\sigma_{i,t}^2 \underset{t-i\to\infty}{\approx} (t-i)\frac{\gamma}{1+r+\gamma}\sigma_0^2$$

at time t*, for a firm born at date* i*,where* γ *is the growth rate of the population,* $r = m_r/m_n$ *is the ratio of random versus network-based meetings, and* σ_0^2 *is the variance of the distance from a firm's random contacts.*

Proof. Plugging the expression for the in-degree distribution of a firm from Proposition 1 into the definition of the variance from a firm's contacts, $\sigma_{i,t}^2$,

$$M_{i,t}\sigma_{i,t}^2 = \int_{\mathbb{R}} x^2 \Bigg\{ \underbrace{\left(\delta(x) + \gamma\frac{m_n}{m_r+m_n}f_0(x)\right) * \ldots * \left(\delta(x) + \gamma\frac{m_n}{m_r+m_n}f_0(x)\right)}_{t-i \text{ times}}$$
$$* \left(\frac{m_r}{m_n}(m_r+m_n)\delta(x) + M_0 f_0(x)\right) - \frac{m_r}{m_n}(m_r+m_n)\delta(x)\Bigg\} dx$$

Rearranging to get a convolution of well defined probability density functions (each summing up to 1), I get,

$$M_{i,t}\sigma_{i,t}^2 = \left(1+\gamma\frac{m_n}{m_r+m_n}\right)^{t-i}\left(\frac{m_r}{m_n}(m_r+m_n)+M_0\right)$$
$$\times \int_{\mathbb{R}} x^2 \Bigg\{ \underbrace{\frac{\delta(x)+\gamma\frac{m_n}{m_r+m_n}f_0(x)}{1+\gamma\frac{m_n}{m_r+m_n}} * \ldots * \frac{\delta(x)+\gamma\frac{m_n}{m_r+m_n}f_0(x)}{1+\gamma\frac{m_n}{m_r+m_n}}}_{t-i \text{ times}}$$
$$* \frac{\frac{m_r}{m_n}(m_r+m_n)\delta(x)+M_0 f_0(x)}{\frac{m_r}{m_n}(m_r+m_n)+M_0}\Bigg\} dx - \int_{\mathbb{R}} x^2 \frac{m_r}{m_n}(m_r+m_n)\delta(x)\,dx$$

Noting the fact that the Dirac delta function has zero variance, I get the following intermediate result,

$$\int_{\mathbb{R}} x^2 (A\delta(x) + Bf_0(x))\,dx = B\sigma_0^2 \text{ for any } (A,B)\in\mathbb{R}^2$$

From this intermediate result, I get,

$$\int_{\mathbb{R}} x^2 \frac{\delta(x)+\gamma\frac{m_n}{m_r+m_n}f_0(x)}{1+\gamma\frac{m_n}{m_r+m_n}}dx = \frac{\gamma\frac{m_n}{m_r+m_n}}{1+\gamma\frac{m_n}{m_r+m_n}}\sigma_0^2$$
$$\int_{\mathbb{R}} \frac{\frac{m_r}{m_n}(m_r+m_n)\delta(x)+M_0 f_0(x)}{\frac{m_r}{m_n}(m_r+m_n)+M_0}dx = \frac{M_0}{\frac{m_r}{m_n}(m_r+m_n)+M_0}\sigma_0^2$$
$$\text{and } \int_{\mathbb{R}} x^2\frac{m_r}{m_n}(m_r+m_n)\delta(x)\,dx = 0$$

Next, I use the fact that if two independent random variables X and Y have respective p.d.f.'s f and g, then $X+Y$ has a p.d.f. equal to the convolution $f * g$. Obviously, if X and Y are independent, the variance of $X+Y$ is equal to the sum of the variances of X and Y. This allows me to get the following result,

$$M_{i,t}\sigma_{i,t}^2 = \left(1+\gamma\frac{m_n}{m_r+m_n}\right)^{t-i}\left(\frac{m_r}{m_n}(m_r+m_n)+M_0\right)\left[(t-i)\frac{\gamma\frac{m_n}{m_r+m_n}}{1+\gamma\frac{m_n}{m_r+m_n}}+\frac{M_0}{\frac{m_r}{m_n}(m_r+m_n)+M_0}\right]\sigma_0^2$$

Plugging in the expression for the total in-degree of a firm from Lemma 1, I get,

$$\sigma_{i,t}^2 = \frac{\left(1+\gamma\frac{m_n}{m_r+m_n}\right)^{t-i}\left(\frac{m_r}{m_n}(m_r+m_n)+M_0\right)}{\left(1+\gamma\frac{m_n}{m_r+m_n}\right)^{t-i}\left(\frac{m_r}{m_n}(m_r+m_n)+M_0\right)-\frac{m_r}{m_n}(m_r+m_n)} \left[(t-i)\frac{\gamma\frac{m_n}{m_r+m_n}}{1+\gamma\frac{m_n}{m_r+m_n}}+\frac{M_0}{\frac{m_r}{m_n}(m_r+m_n)+M_0}\right]\sigma_0^2$$

As firms age, $t-i\to\infty$, the first multiplicative term converges to 1, the second additive term in the square bracket becomes negligible, and I get the proposed approximation,

$$\sigma_{i,t}^2 \underset{t-i\to\infty}{\approx} (t-i)\frac{\gamma\frac{m_n}{m_r+m_n}}{1+\gamma\frac{m_n}{m_r+m_n}}\sigma_0^2$$

■

Proposition 3 (reminded) *The geographic dispersion of a firm's contacts increases with the firm's in-degree. For a population growth rate γ small, the variance of the distance from a firm with M contacts, $\sigma^2(M)$, is given by,*

$$\sigma^2(M) = \left(1+\frac{r\times m}{M}\right)\ln\left(1+\frac{M}{r\times m}\right)\times\sigma_0^2$$

where σ_0^2 is the variance of the distance from a newly born firm's contacts, $r = m_r/m_n$ is the ratio of random versus network based meetings, and $m = (m_r+m_n)$ is the total number of contacts made by newly born firms.

Proof. From the proof of Proposition 2 I get a relation between $t-i$ and $M_{i,t}$,

$$t-i = \ln\left(\frac{M_{i,t}+\frac{m_r}{m_n}(m_r+m_n)}{M_0+\frac{m_r}{m_n}(m_r+m_n)}\right)\Big/\ln\left(1+\gamma\frac{m_n}{m_r+m_n}\right)$$

From Lemma 2, and plugging in the expression in Lemma 1 for the in-degree of a firm, I know that,

$$\sigma_{i,t}^2 = \frac{M_{i,t}+\frac{m_r}{m_n}(m_r+m_n)}{M_{i,t}}\left[(t-i)\frac{\gamma\frac{m_n}{m_r+m_n}}{1+\gamma\frac{m_n}{m_r+m_n}}+\frac{M_0}{\frac{m_r}{m_n}(m_r+m_n)+M_0}\right]\sigma_0^2$$

This expression defines and implicit relationship between σ^2 and M. Plugging the expression for $t-i$ into the previous expression, I can solve for this implicit relation. I get the following expression for σ^2 as a function of M,

$$\sigma^2(M) = \frac{M + \frac{m_r}{m_n}(m_r+m_n)}{M}$$
$$\times \left[\frac{\gamma \frac{m_n}{m_r+m_n}}{\left(1+\gamma\frac{m_n}{m_r+m_n}\right)\ln\left(1+\gamma\frac{m_n}{m_r+m_n}\right)} \ln\left(\frac{M+\frac{m_r}{m_n}(m_r+m_n)}{M_0+\frac{m_r}{m_n}(m_r+m_n)}\right) + \frac{M_0}{\frac{m_r}{m_n}(m_r+m_n)+M_0}\right]\sigma_0^2$$

As γ gets small, the multiplicative term in front of the log gets close to 1. For γ small, $\sigma^2(M)$ is given by,

$$\sigma^2(M) = \frac{M + \frac{m_r}{m_n}(m_r+m_n)}{M}\left(\ln\left(\frac{M+\frac{m_r}{m_n}(m_r+m_n)}{M_0+\frac{m_r}{m_n}(m_r+m_n)}\right) + \frac{M_0}{M_0+\frac{m_r}{m_n}(m_r+m_n)}\right)\sigma_0^2$$

Taking $M_0 = 0$, I get the proposed expression for $\sigma^2(M)$. ■

Proposition 4 *All firms have the same out-degree distribution, f_{out},*

$$f_{out} = m_r \sum_{t=0}^{\infty}\left(\frac{m_n}{m_r+m_n}\right)^t \underbrace{f_0 * f_0 ... * f_0}_{t\ times}$$

Proof. Consider a firm born at time i and located in the origin. This firm forms m_r contacts at random, distributed over $\mathbb{R}$ according to the p.d.f. f_0. These contacts therefore follow the distribution $m_r f_0$. Firm i then forms m_n network-based contacts with the contacts met at random. The firms met by i at random have themselves met a fraction $\frac{m_r}{m_r+m_n}$ of their own contacts at random, distributed according to f_0. These network-based contacts therefore follow the distribution $m_n \frac{m_r}{m_r+m_n} f_0 * f_0 = m_r \frac{m_n}{m_r+m_n} f_0 * f_0 ...$ etc. ■

B Additional economic assumptions and discussions

B.1 A model of trade with informational barriers

In this section, I embed a simple Krugman (1980) model of trade into the model of network formation described in Section 1. The only assumption added to the Krugman model is that firms can only sell their output to a consumer they have met through the directed network described above. In the next section, I propose a simple model with informational asymmetries and moral hazard that would justify such a selective trading strategy.

Preferences: there is a continuum of consumers in each country, that share the same CES preferences, but differ in the set of goods they have access to. Consumer i has the following preferences over the set Ω_i of goods it has access to,

$$U_i = \left(\int_{\Omega_i} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}$$

where the elasticity of substitution σ is larger than 1.

