NBER WORKING PAPER SERIES

THE NETWORK STRUCTURE OF INTERNATIONAL TRADE

Thomas Chaney

Working Paper 16753
http://www.nber.org/papers/w16753

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2011

The Network Structure of International Trade
Thomas Chaney
NBER Working Paper No. 16753
January 2011
JEL No. D85,F1

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 explain the observed cross section of firm level exports. All theoretical predictions are supported by the data.

Thomas Chaney
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
tchaney@uchicago.edu

The Network Structure of International Trade*

Thomas Chaney†

University of Chicago, NBER and CEPR

January, 2011

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 explain the observed cross section of firm level exports. All theoretical predictions are supported by the data.

Introduction

Individual firms differ hugely in their exposure to international trade. Most firms do not export abroad. Of those which do, only few 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 with 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 unexplained. Whereas Bernard, Eaton, Jensen and Kortum

---

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

(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 offer a simple explanation for the heterogeneous ability of individual firms to access foreign markets based on the formation of an international social network. Individual exporters meet foreign importers both at random and through their network of existing foreign contacts. The cross-sectional predictions of the model on both the distribution of the number of foreign contacts, and on the 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 extend the social network model of Jackson and Rogers (2007). 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 ones gives an advantage to firms with many contacts. This generates a fat tailed distribution of the number of contacts across firms. The key parameter that shapes 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 well described by the theoretical model. This allows me to estimate 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. I tie together quantitatively the cross-sectional and time-series evidence on firm level trade.

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 away from Japan as Japanese firms would. 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 expand into more remote countries, so that their exports become geographically more dispersed. The speed at which the geographic dispersion of foreign contacts increases depends on the relative importance of random versus network-based meetings. The geographic dispersion of foreign contacts 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 hits 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.

This paper contributes to the literature on both international trade and social networks.

There is a nascent literature in international trade on the role that social networks and informational barriers play in facilitating or hampering transactions. In a seminal paper, Rauch and Trindade (2002) show that the presence of ethnic Chinese networks facilitates bilateral trade. Moreover, they show that this effect is more pronounced for trade in differentiated goods than for trade in homogeneous goods. They argue that these findings are evidence for the importance of informational barriers, and that social networks mitigate those barriers. In the context of intra-national trade, Combes, Lafourcade and Mayer (2005) show that social and business networks facilitate trade between regions within France. Burchardi and Hassan (2010) show that West German regions that have closer social ties with East Germany experienced faster growth and engaged in more investment into East Germany after the German reunification. On a somewhat related topic, Hidalgo, Klinger, Barabási and Hausmann (2007) show that the product mix manufactured and exported by countries can be described as a network, and that countries move towards more connected sectors when they grow. In this paper, I develop a more general model of the formation of an international network of firms, and show how this network matters for firm level trade patterns, over and beyond the effects analyzed in the relatively narrow special cases studied so far.

This paper is also related to the recent literature that emphasises the importance of informational barriers in international trade and the role of trade intermediaries in overcoming those barriers. Antràs and Costinot (2010) develop a theoretical model of trade that relaxes the assumption of a centralized Walrasian market, and derive predictions for the welfare gains from trade in a setting where trade is intermediated. Ahn, Khandelwal and Wei (2010) demonstrate empirically the importance of trade intermediaries in facilitating trade, especially for smaller exporters and for penetrating less accessible markets. I do not formally introduce trade intermediaries, but I stress the importance of informational barriers, and show how a social network can partially overcome these barriers. The network I describe can be thought of as a formal treatment of the way in which trade intermediaries connect importers and exporters.

This paper is complementary to models of international trade with heterogeneous firms such as Bernard, Eaton, Jensen and Kortum (2003), Melitz (2003) and its extension in Chaney (2008). Those models assume that differences in the ability of individual firms to enter foreign markets are entirely driven by some exogenous productivity differences, and by the configuration of exogenous parameters that govern the accessibility of different foreign markets. These models successfully replicate a series of stylized facts regarding the size distribution of individual firms in different markets and the efficiency of firms entering different sets of countries, as shown by Eaton, Kortum and Kramarz (2010). But they take as exogenous all the parameters that govern the entry of firms into different markets. In other words, these models are successful at explaining the intensive margin of trade at the firm level, but are silent about the determinants of the extensive margin of trade. Moreover, these models are unable to match simultaneously the different stylized facts I uncover regarding the distribution of the number and the geographic location of foreign markets entered by different firms. By contrast, the model I develop offers a parsimonious explanation for the extensive margin of trade at the firm level, but is silent about the intensive margin of trade. Because this model is analytically tractable, it would be easy to combine it with a Melitz-type model, and explain both the extensive and intensive margins of trade. In that sense, this model is complementary to the models of trade with heterogeneous firms.