Technology: there is a continuum of monopolistic firms that face the same increasing returns technology. The labor required to produce q units is

$$l(q) = \alpha + \beta q$$

Informational frictions: the only departure from the classical Dixit-Stiglitz-Krugman set-up is that firms can only sell their output to consumers they know (through the network described above), and consumers can only buy from firms they know (through the same network). More precisely, each consumer has access to a mass M of goods ($M_i = \int_{\Omega_i} d\omega$ for consumer i), with M distributed within the population according to the distribution $F(M)$. In the same way, each firm has access to a mass M of consumers, again with M distributed within the population according to the distribution $F(M)$. For simplicity, I assume that there are no additional barriers to trade, in the form of either a fixed or a variable trade cost.

Prices, quantities and utilities: given that each consumer has access to a continuum of differentiated goods (only the measure of those goods varies across consumers), their demand for each good is iso-elastic.[42] Facing an iso-elastic demand function, each firm charges a constant mark-up over marginal cost, $p_i(\omega) = p = \frac{\sigma}{\sigma-1}\beta w$ for any (i, ω), where w is the wage rate, which I normalized to 1. Without loss of generality, I can normalize $\beta = \frac{\sigma-1}{\sigma}$ so that $p = 1$.

Facing those prices, if consumer i that has access to M_i goods, she will buy $q_i = 1/M_i$ units of each good $\omega \in \Omega_i$. The welfare of this consumer is then simply $U_i = U(M_i) = 1/M_i^{\frac{1}{1-\sigma}}$, where $M_i^{\frac{1}{1-\sigma}}$ is the ideal price index that this consumer faces.

General equilibrium: Imposing free entry of firms in each location will pin down the number of firms. This is left as an exercise. Alternatively, I can assume that firms are born from the process described above, and that aggregate profits are redistributed lump sum to the consumers who collectively own all the firms in the economy. Alternatively, breaking the symmetry between all

[42]In a model with a finite discrete number of goods, the price elasticity of demand would depend on the number of accessible goods, asymptoting to the constant elasticity case only for a large number of goods. Having a continuum of goods for each consumer assumes away this complication.

locations, one would have to impose trade balance between locations in order to solve for relative wages.

Aggregate welfare: Given the simple structure of the economy, I can perform a series of comparative statics experiments. First, I will describe the aggregate welfare of this economy, or equivalently, the average utility reached by consumers in this economy,

$$\mathbb{E}_F [U] = \int_{M \geq 0} M^{\frac{1}{\sigma-1}} dF(M)$$

Jackson and Rogers (2007) derive a series of properties of the distribution $F(M)$, which allow me to describe the impact of changing the technological parameters of the network formation (m, r) on aggregate welfare.

Proposition 5 *Aggregate welfare increases with the total number of links made by newborn firms,* m*.*

Proof. For r fixed, if $m > m'$, then the in-degree distribution F associated with m first order stochastically dominates the distribution F' associated with m' (see Jackson and Rogers (2007), Theorem 7 page 905). Since the utility associated with having access to a mass M of goods, $M^{\frac{1}{\sigma-1}}$, is increasing in M, the aggregate welfare $\mathbb{E}_F [U]$ associated with m is higher than the aggregate welfare $\mathbb{E}_{F'} [U]$ associated with m'. ■

Proposition 6 *If goods are sufficiently substitutable,* $\sigma > 2$*, then aggregate welfare increases with the ratio of random to network-based meetings,* r*. Otherwise, for* $1 < \sigma < 2$*, aggregate welfare decreases with* r*.*

Proof. For m fixed, if $r > r'$, then the in-degree distribution F associated with r second order stochastically dominates the distribution F' associated with r' (see Jackson and Rogers (2007), Theorem 6 page 903). If $\sigma > 2$, the utility associated with having access to a mass M of goods, $M^{\frac{1}{\sigma-1}}$, is concave in M. The aggregate welfare $\mathbb{E}_F [U]$ associated with r is therefore higher than the aggregate welfare $\mathbb{E}_{F'} [U]$ associated with r'. If $\sigma < 2$, the utility is convex in M, and the opposite holds. ■

The intuition for those results is rather simple. First, increasing the number of links formed by newly born firms, m, increases the number of goods accessible to all consumers. Since consumers in this simple model have a love for variety, more links will unambiguously increase welfare for all consumers.

Second, increasing the ratio of random to network-based meetings decreases the dispersion of the number of contacts across consumers. As explained in the main body of the text, network-based meetings give an advantage to agents who already have many contacts, which makes the access to new contacts more unequal. If goods are sufficiently substitutable, increasing the number of contacts brings about a smaller and smaller welfare gain. As a consequence, aggregate utility is higher for a less "unequal" network. On the other hand, if goods are less substitutable, increasing the number of contacts brings about a larger and larger welfare gain. In that case, aggregate utility will be higher for a more "unequal" network.

Sales distribution: I can derive similar predictions for the distribution of sales across firms, as well as for aggregate production.

Firms differ in the mass of consumers they have access to, M. Moreover, each of their consumers themselves differ in the number of goods they have access to. Since by assumption all consumers have the same income, and since all goods have the same price, consumers with access to more goods will buy less of each good. The quantity of each good bought by a consumer who has access to μ goods is then simply $1/\mu$. The expected quantity sold to each consumer is therefore given by $\int_{\mu\geq 0} \frac{1}{\mu} dF(\mu)$, and the total expected sales of a firm that has access to M consumers is,

$$pQ_F(M) = M \int_{\mu\geq 0} \frac{1}{\mu} dF(\mu)$$

As for aggregate welfare, the characterization of the properties of the distribution $F(M)$ in Jackson and Rogers (2007) allows me to describe both aggregate sales and the sales distribution across firms.

First, the higher the total number of links formed at birth, m, the less a firm will sell to any single consumer. This result simply derives from the fact that the more alternatives a consumer has, the fewer goods she will buy from any single supplier. Second, the higher the ratio of random to network-based meetings, r, the less a firm will sell to any single consumer. This result derives from the fact that the higher r, the less dispersed the distribution $F(M)$ is; moreover, a firm can increase its expected sales by shifting away from consumers who have many alternatives towards consumers who have few alternatives; as a consequence, the more dispersed the distribution $F(M)$ is, the more a firm can sell to individual consumers on average.

Note however that this simple model does not generate any interesting prediction on the intensive margin of sales, i.e. the average sales per consumers. More precisely, in expectation, the consumers reached by any firm have access to the same number of goods, irrespective of the number of consumer this firm reaches. Therefore, in expectation, all firms will sell the same quantities (and values) per consumer, irrespective of how many consumers a firm reaches. This result is obviously at odds with the fact that firms that sell to many markets tend to sell large quantities in each of these markets.

I leave an extension of this model that would incorporate a meaningful intensive margin of trade for future research.

B.2 Trading under the threat of moral hazard

In this section, I propose a simple model with informational asymmetries and moral hazard that explains why a given firm would only trade with firms it has met through the network described in Section 1. Note that this model is meant only as an illustration of a possible economic mechanism that would support the proposed dynamic network formation. As a consequence, the model is purposefully simple.

Set-up: there is a continuum of firms of mass 1. Each firm produces a differentiated good. Each firm can both buy differentiated inputs from other firms and sell its differentiated output to other firms. A good can be of either high quality (q_H) or low quality (q_L). Producing high quality goods is costly. The quality of a good is observable and can be contracted upon.

When a supplier meets a buyer, the match specific cost to the supplier of customizing its good for the client is c. The cost c is drawn over $\mathbb{R}^+$ from a known probability distribution G,

$$\Pr\left(\tilde{c} < c\right) = G\left(c\right)$$

For simplicity, I assume that the distribution of customization costs is independent across matches. This cost c is only observable to the supplier, and cannot be contracted upon. I normalize the cost of producing a low quality good to zero.

A high quality input has a value V for the client. The value of a low quality good for a client is normalized to zero. I assume for simplicity that those values are the same for all firms.

Upon a successful match, one unit of output is traded. All firms are risk neutral.

Upon meeting, the timing of the game played by a supplier and its client is as follows:

1. The client offers a price for a high quality good, and a price for a low quality good.

2. The supplier receives a customization cost draw, c, and decides whether to produce a high or low quality good.

3. After observing the good's quality, the client and supplier trade at the agreed prices.

I will look for sub-game perfect Nash equilibria of this game.

Solution to the match specific game: The supplier will produce a high quality good for any price above its cost draw. This happens with probability $G(p_H)$. Conditional on receiving a high quality good, the surplus of the client is $(V - p_H)$. The client therefore chooses p_H^* so as to maximize her expected profits,[43]

$$p_H^* = \arg\max_{p_H} (V - p_H) G(p_H)$$

Since a low quality good has no value, the client sets $p_L^* = 0$, and no low quality good is ever produced or traded. Facing those prices, the expected surplus of a match for a supplier, S, is given by,

$$S = \int_0^{p_H^*} (p_H^* - c)\, dG(c)$$

Random meetings and search frictions: each period, suppliers engage in a costly search for potential clients. The marginal cost of finding a new client for a supplier is given by $s(m)$, with $s(0) = 0$, $s \geq 0$, $s' > 0$ and $\lim_{m \to \infty} s(m) = +\infty$. Given the expected profit from finding a successful match S, a supplier will sample a mass M_r of firms at random, defined by $s(M_r) = S$. Given that a successful match is formed with a probability $G(p_H^*)$ with each client met, a supplier forms a mass m_r of successful matches,

$$m_r = \frac{M_r}{G(p_H^*)}$$

Note that I do not consider the role of geography in this simple example. It would be trivial to add a geographic dimension, where the M_r matches, and the subset of m_r successful matches, are distributed over space according to the p.d.f. f_0.

Network based meetings: given the search frictions, clients who themselves are suppliers to other clients are in a privileged position to leverage the information about their own network of clients. The game played by the upstream supplier, her client, and the downstream clients of her

[43] Note that an interior solution exists under some regularity conditions on V and $G(\cdot)$. $V = 4$ and $G(c) = 1 - 1/c$ for instance admits the simple solution $p_H^* = 2$.

client is similar to the game above. The initial client makes a take-it-or-leave-it offer to each of her supplier to reveal information about each of her own clients for a fee $S = \int_0^{p_H^*} (p_H^* - c)\, dG(c)$. After being connected, the supplier and her new client bargain as above, the new client sets a price p_H^* for a high quality good, which the supplier accepts only if her match specific cost draw is below p_H^*. S is exactly the surplus that a supplier can expect from meeting with a random client. Since the initial client does not observe the cost c that the supplier would have to incur to customize her good for a new client, no additional information permeates. Moreover, the initial client extracts all the information rent from the supplier.

Each period, a supplier will therefore meet a fraction $G(p_H^*)$ of the clients of each of her own clients. Note that given the implicit assumption about constant returns to scale, there is no strategic consideration for initial clients to reveal (at a cost) their own client base to their suppliers. They simply extract a fee for each contact they reveal, without the fear of losing their own contacts.

Discussion: the proposed model would explain why each period, firms search for a fixed number of contacts at random, and in addition, get connected to a subset of the contacts of their contacts. The dynamic network that arises from this set-up is more realistic but also more complicated than the one proposed in Section 1, since each period, any new contact acquired by a firm in the existing network of a supplier may be revealed to the supplier. The simplifying assumption of having the network evolving only on the in-degree side (with the out-degree acquired once and for all at birth) is lost. However, the main force that generates a fat tailed network remains: firms that already have many contacts are more likely to acquire new contacts. And the force that keeps the network from being a scale-free network also remains: each period, firm always have an incentive to look for contacts at random.

The study of such a more realistic but also more complex dynamic network is left for future research.

B.3 Comparison with existing trade models

In this section, I derive formally the predictions of the two most prominent existing firm level trade models, Melitz (2003) and Bernard, Eaton, Jensen and Kortum (2003), regarding the cross-sectional distribution of the number of foreign markets reached by French exporters. I show that neither of those models makes any robust prediction regarding this distribution, unless some ad-hoc assumption about all the exogenous parameters of those models is made. I argue that there is no a priori reason to make any such ad hoc assumption.

Comparison with Melitz (2003) and Chaney (2008)

Strictly speaking, the model developed in Melitz (2003) predicts a degenerate distribution $f(M)$ where all exporters export to each and every country in the world. This is due obviously to the simplifying assumption that all trade barriers, country sizes and labor productivities are perfectly symmetric. Chaney (2008) develops a simple extension of Melitz (2003) with asymmetric countries and trade barriers. I will describe the prediction of this model regarding the p.d.f. $f(M)$.

Set-up: as a reminder, the set-up in Chaney (2008) is as follows. I will only describe the patterns of entry of French exporters into all foreign countries. Preferences are CES with an elasticity of substitution σ. The cost of selling q units of good in country m for a French firm with productivity φ is

$$c_m(q) = \frac{w_F \tau_{F,m}}{\varphi} q + f_{F,m}$$

where w_F is the French wage, $\tau_{F,m}$ and $f_{F,m}$ are respectively the variable and fixed cost of export to country m for a French firm. Productivities are distributed Pareto,

$$\Pr(\Phi \leq \varphi) = 1 - \varphi^{-\gamma}$$

Chaney (2008) proves that all firms with a productivity above $\overline{\varphi}_{F,m}$ export to country m, with $\overline{\varphi}_{F,m}$ defined as,

$$\overline{\varphi}_{F,m} = \lambda \left(\frac{Y}{Y_m}\right)\left(\frac{w_F \tau_{F,m}}{\theta_m}\right) f_{F,m}^{1/(\sigma-1)}$$

with λ a constant, $\theta_m^{-\gamma} = \sum_{n=1}^{N} (Y_n/Y)(w_n \tau_{n,m})^{-\gamma} f_{n,m}^{-(\gamma/(\sigma-1)-1)}$, Y the world *GDP*, and Y_m the *GDP* of country m. There is a strict ordering of the productivity thresholds faced by French exporters. Without loss of generality, I can rearrange countries in increasing productivity thresholds, so that $\overline{\varphi}_{F,m} > \overline{\varphi}_{F,m'}$ iif $m > m'$. Any firm with a productivity above $\overline{\varphi}_{F,m}$ exports at least to all markets $m' \leq m$.

Prediction regarding the p.d.f. $f(M)$: the fraction of firms that exports to exactly M markets is simply given by the probability of receiving a productivity φ above $\overline{\varphi}_{F,M}$ but strictly below $\overline{\varphi}_{F,M+1}$. A firm with such a productivity will export to all countries $m \leq M$ (there are exactly M of them), but not to any country $m > M$. Using the assumption of Pareto distributed productivity shocks, I can derive a simple expression for the fraction of firms that export to exactly M countries,

$f(M)$,

$$f(M) = (\lambda w_F)^{-\gamma}\left[\left(\frac{Y}{Y_M}\right)^{-\gamma}\left(\frac{\tau_{F,M}}{\theta_M}\right)^{-\gamma} f_{F,M}^{-\gamma/(\sigma-1)} - \left(\frac{Y}{Y_{M+1}}\right)^{-\gamma}\left(\frac{\tau_{F,M+1}}{\theta_{M+1}}\right)^{-\gamma} f_{F,M+1}^{-\gamma/(\sigma-1)}\right]$$

The only prediction of this model is that this density is non negative (it may be zero if the knife-edge case where two productivity thresholds are exactly equal arises). It is easy to see from this formula that one can make any $f(M)$ either arbitrarily small (by having $\left(\frac{Y}{Y_{M+1}}\right)\left(\frac{\tau_{F,M+1}}{\theta_{M+1}}\right) f_{F,M+1}^{1/(\sigma-1)}$ arbitrarily close to $\left(\frac{Y}{Y_M}\right)\left(\frac{\tau_{F,M}}{\theta_M}\right) f_{F,M}^{1/(\sigma-1)}$), or arbitrarily close to 1 (by having $\left(\frac{Y}{Y_M}\right)\left(\frac{\tau_{F,M}}{\theta_M}\right) f_{F,M}^{1/(\sigma-1)}$ arbitrarily small compared to $\left(\frac{Y}{Y_{M+1}}\right)\left(\frac{\tau_{F,M+1}}{\theta_{M+1}}\right) f_{F,M+1}^{1/(\sigma-1)}$). There is no reason a priori that the function $f(M)$ is even decreasing in M: if the thresholds of entry for two relatively accessible countries are arbitrarily close, and the thresholds of entry into two relatively inaccessible countries are arbitrarily distant, then $f(M)$ will be upward sloping.

One may argue that country sizes (the Y_M's) are approximately Pareto distributed. One may further argue that the fixed export costs are approximately linear in country size, or at least linear in logs. In such a case, and in the absence of variable trade barriers, the thresholds of entry into different foreign markets would be Pareto distributed. Formally, in such a case, one can write $\overline{\varphi}_M = \alpha M^{\beta}$, so that $f(M) = \alpha\left(M^{-\beta\gamma} - (M+1)^{-\beta\gamma}\right)$. For $\beta\gamma \approx 1.5$, this relationship would describe the data relatively well. Whereas this relationship does not exhibit the curvature in a log-log scale that we see in the data, the predicted line is close to the empirical distribution $f(M)$.

However, this arguments abstracts entirely from the existence of variable trade barriers. Such a model would make the counter-factual prediction that the number of French exporters is log-proportional to country size. Eaton, Kortum and Kramarz (2010) show that French exporters tend to cluster in countries that are geographically close to France, whether large or small, and not in potentially much larger countries that are faraway from France. In other words, they show evidence that variable trade barriers do play an important role in shaping the entry of French exporters into foreign market, or that fixed export costs are not proportional to country size. For instance, in 1986, 17,695 French firms export to Belgium, with a *GDP* of \$144 billion; 14,579 export to Germany, with a *GDP* of \$1.01 trillion; and 7,608 export to the U.S., with a *GDP* of \$4.43 trillions. So Belgium, which is 31 times smaller than the U.S. attracts more than twice as many French exporters as the U.S. It is 7 times smaller than Germany but still attracts 20% more French exporters. Germany, which unlike Belgium is a non French speaking country, is 4.5 times smaller than the U.S. but receives twice as many French exporters. Even comparing two English speaking countries that are not contiguous to France, the U.K. with a GDP of \$570 billion is almost 8 times smaller than the U.S., but receives 30% *more* French exporters, not 87% *less*. These massive departures from a linear relationship between country size and number of French exporters are not restricted to these 4 countries, but occur systematically. Empirically, there is no systematic correlation between country size (measured as GDP) and the distance from France. So there is no reason to believe that the thresholds of entry into different markets are themselves Pareto distributed.

To illustrate this point, using the very same data on French manufacturing exporters in 1986 as Eaton, Kortum and Kramarz (2010), I get the following numbers for the density $f(M)$ for $M = 1, \ldots, 6$:

$$\begin{array}{ll} f(1) = 3,120 & f(4) = 891 \\ f(2) = 406 & f(5) = 1,458 \\ f(3) = 3,530 & f(6) = 686 \end{array}$$

Evidently, the theoretical predictions of the Chaney/Melitz model regarding the number of foreign markets reached by individual firms are at odds with the data. The predicted density $f(M)$ is not even decreasing in M, as it is in the data.

To further illustrate this point, I calibrate the Melitz/Chaney model so as to match exactly the number of French exporters in every foreign market. I use the same data on all French exporters in 1992 that I used in Section 2. Given the precise ordering of foreign markets predicted by the model, I can rank foreign markets in decreasing order of accessibility for French exporters. The number of firms that export to exactly M markets is then the difference between the number of firms that export to the M^{th} and $(M+1)^{th}$ market. As can be seen visually on Figure 3, the Melitz/Chaney model cannot replicate the empirical distribution of the number of foreign markets accessed by individual firms.