Finally, this paper is related to the literature on social networks. I develop a spatial extension of the model of social network in Jackson and Rogers (2007). They propose a tractable way to combine the features of a random network and a preferential network.[1] The main theoretical

---

[1] See Erdös and Rényi (1959) for a seminal description of random networks, and Barabási and Albert (1999) for a description of preferential networks that exhibit scale-free degree distributions. See Jackson (2010) for a thorough overview of models of social networks.

innovation of my model is to embed this general network into space. For the purpose of this paper, I assume that this space corresponds to the physical geographic space. But it could alternatively correspond to any other space that describes some of the attributes of the agents connected through that network. Existing models of social networks have been able to characterize the links formed by agents that differ only according to a small, discrete set of attributes.[2] By contrast, the model I develop is embedded in a large and potentially multi-dimensional space. I am able to characterize who is connected with whom along dimensions outside of the network. This technical innovation allows to treat formally various measures of social distance in social networks in a more general way than has been done so far. I also offer an empirical application of a network model to a data-set much larger than has typically been used in the social network literature.

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. The actual trade that occurs once two firms are linked can be derived in a simple extension of the Krugman (1980) model.[3] 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 network of foreign contacts, they can meet some of the contacts of their contacts.

---

[2]See McPherson, Smith-Lovin and Cook (2001) for an overview of various situations where agents tend to connect to each other according to some attributes outside of the network, which is generally described as homophily. See Bramoullé and Rogers (2010) for a recent model of network with homophily between two groups.

[3]See Appendix B page 41.

This model delivers a series of predictions that are supported by data on firm level trade. First, the model replicates 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 quantitatively matches the increase of the geographic dispersion of exports with the number of foreign markets a firm serves (see Proposition 3 and the empirical test in Section 2.3). Third, the main assumptions of the model on the dynamics of firm level trade flows are supported by the data (see the empirical tests in Section 2.4).

1.1 Set-up

The formal set-up is as follows. Firms are uniformly[4] distributed over a one-dimensional[5] infinite space, represented by $\mathbb{R}$.[6] Time is discrete. The population in each location grows at the same constant growth rate γ.[7] I normalize the population at time $t = 0$ to 1 in each location.[8] I designate a firm by the pair $(x, \tau) \in \mathbb{R} \times \mathbb{N}$, where $x \in \mathbb{R}$ is the firm's location, and $\tau \geq 0 \in \mathbb{N}$ is the firm's age. Once born, a firm never changes location. Firms form directed links with one another.[9] Neither firms nor links ever die.[10]

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, τ) at time t is described by a continuous distribution $f_{x,\tau,t}$,

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

---

[4]I show numerically in Section 1.3 that all results hold for a non-uniformly distributed population.

[5]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.

[6]Note that I consider an infinite continuous space for analytical tractability. All the results hold numerically in a finite discrete world, as shown in Section 1.3. I will however 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, τ) is known by exactly a number $f_{x,\tau,t} (y - x) \, dy$ of different firms location y (strictly speaking, in a small neighborhood dy around y), it is known by a total number of $\int_a^b f_{x,\tau,t} (y - x) \, dy$ different firms in an entire interval $[a, b]$, and by a total number of $M_{x,\tau,t}$ different firms worldwide. This language is both formally rigorous, and intuitively accessible.

[7]All the results hold if population growth is arithmetic instead of exponential.

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

[9]The links are directed in the sense that I will keep track of which firm initiated a contact. I will later assume that when a firm initiates a contact with another, it imports some amount from it. I show in Appendix B page 41 how to derive this result in a Krugman (1980) model with informational frictions.

[10]All results hold in a more general form if death hits firms and/or links with a constant Poisson probability.

so that the mass of firms located in $[a, b] \subset \mathbb{R}$ that know[11] firm (x, τ) at time t is $\int_a^b f_{x,\tau,t} (y - x) \, dy$.[12]

The *in-degree* of firm (x, τ) at time t, defined as the total mass of firms that know (x, τ) at time t, is then simply $M_{x,\tau,t}$.[13]

The process through which firms acquire both an out-degree and an in-degree distribution is 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 is trivial and of no interest. All the action takes 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. A firm in location x forms $m_r g (y - x) \, dy$ successful links with firms in location y, where g is a well defined symmetric probability density function. Except for the assumption of symmetry, the p.d.f. g can have any arbitrary shape.[14]

Second, out of the union of the out-degree distributions of all m_r firms met at random, a newly born firm samples a mass m_n of firms (where m_n stands for *network-based meetings*).[15] Trivially, all firms have the same out-degree equal to $(m_r + m_n)$.[16]

To further simplify the model, I use the following mean-field approximation. I assume that the number of links any firm receives is exactly equal to 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. I show numerically in Section 1.3 that this approximation is precise.