I can also describe the predictions of the Melitz/Chaney model regarding the geographic dispersion of exports. In the Melitz/Chaney model, more productive firms are able to enter both more markets, and less accessible markets. Since less accessible markets tend to be geographically more remote, this model predicts that the geographic dispersion of exports tends to be larger among firms that export to many markets than among firms that export to few markets. However, to generate such a prediction, the Melitz/Chaney model again relies on a series of exogenous parameters that this model has nothing to say about a priori. Therefore, this model does not deliver any precise prediction regarding the shape of the relationship between geographic dispersion and the number of markets accessed. To illustrate this point, I calibrate the geographic dispersion

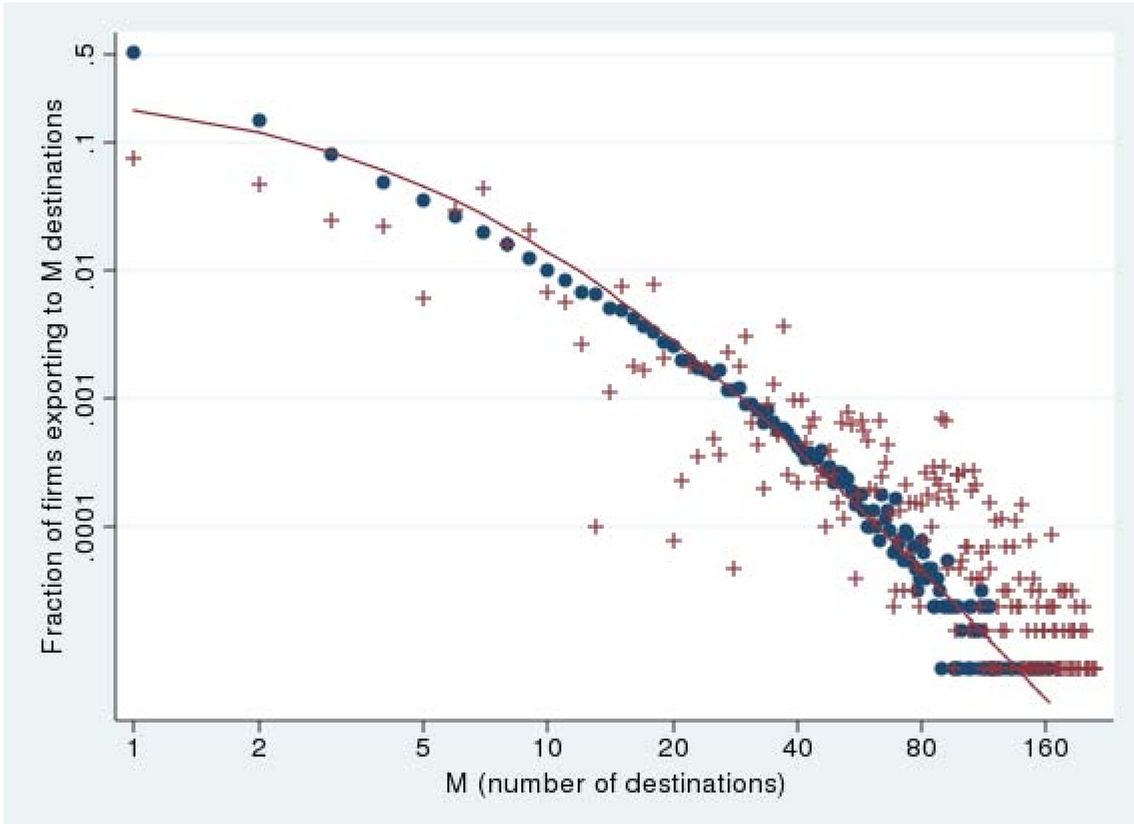

Figure 3: Network versus Melitz/Chaney model, $f(M)$ versus M.

Notes: $f(M)$ is the fraction of firms exporting to M destinations; blue dots: data, all French exporters in 1992; red line: calibrated network model; red plus signs: calibrated Melitz/Chaney model.

of exports, $\sigma^2(M)$, in the Melitz/Chaney model. I control for country size in the same way as I do when constructing the empirical measure of $\sigma^2(M)$ presented in Appendix C.2 on page 56. I order foreign markets in decreasing order of accessibility for French exporters, controlling for market size as follows: call N_M the number of French firms that export to market M, where M is defined such that $\frac{N_1}{GDP_1} > \frac{N_2}{GDP_2} > \frac{N_3}{GDP_3} > \ldots$ The Melitz/Chaney model predicts that any firm that exports to exactly M markets will export to all countries $c \leq M$. The predicted geographic dispersion of exports among firms that export to M markets is therefore given by,

$$\sigma^2_{Melitz/Chaney}(M) = \frac{\sum_{c \leq M} (Distance_{Fr,c})^2}{M}$$

As can be seen visually on Figure 4, the Melitz/Chaney can only predict that the geographic dispersion of exports tends to increase with the number of foreign markets accessed, but it has nothing to say about the specific shape of this relationship. Not controlling for market size, the

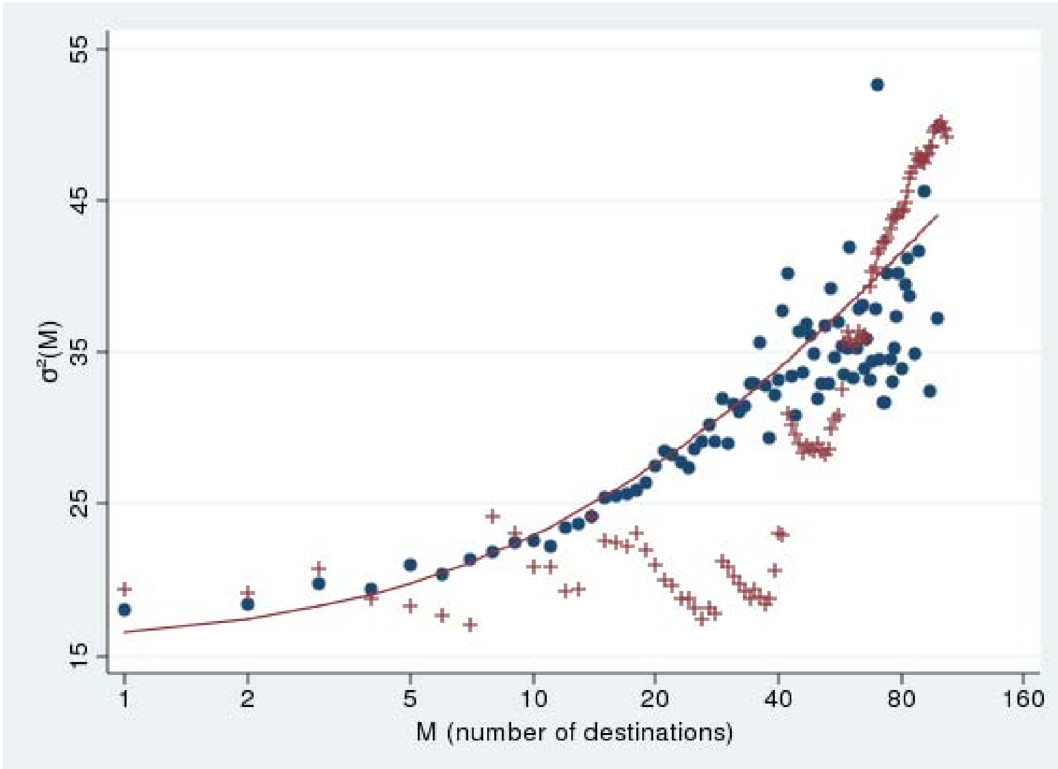

Figure 4: Network versus Melitz/Chaney model, $\sigma^2(M)$ versus M.

Notes: $\sigma^2(M)$ is the second moment of the distance from a firm's export destinations, among firms exporting to M destinations; blue dots: data, all French exporters in 1992; red line: calibrated network model; red plus signs: calibrated Melitz/Chaney model.

empirical fit of the Melitz/Chaney model would be substantially worse.[44]

Note of course that the Melitz/Chaney model is not only meant to explain the extensive margin of international trade (the number of foreign markets accessed by exporters), but it also delivers a series of predictions on the intensive margin of international trade (the size of firm level exports), on how the size of sales in the domestic market helps predict which markets a firm enters, and how much it exports there. As shown by Eaton, Kortum and Kramarz (2010), these predictions are strongly supported by the data.

To conclude, not only would the Melitz/Chaney model require ad hoc assumptions regarding the exogenous parameters of the model (which the Melitz/Chaney model has nothing to say about) to match the data, but there is no empirical ground for making such ad hoc assumptions.

Comparison with Bernard, Eaton, Jensen and Kortum (2003)

[44]Calibration available from the author upon request.

Whereas in Bernard, Eaton, Jensen and Kortum (2003), the formula for the fraction of firms that export to country m is almost identical to the formula in Chaney (2008), there is not a strict hierarchy of foreign markets in terms of accessibility to French exporters. As a consequence, the formula for the fraction of firms that export to exactly M markets is substantially more complicated.

In the interest of clarity, I will therefore solve a simple special case with 3 countries (potentially asymmetric in labor productivity), and symmetric trade barriers. All the intuition derived in this special case carry over to the case of many countries with asymmetric bilateral trade barriers.

Set-up: as a reminder, the set-up in Bernard, Eaton, Jensen and Kortum (2003) is as follows. There is a continuum of sectors that produce differentiated goods. The distribution of labor productivity z of the most productive firm in country m in any of those sectors is Fréchet,

$$\Pr\left(Z_m \leq z\right) = \exp\left(-T_m z^{-\theta}\right)$$

The parameter θ is the same across countries, but T_m differs across countries. Note that there is a one-to-one mapping between the parameter T_m and the *GDP* of country m. Productivity draws are independent across countries. Firms face no extra cost of selling domestically, but they face an iceberg cost τ when exporting to any foreign country.

Prediction regarding the p.d.f. $f\left(M\right)$: Let us set country 1 as France, and compute the fraction of French firms that sell in 1, 2 or 3 markets.

All firms with a productivity $z_1 > \max\left\{\tau z_2, \tau z_3\right\}$ sell in all three markets. With independent Fréchet distributions, the probability of such an event occurring, and therefore the fraction of firms from 1 that export to exactly three markets, $f\left(3\right)$, is given by,

$$f\left(3\right) = \frac{T_1}{T_1 + \tau^\theta T_2 + \tau^\theta T_3}$$

A firm in country 1 sells in exactly two markets if either it is the best in country 1 and 2, but not in 3, or it is the best in country 1 and 3, but not in 2. Formally, a firm with a productivity z_1 sells in exactly two markets if either $\left\{z_1/\tau > z_2 \text{ and } z_1/\tau < z_3 < z_1\right\}$ or $\left\{z_1/\tau > z_3 \text{ and } z_1/\tau < z_2 < z_1\right\}$ are true. The respective probabilities of each of those mutually exclusive events are,

$$\begin{aligned}
\Pr\left\{z_1/\tau > z_2 \text{ and } z_1/\tau < z_3 < z_1\right\} &= \frac{T_1}{T_1 + \tau^\theta T_2 + T_3} - \frac{T_1}{T_1 + \tau^\theta T_2 + \tau^\theta T_3} \\
\Pr\left\{z_1/\tau > z_3 \text{ and } z_1/\tau < z_2 < z_1\right\} &= \frac{T_1}{T_1 + T_2 + \tau^\theta T_3} - \frac{T_1}{T_1 + \tau^\theta T_2 + \tau^\theta T_3}
\end{aligned}$$

The fraction of firms from 1 that export to exactly two market, $f(2)$, is given by,

$$f(2) = \frac{T_1}{T_1 + \tau^\theta T_2 + T_3} - \frac{T_1}{T_1 + \tau^\theta T_2 + \tau^\theta T_3} + \frac{T_1}{T_1 + T_2 + \tau^\theta T_3} - \frac{T_1}{T_1 + \tau^\theta T_2 + \tau^\theta T_3}$$

The formula for the fraction of firms that sell in exactly one market (in the home market necessarily) is even more complicated, and contains a total of 16 terms. I will spare the reader and skip this formula, concentrating instead on exporters only. Note the slight abuse of notation due the fact that there is a non empty set of firms which despite being the most productive among home firms do not sell in any market (not even at home). The exact fractions of firms selling to exactly 2 and 3 markets are the above formulas divided by the same number, the probability of selling in at least one market.

From the formulas above, it is easy to see that the distribution $f(M)$ does not even have to be downward sloping. As trade barriers become infinitely large ($\tau \to +\infty$), no firm is able to export anywhere, $f(2) = f(3) = 0$, and all firms sell at home, $f(1) = 1$. On the opposite extreme, when trade barriers vanish ($\tau \to 1$), any firm that survives will sell in all three markets, $f(3) = \frac{T_1}{T_1+T_2+T_3}$, but no firm sells to exactly 1 or 2 markets, $f(1) = f(2) = 0$. In the first case, the distribution $f(M)$ is decreasing in M, whereas in the second case, it is increasing in M. Moreover, whereas the fraction of firms able to enter all 3 markets monotonically decreases with the level of trade barriers, τ, the fraction of firms that sell to exactly 2 markets is not monotone in the level of trade barriers. $f(2)$ increases in τ for τ small, and decreases in τ for τ large. For different levels of trade barriers, the fraction of firms that sell to exactly 2 markets will be alternatively larger or smaller than the fraction of firms selling to 3 markets.

As in the Melitz/Chaney model, the distribution $f(M)$ in the Bernard, Eaton, Jensen and Kortum model not only can take any shape, but it does not even have to be downward sloping. The specific shape of this distribution depends on ad hoc assumptions regarding the distribution of exogenous parameters (country sizes or relative productivities, T_m's, and bilateral trade barriers, τ_{nm}'s), about which the model has nothing to say.

C Data

In this section, I describe the source and construction of the data, I provide some descriptive statistics in addition to what is presented in the main body of the paper, and I perform a series of robustness checks.

C.1 Data sources

In this section, I describe the various sources for the data that I use in Section 2.

Firm level export data

The data on firm level exports comes from the French customs, and is described in greater details in Eaton, Kortum and Kramarz (2010). Until 1992, all shipments crossing the French border are reported, either by the owner of the (exporting) firm, or by authorized customs commissioners. Information about the identity of the exporting firm, the value of the shipment, the industrial sector, and the destination country is recorded. This information is then aggregated over a year. I use data on all French exporters (including non manufacturing firms).[45] A data point is therefore a firm, year, destination country and value of exports (in French Francs) vector. Since I am primarily interested in the extensive margin of exports, I do not use information on the value of exports, except in Appendix C.4.

In addition, the customs data are matched with balance sheet information collected by the French fiscal authorities. All firms with a turnover of 1,000,000 French Francs in services, or 3,000,000 French Francs in manufacturing are mandated to report this information. Virtually all exporters are included in this data set. I use information on annual sales, employment, and capital expenditure.

I restrict my sample of firms to exporters only.

Distance data

I use data on bilateral distances between countries collected and constructed by the CEPII. The distance between two countries is calculated as a weighted arithmetic average of the geodesic distances between the main cities in these countries, where population weights are used. Data on the location of the main cities in each country (latitude and longitude), as well as the population of those main cities are used to compute those distances. The construction of the data is described in further details by Mayer and Zignago (2006) at `http://www.cepii.fr/distance/noticedist_en.pdf`.

Country size data

I use as a measure of a country's size its nominal *GDP* (in US$) in the current year. The data are collected from the Penn World Tables and are described in further details at `http:`

[45]Restricting the sample to manufacturing firms does not alter the results significantly.

//pwt.econ.upenn.edu/.

Bilateral trade flows

To proxy for the intensity of firm level contacts between countries other than France, I use data on bilateral trade flows between countries. The data corresponds to the nominal value (in US$) of aggregate trade flows between country pairs. The data are collected from the NBER, and are described in further details in Feenstra et al. (2004).

C.2 Data construction

In this section, I describe and discuss in details how I construct several empirical measures that I use in Section 2.

Construction of the p.d.f. $f(M)$ used in Section 2.2

The empirical test of Proposition 2 simply uses the fraction of firms exporting to exactly M markets as a proxy for the fraction of firms with exactly M contacts. Below, I show formally how to correct for the fact that some firms may have more than 1 contact per market.

Consider the following simplified set-up. There are N distinct countries, each one populated by the same number of firms. These firms are connected to one another by the directed network described in Section 1. The distribution of the number of contacts M is therefore given by the c.d.f. $F(M)$ given in Proposition 2. Assume for simplicity that physical geography does not matter, so that any contact is equally likely to be located in any of the N countries, and that the location of any two contacts are independent from one another. In other words, there is a probability $1/N$ that a given contact is located in a given country, and those probabilities are independent across contacts.

Consider now the following question: what is the probability, $\Pr(n|M)$, that a firm with M contacts has contacts in n distinct countries?

For a given number of contacts M, there are N^M equiprobable distinct ways of distributing M contacts into the N countries of the world, so that each configuration is realized with a probability $\left(\frac{1}{N}\right)^M$. There are $\binom{N}{n}$ different ways of choosing n countries out of the total of N countries. Call $G(M,n)$ the number of distinct ways to assign each of the M contacts into n countries. The probability $\Pr(n|M)$ that a firm with M contacts has contacts in n distinct countries is then simply,

$$\Pr(n|M) = \binom{N}{n} \left(\frac{1}{N}\right)^M G(M,n)$$

The number $G(M,n)$ is defined recursively as follows. There are two mutually exclusive cases. In the first case, the first $M-1$ contacts are assigned to only $n-1$ countries. In that case, the last M^{th} contact must necessarily be assigned to the n^{th} country. There are n equiprobable such cases, one for each n^{th} missing last country. There are $G(M-1,n-1)$ ways to assign $n-1$ countries to $M-1$ contacts, n candidate countries that can be missing for the last contact. There are therefore $nG(M-1,n-1)$ distinct ways of assigning n countries to M contact in that first case. In the second case, the first $M-1$ contacts are assigned to n countries. We can then assign the last contact to any one of the n countries. There are therefore $nG(M-1,n)$ distinct ways of assigning n countries to M contacts in that second case. The number $G(M,n)$ is defined recursively as,

$$G(M,n) = n\left[G(M-1,n-1) + G(M-1,n)\right]$$

with the initial conditions $G(M,M) = M!$ and $G(M,1) = 1$. Noting that the known Stirling number of the second kind, $S_2(M,n)$, is defined recursively in a similar fashion,

$$S_2(M,n) = S_2(M-1,n-1) + nS_2(M-1,n)$$

with the initial conditions $S_2(M,M) = 1$ and $S_2(M,1) = 1$, I get the following relationship between the numbers $G(M,n)$ and $S_2(M,n)$,

$$G(M,n) = n!S_2(M,n)$$

I can now answer the question of interest: given that the number of contacts, M, is distributed according to the distribution $F(M)$, and that there are N equal sized countries, what is the distribution of the number of countries, n, accessed by different firms?

Obviously, a firm that has M contacts can at most export to M different countries. But for any $M > 1$, the probability that two different contacts fall into the same country is positive, and it increases with M. In other words, among firms with contacts in n distinct countries, there are firms with n, $n+1$, $n+2$, ... contacts.