With this simple process for link formation, I can now describe how the in-degree distribution

---

[11] I will systematically use the expression "X knows Y" in the sense that X initiated a contact with Y, and "X is known by Y" in the sense that Y has initiated a contact with X.

[12] 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 space of the model.

[13] Note that $f_{x,\tau,t}$ is not a probability density function, since it sums up to $M_{x,\tau,t}$ which differs from 1 in general.

[14] Note that geographic distance will actually *hinder* the formation of random links only if $\partial g(|x|) / \partial |x| < 0$. 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.

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

[16] See Proposition 4 on page 38 in Appendix A for a formal derivation of the out-degree distribution. I am grateful to Enghin Atalay for solving for this distribution.

of a firm evolves.[17] Given the geographic symmetry of the network, I consider a firm located at the origin ($x = 0$) without loss of generality.[18] Moreover, with the mean-field approximation, the only state variable that matters is the age of a firm. To ease notations, I will drop the location ($x = 0$) and the time (t) subscripts from now on.

Assume that a firm of age τ has an in-degree distribution $f_\tau(\cdot)$. The following period, some newly born firms will meet firm τ at random. There are γN_t newly born firms in each location. Newly born firms located in x meet $m_r g(0 - x)\, dx = m_r g(x)\, dx$ firms in the origin. Since there are N_t firms at the origin, the number of new random meetings received by firm τ originating from location x is given by $\frac{\gamma N_t m_r}{N_t} g(x)\, dx = \gamma m_r g(x)\, dx$.

In addition, some of the newly born firms will meet firms in the in-degree distribution of firm τ and form a link with τ 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 τ and form a link with τ 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 $(m_r + m_n)$. Each firm in 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-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 $\frac{\gamma N_t m_r}{N_t} g(y - x)\, dy = \gamma m_r g(x - y)\, dx$ firms from location x. The number of firms in location y that already know τ is given by $f_\tau(y)\, dy$, inherited from the previous period. The number of firms located in x that form network-based meetings with firm τ via a firm in y is then equal to $\gamma m_r g(x - y)\, dx \times f_\tau(y)\, dy \times \frac{m_n}{m_r(m_r + m_n)}$. A network-based meeting can potentially be intermediated by firms in any location $y \in \mathbb{R}$. So the in-degree distribution of a firm of age τ evolves recursively according to,

$$f_{\tau+1}(x)\, dx = f_\tau(x)\, dx + \gamma m_r g(x)\, dx + \gamma m_r \frac{m_n}{m_r(m_r + m_n)} \int_{y \in \mathbb{R}} g(x - y) f_\tau(y)\, dy dx$$

$$\text{or } f_{\tau+1} = f_\tau + \gamma m_r g + \frac{\gamma m_n}{(m_r + m_n)} g * f_\tau \tag{1}$$

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

---

[17]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 $M_0 \geq 0$. This initial in-degree, if present, is distributed over space according to the density g. 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.

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

this distribution.

Proposition 1 *The in-degree distribution of a firm of age τ, $f_{\tau,t}$, is given by,*

$$f_\tau = \frac{m_r}{m_n}(m_r + m_n)\left(\left(\delta + \gamma\frac{m_n}{m_r + m_n}g\right)^{*\tau} - \delta\right)$$

where δ is the Dirac delta function, $*$ is the convolution product,[19] γ is the growth rate of the population, g is the geographic distribution of random contacts for newborn firms, and m_r and m_n are respectively the number of random and network-based meetings of newly born firms.

f_τ admits a closed form solution in the special cases where g is a Gaussian or a Cauchy distribution.

Proof. See Appendix A page 32. ∎

This analytical solution for the in-degree distribution of firms allows me to describe several moments of this distribution. In the next section, I analyze two of these moments.

1.2 The number and geography of contacts

Having characterized the in-degree distribution of each firm, the following Proposition 2 characterizes the distribution of the number of contacts across firms, while Proposition 3 characterizes the geographic dispersion of these contacts.[20]

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.

---

[19]The exponent $*\tau$ stands for a distribution convoluted with itself τ times.

[20]Note that this model encompasses the space-less model of Jackson and Rogers (2007). Formally, this would correspond to the case where $g = \delta$, the Dirac delta function. A formal proof of this result can be found in the proof of Proposition 1 in Appendix A page 32.