The fraction of firms that have exactly M contacts is simply given by the p.d.f. $f(M)$, associated with the c.d.f. $F(M)$ defined in Proposition 2. Given the distribution above of the number of countries reached by a firm with M contacts, $\Pr(n|M)$, the fraction of firms that have contacts in exactly n distinct countries, is given by,

$$\begin{aligned}\varphi(n) &= \sum_{M=n}^{+\infty} f(M)\Pr(n|M) \\ &= \sum_{M=n}^{+\infty} (1+r)\left(\frac{r\times m}{M+r\times m}\right)^{2+r} n! S_2(M,n)\begin{pmatrix} N \\ n \end{pmatrix}\left(\frac{1}{N}\right)^M\end{aligned}$$

For a given total number of countries, N, observing only the number of countries where a firm exports to, n, but not directly the total number of contacts of the firm, M, one can estimate the parameters $\widehat{(r,m)}$ that govern the underlying distribution of the number of contacts, $F(M)$. I present the results of this corrected estimation in Appendix C.3 on the next page.

Construction of $\widehat{\sigma^2(M)}$ used in Section 2.3

To bring Proposition 3 to the data, I need to measure the geographic dispersion of a firm's foreign contacts, $\sigma^2(M)$. The construction of this measure is as follows.

Using data for the year 1992, I calculate the following empirical counterpart to $\sigma^2(M)$, the second moment of the distance from a firm's export destinations for firms with exactly M foreign contacts,

$$\widehat{\sigma^2(M)} = \frac{\sum_{(i,c)\in E(M)\times C}\left(\frac{1}{GDP_c}\right)\times\left(Distance_{Fr,c}\right)^2\times\mathbb{I}\left\{export_{i,c}>0\right\}}{\sum_{(i,c)\in E(M)\times C}\left(\frac{1}{GDP_c}\right)\times\mathbb{I}\left\{export_{i,c}>0\right\}}$$

where GDP_c is country c's GDP, $\left(Distance_{Fr,c}\right)^2$ is the distance (in $1,000$'s of km's) between France and country c squared, $\mathbb{I}\left\{export_{i,c}>0\right\}$ is an indicator function taking the value 1 if firm i exports to country c, $E(M)$ is the set of firms exporting to exactly M foreign markets, and C is the set of all the countries for which I have information on GDP and distance from France. This is the exact empirical counterpart of the second moment of the distance from a firm's export, with each observation weighted by the inverse of the country size.

The justification for using this empirical measure of the geographic dispersion of exports for different types of firms is twofold.

First, following the guidance of the model, I simply assume that among all firms exporting to M foreign countries, each observed export destination is an independent draw from the same distribution f_M.[46] $\widehat{\sigma^2(M)}$ is simply the empirical second moment of this distribution f_M. Note

[46] f_M is the distribution $f_{i,t}$ for (i,t) such that $M_{i,t}=M$.

that I have a large number of observations to calculate each $\widehat{\sigma^2(M)}$. When M is small, there are many firms in $E(M)$, and when M is large, even though few firms are able to export to many markets, I can observe the geographic distribution of their exports into many different foreign markets.

Second, I weight each observation by the inverse of the destination country size ($1/GDP_c$). The reason is that larger countries are, not surprisingly, more likely export destinations than small countries.[47] In order not to give any systematic salience to large countries, that may not be evenly distributed over space, I correct for the impact of country size. Note however that since country size (GDP_c) and distance from France ($Distance_{Fr,c}$) are not systematically correlated, this correction does not change the results in a significant way. See Appendix C.3 below for such robustness checks.

C.3 Robustness checks

In this section, I perform a series of robustness checks on the main empirical findings of Section 2.

Cross-section: $f(M)$ versus M

To check the robustness of the results presented in Section 2.2, I test Proposition 2 and estimate the relevant parameters (r, m) using different estimation procedures, corrections, and samples of firms.

First, I estimate the parameters (r, m) that govern the distribution of the number of foreign contacts in Proposition 2 using a Maximum Likelihood Estimation procedure, as opposed to the Non Linear Least Squares estimates presented in Section 2.2. In this estimation, as in Section 2.2, I assume that the presence in one foreign country corresponds to one single foreign contact in that country.

Second, I take into account the fact that I observe the number of foreign countries reached by French exporters, not the number of actual foreign contacts. To do so, I use the correction developed in Appendix C.2 on page 54. I then estimate the parameters of interest (r, m) using a Maximum Likelihood Estimation procedure.

Third, I compare the fit of the distribution in Proposition 2 to a simpler benchmark where contacts are simply acquired at random. Such a benchmark would correspond exactly to the

[47] This would be the case in most existing firm level trade models. It is the case in the proposed model where a large country is populated by many firms, so that the probability of acquiring a contact there is large.

Table 4: Empirical fit of Proposition2, robustness checks (I)

	(1) (NLLS)	(2) (MLE)	(3) (MLE w. correction)	(4) (MLE)
r	1.58***	1.82***	1.65***	$+\infty$
	(.14)	(.025)	(.022)	
m	4.86***	3.71***	3.83***	3.84***
	(.44)	(.017)	(.019)	(.011)
Adj. R^2/ log (lik.)	0.98	-282,022	-282,098	-296,916
lik. ratio test: (4) vs. (2)		$\Lambda = 29,788$, p-value $< .0001$		

Notes: This table presents the estimates of parameters r and m using different procedures. These two parameters govern the distribution of the number of foreign contacts, $F(M)$, in Proposition 2. I use the same data on French exporters in 1992 for all procedures. Standard errors are in parentheses. *, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.

special case $r \to +\infty$ in Proposition 2. I estimate the parameter m in this benchmark using a Maximum Likelihood Estimation procedure, and propose a likelihood-ratio test to compare this simple benchmark to my model.

The results of these estimations are presented in Table 4. Column (1) shows the results of the non linear least squares estimation of Equation (2), already shown in Section 2.2. Column (2) estimates this equation using a maximum likelihood estimation instead. Column (3) estimates this equation using a maximum likelihood estimation, and controlling for the fact that foreign countries, and not foreign contacts are observed by the econometrician. I use the correction described in Appendix C.2. Specifically, in the formula on page 54 for $\varphi(n)$, the fraction of firms that export to n markets, I sum from n to 310 (and not all the way to $+\infty$). Increasing this upper bound does not affect the results. Finally, column (4) estimates using maximum likelihood Equation (2) in the special case where $r \to +\infty$, that is in the special benchmark case where all contacts are formed at random.

The parameters estimates vary little across the 3 different estimation procedures (1)-(3). The maximum likelihood estimates give a somehow higher r and a lower m, but the difference is small. Moreover, the correction in (3) does not alter the estimates much, which suggests that the simplifying assumption of one country-one contact is a relatively good approximation. Finally, the likelihood ratio test unambiguously rejects the simpler purely random model. As was already visually obvious, network-based meetings do seem to play an important role in shaping the distribution of the number of foreign contacts across exporters, giving it an unmistakable fat upper tail.

I also check that the results presented in Section 2.2 are stable not only across different estimation procedures, but over time. To do so, I replicate the estimations in columns (1)-(3) in Table 4 for each year, from 1986 to 1992. That is, I estimate separately for each year the prediction from Proposition 2 regarding the cross-sectional distribution of the number of foreign contacts using non linear least squares, maximum likelihood, and maximum likelihood after controlling for the fact that I observe countries and not contacts.

The results are presented in Table 5. All estimated parameters are statistically different from zero at the 1% level of significance.[48] Except for the year 1986 where the estimated parameters are somewhat higher in the NLLS estimation, the coefficients are virtually identical across years, and do not differ much across the different estimation procedures. This suggests first that the results are robust, and second that the system of French exporters is in a steady state over the period considered.

Cross-section: $\sigma^2(M)$ versus M

To check the robustness of the results presented in Section 2.3, I test Proposition 3 and estimate the relevant parameter σ_0^2 using different empirical measures of $\sigma^2(M)$ and different samples of firms.

The results are presented in Table 6. For each year between 1986 and 1992, I estimate separately the coefficient σ_0^2. To do so, I use the formula for $\sigma^2(M)$ derived in Proposition 3, imposing the parameter $r \times m$ estimated on the cross-sectional distribution of the number of foreign contacts in the top panel of Table 5. I estimate σ_0^2 using non linear least squares, weighting each observation by the precision of the empirical estimate of $\sigma^2(M)$. I use two alternative empirical measures for $\sigma^2(M)$.

The top panel of Table 6 uses the same empirical measure of the second moment of the distance from a firm's export destinations, $\sigma^2(M)$ as in Section 2.3, where I correct for differences in GDP across countries. The formula for calculating the empirical counterpart of $\sigma^2(M)$ is as follows,

$$\widehat{\sigma^2(M)} = \frac{\sum_{(i,c)\in E(M)\times C}\left(\frac{1}{GDP_c}\right)\times\left(Distance_{Fr,c}\right)^2\times\mathbb{I}\left\{export_{i,c}>0\right\}}{\sum_{(i,c)\in E(M)\times C}\left(\frac{1}{GDP_c}\right)\times\mathbb{I}\left\{export_{i,c}>0\right\}}$$

[48] Estimated standard errors and statistical tests available from the author upon request.

Table 5: Empirical fit of Proposition2, robustness checks (II)

Year	(1986)	(1987)	(1988)	(1989)	(1990)	(1991)	(1992)
	Non Linear Least Squares Estimation of $f(M)$:						
r	1.75	1.54	1.56	1.60	1.59	1.61	1.58
	(.15)	(.12)	(.15)	(.15)	(.15)	(.14)	(.14)
m	5.62	5.20	4.79	4.85	4.72	4.76	4.86
	(.46)	(.42)	(.48)	(.47)	(.48)	(.44)	(.44)
Adj. R^2	0.98	0.98	0.97	0.97	0.97	0.98	0.98
	Maximum Likelihood Estimation of $f(M)$:						
r	1.61	1.64	1.81	1.80	1.86	1.89	1.82
	(.022)	(.022)	(.026)	(.025)	(.026)	(.026)	(.025)
m	4.04	3.98	3.79	3.82	3.76	3.68	3.71
	(.020)	(.019)	(.018)	(.018)	(.018)	(.017)	(.017)
log (lik.)	-285,653	-297,774	-268,388	-280,602	-270,395	-277,770	-282,022
	Maximum Likelihood Estimation with Correction of $\varphi(n)$:						
r	1.45	1.49	1.64	1.64	1.68	1.71	1.65
	(.020)	(.020)	(.023)	(.023)	(.024)	(.024)	(.022)
m	4.19	4.12	3.91	3.93	3.88	3.81	3.83
	(.022)	(.021)	(.020)	(.019)	(.019)	(.019)	(.019)
log (lik.)	-285,711	-297,839	-268,454	-280,670	-270,464	-277,844	-282,098

Notes: This table shows the results of the Non Linear Least Square and Maximum Likelihood estimations of Equation (2) derived from Proposition 2 for French exporters in all years from 1986 to 1992. The top panel uses the log of the fraction of firms that export to M markets as the dependent variable, using a non linear least squares estimation. The middle panel estimates by maximum likelihood the p.d.f. $f(M)$ assuming that the number of contacts of a firm, M, is equal to the number of foreign countries. The bottom panel estimates by maximum likelihood the p.d.f. $\varphi(n)$ using the correction proposed in Appendix C.2 on page 54 for the fact that the number of countries, n, and not contacts are observed. Standard errors are reported in parentheses. All coefficients on this table are statistically different from zero at the 1% level of significance.

Table 6: Empirical fit of Proposition 3, robustness checks

Year	(1986)	(1987)	(1988)	(1989)	(1990)	(1991)	(1992)
				$\sigma^2(M)$ corrected for GDP differences:			
σ_0^2	16.21	15.31	15.15	15.56	15.53	15.41	15.58
	(.09)	(.10)	(.11)	(.10)	(.09)	(.08)	(.10)
Adj. R^2	0.91	0.92	0.89	0.90	0.90	0.93	0.89
				$\sigma^2(M)$ not corrected for GDP differences:			
σ_0^2	16.06	15.06	14.91	15.38	14.90	14.86	15.03
	(.39)	(.37)	(.36)	(.35)	(.39)	(.43)	(.41)
Adj. R^2	0.75	0.76	0.77	0.78	0.75	0.71	0.73

Notes: This table shows the non linear least square estimate of σ_0^2 from Proposition 3, imposing the parameter $r \times m$ estimated by NLLS in the top panel of Table 5. The estimation is run separately for each year from 1986 to 1992. The top panel corrects the empirical measure of $\sigma^2(M)$ for differences in GDP across countries, while the bottom panel does not. Each point weighted by the square root of the number of observations used to compute $\sigma^2(M)$. Standard errors are in parentheses. All coefficients on this table are statistically different from zero at the 1% level of significance.

where $E(M)$ is the set of firms that export to M countries, and C is the set of all countries. In this measure, I discount the number of exporters to a given country c by the GDP of country c, using the fact that the number of exporters to a country is approximately proportional to country size.

In the second panel of Table 6, I do not correct for differences in GDP across countries. The formula for calculating $\sigma^2(M)$ is as follows,

$$\widehat{\sigma^2(M)} = \frac{\sum_{(i,c)\in E(M)\times C} (Distance_{Fr,c})^2 \times \mathbb{I}\{export_{i,c} > 0\}}{\sum_{(i,c)\in E(M)\times C} \mathbb{I}\{export_{i,c} > 0\}}$$

where $E(M)$ is the set of firms that export to M countries, and C is the set of all countries.

However, since there is no systematic correlation between country size and the geographic distance from France, this correction does not affect the estimated σ_0^2 substantially. The statistical significance of the estimated σ_0^2 is reduced, and the R^2 goes down from about 90% to 75% when I do not control for differences in GDP. But all the coefficients remain highly significant (at the 1% confidence level), and the level of the coefficients does not differ much across both measures.

Except in 1986 when the estimated σ_0^2 is somewhat larger than in other years, the estimated coefficients are very stable across the different years, and across the two alternative measures of

$\sigma^2(M)$. This suggests that the results presented in Section 2.3 are robust. Moreover, given that I only allow for a single degree of freedom when estimating this relationship, Proposition 3 finds a remarkably strong support in the data.

Time-series: PROBIT regression

Probit, additional controls (aggregate trade with 3rd countries for year t-1, t-2, t-3; sales growth)

Time-series: structural estimation

The theoretical prediction from the model regarding the law of motion of the number of foreign contacts can be derived by integrating Equation (1) over $\mathbb{R}$, as presented in Section 2.4,

$$M_{i,t+1} - M_{i,t} = \gamma m_r + \frac{\gamma}{1+r} M_{i,t}$$

To structurally test for the assumptions of the model, I estimate by Ordinary Least Squares the following equation,

$$(\text{Number of new contacts})_{i,t+1} = \alpha + \beta M_{i,t} + \text{Controls}_{i,t} + \epsilon_{i,t} \tag{4}$$

where the dependent variable is the number of new countries entered by firm i between year t and $t+1$, $M_{i,t}$ is the number of countries where firm i exports at time t, and $\epsilon_{i,t}$ is a normally distributed error term. Since I do not allow for the death of contacts in the theoretical model, I focus my empirical analysis on the creation of new contacts.

The estimation of Equation (4) allows me to recover the structural parameter $r \times m = \frac{\alpha}{\beta}$.

The results of this estimation are presented in Table 7. In the different specifications, I control for various measures of the growth trajectory a firm is in. The idea is that firms that grow on the domestic market may as well expand abroad, for reasons that are orthogonal to my model of network formation. I control for different combinations and measures of domestic sales growth, domestic employment growth, and domestic investment growth at the firm level.

Firms whose domestic sales are more likely to enter new foreign markets. Employment growth has some limited but non robust positive impact on the entry into new foreign markets, whereas investment growth does not seem to have any impact on the entry into foreign markets. If anything, investment growth deters entry into foreign markets.

The estimation of the various specifications of Equation (4) give an estimate for $r \times m$ that ranges between 2 (column (1) without any control) to 5.68 (column (7)). The specification that

Table 7: Entry into new markets, robustness checks

	Dependent variable: number of new foreign markets entered by firm i between t and $t+1$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$M_{i,t}$	.182***	.165***	.166***	.167***	.167***	.165***	.165***	.167***	.165***
	(0.00027)	(0.00040)	(0.00040)	(0.00040)	(0.00040)	(0.00040)	(0.00042)	(0.00040)	(0.00042)
$Sales_{i,t+1} - Sales_{i,t}$		8.37E-8***				1.70E-7***			
		(7.52E-9)				(1.06E-8)			
$\frac{Sales_{i,t+1}}{Sales_{i,t}}$			4.95E-6**				.000111***		
			(1.81E-6)				(1.81E-5)		
$Sales_{i,t} - Sales_{i,t-1}$				3.47E-8***				3.65E-8***	
				(5.85E-9)				(8.25E-9)	
$\frac{Sales_{i,t}}{Sales_{i,t-1}}$					4.87E-6**				5.34E-5**
					(1.69E-6)				(1.68E-5)
$Emp_{i,t+1} - Emp_{i,t}$						4.71E-5***			
						(8.60E-6)			
$\frac{Emp_{i,t+1}}{Emp_{i,t}}$							-3.30E-6		
							(2.21E-5)		
$Emp_{i,t} - Emp_{i,t-1}$								5.31E-6	
								(7.59E-6)	
$\frac{Emp_{i,t}}{Emp_{i,t-1}}$									-1.35E-5
									(2.06E-5)
$Inv_{i,t+1} - Inv_{i,t}$						-6.67E-7***			
						(3.67E-8)			
$\frac{Inv_{i,t+1}}{Inv_{i,t}}$							-6.87E-9		
							(7.32E-7)		
$Inv_{i,t} - Inv_{i,t-1}$								-4.15E-08	
								(3.50E-08)	
$\frac{Inv_{i,t}}{Inv_{i,t-1}}$									2.54E-6***
									(6.80E-7)
Constant	.360***	.876***	.874***	.675***	.674***	.876***	.908***	.675***	.708***
	(0.0022)	(0.0048)	(0.0048)	(0.0048)	(0.0048)	(0.0048)	(0.0053)	(0.0048)	(0.00512)
N. obs.	743916	292773	291726	263492	262472	292773	258336	263492	232501
R^2	0.38	0.37	0.37	0.40	0.40	0.37	0.37	0.40	0.40

Notes: This table shows the results of the OLS estimation of different specifications of Equation (4) for French exporters over the period 1986 to 1992. The dependent variable is the number of new foreign markets entered by firm i between year t and $t+1$. $M_{i,t,}$is the number of foreign markets where firm i exports at time t. $Sales_{i,t}$ is aggregate domestic sales (in French Francs, in France) of firm i at time t. $Emp_{i,t}$ is the total number of employees of firm i at time t. $Inv_{i,t}$ is the total capital expenditure (in French Francs) of firm i at time t. Standards errors are clustered at the firm level. *, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.

is characterized by the most significant combination of controls, column (6), gives an estimate of 5.5. Given that an entirely different set of data is used, this is surprisingly close to $r \times m \approx 7.68$ estimated from the cross-sectional distribution of the number of foreign contacts, or 6.32 once I control for the fact that I observe the number of countries, and not directly the number of contacts of French exporters.

These findings suggest that my model of network formation is able to identify precisely the link between the time-series and the cross-section of entry of individual exporters into foreign markets.

C.4 Explaining firm level trade flows

In this section, I document several differences in the export behavior across exporters of different sizes. First, I show that conditional on exporting to a country, the average size of exports is independent of country size for small exporters, whereas it becomes positively correlated with country size only for large exporters. Second, I show evidence that very large exporters are less sensitive to distance than small exporters. I argue that both sets of evidence are qualitatively in line with the predictions of the theory.

I estimate different specifications of the standard gravity equations for the year 1992, across different types of exporters,

$$lnX_{n,c} = \alpha_n \ln GDP_c - \beta_n \ln Distance_{Fr,c} + \epsilon_{n,c} \tag{5}$$

where GDP_c is the nominal GDP of country c, $Distance_{Fr,c}$ is the geographic distance between France and country c, and $\epsilon_{n,c}$ is a normally distributed error term. The definitions of n and $X_{n,c}$ are as follows.

French exporters are ordered by the total value of their exports (in French Francs), and grouped into 10 deciles of increasing export sizes, $n = 1, \ldots, 10$, where each decile contains the same number of firms. I use three different definitions for the dependent variable $X_{n,c}$: (i) aggregate exports (in French Francs) to country c for firms in decile n, (ii) average firm level exports sales (in French Francs) to country c for firms in decile n, i.e. the intensive margin of exports to c for firms in n, and (iii) number of exporters to country c in decile n, i.e. the extensive margin of exports to c for firms in n. I estimate Equation (5) using Ordinary Least Squares separately for all three definitions of the dependent variable, for each exporter size decile, as well as for all exporters together. I therefore run a total of $3 \times (10 + 1) = 33$ different regressions.

Table 8: Country size and firm level exports

Export size decile	1	2	3	4	5	6	7	8	9	10	All
Dep. Var.	ln (aggregate exports to country c) (within decile)										
$-\ln(Distance_{Fr,c})$	1.06***	1.02***	1.15***	1.06***	1.21***	1.18***	1.36***	1.23***	1.11***	.84***	.85***
	(.14)	(.15)	(.14)	(.15)	(.16)	(.17)	(.17)	(.17)	(.16)	(.12)	(.12)
$\ln(GDP_c)$	.48***	.57***	.69***	.71***	.82***	.80***	.85***	.87***	1.00***	.91***	.91***
	(.07)	(.07)	(.07)	(.07)	(.08)	(.08)	(.08)	(.08)	(.07)	(.05)	(.05)
Obs.	85	89	92	91	92	97	96	97	101	101	101
R^2	.77	.76	.79	.78	.79	.75	.78	.77	.81	.83	.84
Dep. Var.	Intensive margin: ln (average size of firm level exports to country c) (within decile)										
$-\ln(Distance_{Fr,c})$	−.008	−.04**	.02	−.03	.11*	.12**	.21***	.18***	.21***	.31***	.14**
	(.01)	(.02)	(.04)	(.05)	(.08)	(.07)	(.07)	(.09)	(.08)	(.08)	(.08)
$\ln(GDP_c)$	.02**	.01	.03*	.06***	.12***	.09***	.15***	.14***	.24***	.31***	.26***
	(.01)	(.01)	(.02)	(.02)	(.04)	(.03)	(.03)	(.04)	(.04)	(.04)	(.04)
Obs.	85	89	92	91	92	97	96	97	101	101	101
R^2	.15	.06	.07	.12	.18	.14	.34	.24	.44	.65	.60
Dep. Var.	Extensive margin: ln (number of exporters to country c) (within decile)										
$-\ln(Distance_{Fr,c})$	1.07***	1.06***	1.13***	1.08***	1.09***	1.06***	1.15***	1.04***	.90***	.53***	.71***
	(.14)	(.14)	(.14)	(.14)	(.14)	(.15)	(.14)	(.14)	(.15)	(.11)	(.12)
$\ln(GDP_c)$	.46***	.56***	.66***	.65***	.70***	.71***	.70***	.73***	.76***	.60***	.65***
	(.07)	(.07)	(.07)	(.06)	(.07)	(.07)	(.07)	(.06)	(.07)	(.05)	(.05)
Obs.	85	89	92	91	92	97	96	97	101	101	101
R^2	.77	.77	.79	.80	.80	.77	.78	.79	.75	.72	.76

Notes: This table shows the results of the OLS estimation of the different specifications of Equation (5) for French exporters in 1992. Firms are ordered in increasing size of total exports, and grouped into 10 deciles containing a equal number of firms. The dependent variable for the top panel is the total exports to country c (in French Francs) for firms in each 10 deciles and for all firms together. The dependent variable in the middle panel is the average firm level size of exports to country c (in French Francs) for firms in each 10 deciles and for all firms. The dependent variable in the bottom panel is the number of exporters to country c for firms in each 10 deciles and for all firms. Standard errors are in parentheses.*, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.

The results for those regressions are reported in Table 8.

The top panel (rows 2 to 6) shows the regression results for aggregate exports across all 10 deciles, and for all exporters together; the middle panel (rows 7 to 11) shows the regression results for the intensive margin of trade across all 10 deciles, and for all exporters together; the bottom panel (rows 12 to 16) shows the regression results for the extensive margin of trade across all 10 deciles, and for all exporters together. The first 10 columns correspond to each decile, in increasing order of export sizes, and the last to all exporters together.

The most salient feature of the data is that aggregate trade is much more sensitive to the size of the destination country among large exporters than among small exporters. The elasticity of aggregate trade flows with respect to the size of the destination country increases approximately monotonically from around .5 for the smallest exporters to around 1 for the largest exporters.

Most of the increase in this elasticity is due to an increase of the elasticity of the intensive margin of trade with respect to country size. Whereas for the smallest 4 deciles, the elasticity of the intensive margin of trade with respect to country size is not statistically different from zero, or very close to zero, this elasticity starts increasing for the last deciles, to reach about .3 for the largest exporters. I interpret this finding as evidence that small exporters only have a single contact per market. Therefore, conditional on exporting to a given country, the value of exports is the same irrespective of the size of the destination country size. As firms acquire more contacts, it becomes more and more likely that they have more than one contact per market. Firms with many contacts will typically have more contacts in large countries than in small countries.[49] This would explain why the intensive margin of trade is insensitive to country size for small exporters, and becomes gradually more and more sensitive for larger exporters. A similar argument is presented formally in Armenter and Koren (2010).

The model I developed does not have anything to say about the (substantially smaller) increase in the elasticity of the extensive margin of trade with respect to country size. I leave the study of this empirical fact for future research.

The second most salient feature of the data is that whereas the distance elasticity of trade is large (in absolute value) for most firms, it start falling for the largest exporters. Here, most of the difference is driven by the extensive margin of trade. This finding is consistent with the

[49]Taking the limit of a firm that would have many more contacts than the total number of countries in the world, one can simply use the law of large numbers. Controlling for the impact of distance, such a firm with many contacts should have a number of contacts in any given country that is proportional to the size of the country. The total value of its exports should therefore be roughly proportional to country size.

qualitative predictions of the model. The model predicts that firms with more foreign contacts will also have foreign contacts that are geographically more dispersed, or in other words that the distance elasticity of trade should fall (in absolute value) with the number of foreign contacts. Whereas the distance elasticity of the extensive margin is above but close to 1 for most 80% of the exporters, it falls to .53 for the 10% largest exporters. Given the skewness of the distribution of the number of contacts, it is not surprising that this fall in the distance elasticity of trade only happens for the very largest exporters. Note that even within the 9^{th} decile of firms (ranked by total size of exports), the mean number of foreign markets accessed is only 6 (the median is 5). The mean number of contacts increases to 16 (median 11) for the 10^{th} decile of firms.

The model developed above has nothing to say about the (substantially smaller) increase in the distance elasticity of the intensive margin of trade.

To further investigate this fact, I run three new sets of specifications of Equation (5), for different samples of firms. First, I only consider firms that export to a single foreign market. Second, I only consider firms that export to more than one foreign market. Third, I pull all firms together. As previously, I use three different dependent variables: first, the log of aggregate exports to country c (within the sample); second, the log of the number of exporters to country c; and third, the log of the average size of firm level exports to country c.

The results are presented on Table 9. As predicted by the theory, the exports of firms that have more than one contact are less sensitive to distance than those of firms that export to a single country. All of the decrease in the distance elasticity of exports (in absolute value) comes for a reduction of the distance elasticity of the extensive margin of exports. This elasticity of the extensive margin of exports falls from 1.16 for single market to .68 for firms that export to more than one market. The distance elasticity of the intensive margin of trade remains unchanged, even though it is statistically more significant for firms that export to more than one market.

Table 9: Distance and firm level exports

Dep. var.	ln (aggregate exports to c)			ln (number of exporters to c)			ln (average firm exports to c)		
Sample	$M=1$	$M>1$	All	$M=1$	$M>1$	All	$M=1$	$M>1$	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$-\ln(Distance_{Fr,c})$	.1.29***	.82***	.85***	1.16***	.68***	.71***	.13	.14**	.14**
	(.22)	(.12)	(.12)	(.14)	(.12)	(.12)	(.14)	(.08)	(.08)
$\ln(GDP_c)$	.88***	.96***	.91***	.71***	.65***	.65***	.17***	.31***	.26***
	(.10)	(.05)	(.05)	(.06)	(.05)	(.05)	(.06)	(.03)	(.04)
Obs.	99	101	101	99	101	101	99	101	101
R^2	.71	.85	.84	.80	.75	.76	.18	.63	.60

Notes: This table shows the results of the OLS estimation of the different specifications of Equation (5) for French exporters in 1992. Firms are grouped in three samples: firms that export to a single country, in column (1), (4) and (7); firms that export to more than one country, in columns (2), (5) and (8); and all firms, in columns (3), (6) and (9). The dependent variable for the left panel (columns (1)-(3)) is the total exports to country c (in French Francs) for firms in each sample. The dependent variable in the middle panel (columns (4)-(6)) is the average firm level size of exports to country c (in French Francs) for firms in each sample. The dependent variable in the right panel (columns (7)-(9)) is the number of exporters to country c for firms in each sample. Standard errors are in parentheses. *, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.