Proof. See Appendix A page 34. ■

I show in Section 2.2 that Proposition 2 matches the data on firm level exports, with random meetings accounting for roughly 60% of all new meetings.[21]

Let me briefly describe the properties of the cross sectional distribution of the number of contacts, and provide some intuition for those properties.[22] The upper tail of the distribution asymptotes to a scale-free Pareto distribution, whereas the lower tail is close to an exponential distribution.[23] Firms that already have acquired many contacts can use each and every one of those existing contacts as a bait to "fish" for new network-based contacts. By comparison, 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 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 distribution is Pareto versus exponential changes. In the polar case where almost all meetings are random $(r \to +\infty)$, the whole distribution is exponential, whereas in the opposite polar case where almost all meetings are network-based $(r \to 0)$, the distribution follows a Zipf law.[24] In cases in-between, the distribution is given by some mixture of these polar cases.

The characterization of the in-degree distribution in Proposition 1 not only allows me to characterize the behavior of the total number of contacts of a firm, but also the geography of these contacts. I define the average squared distance from firm τ's contacts, Δ_τ^2, as,

$$\Delta_\tau^2 \equiv \int_{\mathbb{R}} x^2 \frac{f_\tau}{M_\tau}(x)\, dx$$

where f_τ / M_τ is the well defined probability density function of the location of a firm's contacts. The following proposition describes the relationship between the number of contacts and the distance from these contacts.

---

[21] I estimate using firm level trade data that $m_r / m_n \approx 1.65$, so that $m_r / (m_r + m_n) \approx .6$.

[22] The dynamics for the number of contacts is close to the model in Steindl (1965). It is also related to the more elaborate models of Gabaix (1999) and Luttmer (2007).

[23] 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.

[24] See the proof of Proposition 2 in Appendix A page 34 for a formal proof of this statement.

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

$$\Delta^2(M) = \left(1 + \frac{r \times m}{M}\right) \ln\left(1 + \frac{M}{r \times m}\right) \times \Delta_0^2$$

where $\Delta_0^2 \equiv \int_{\mathbb{R}} x^2 g(x)\, dx$ is the average squared distance of random 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 37. ■

Note that all the results derived hold for any arbitrary symmetric probability density g with a finite variance.[25] Note also that I only need information about a single moment of the geographic dispersion of random contacts, Δ_0^2, to explain the dispersion of the contacts of all firms. For any two distributions g and g' that have the same Δ_0^2, the average squared distance from a firm's contact will evolve in the same way. This result will prove useful for guiding the empirical strategy, as I only need to estimate the single moment, Δ_0^2, and do not have to characterize the entire geographic distribution of exports. Moreover, I show numerically in Section 1.3 that Proposition 3 holds even if different points in space have different g distributions, with the same Δ_0^2 on average, and if space is not uniformly populated.

The reason why the geographic dispersion of a firm's contacts increases with this firm's number of contacts 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 ($r \to +\infty$), the average squared distance $\Delta^2(M)$ is constant and equal to Δ_0^2. As the relative share of network-based meetings increases (r shrinks), not only does $\Delta^2(M)$ become larger for all M's, but $\Delta^2(M)$ increases faster with M.

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 (g with average squared distance Δ_0^2). Network-based meetings follow a different spatial dynamic. Firms use their existing

---

[25]As noted earlier, the special case of zero variance ($g = \delta$, the Dirac delta function) collapses exactly to the space-less model of Jackson and Rogers.

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 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,[26] and simply calibrating the units of measurement,[27] I describe how the geographic dispersion of contacts increases as firms get more contacts.

1.3 Robustness

In this section, I show numerically that the predictions of the model are robust to removing the mean-field approximation, as well as relaxing the assumptions that the population is uniformly distributed over space and that the same g function governs random meetings for all firms.

To do so, I randomly generate a non-uniform population distribution over a finite interval.[28] The artificial world I generate features areas of various sizes (islands) separated by empty spaces of various sizes (oceans), as well as a varying population density. A growing population of firms is randomly allocated to these different locations, and forms meetings with existing firms according to the process described in Section 1.1. However, firms in different locations face different geographies: some firms are close to the boundaries of this world, some are in the middle of densely populated areas, and some are in isolated and sparsely populated islands. So firms form random meetings according to different g distributions, with the same dispersion Δ_0^2 on average only. Finally, since I consider a world with a discrete number of firms, I do not impose a mean-field approximation on the process of link formation, and actual meetings are governed in large part by luck.

In other words, using numerical simulations, I relax the mean-field approximation, as well as the assumptions of an unbounded, continuous and uniformly populated world. The results of these numerical simulations is presented in Figure 1.

The numerical simulations show that the mean-field approximation is precise,[29] and that the convenient assumptions that the world is unbounded, continuous and uniformly populated, and that all firms face the same distribution g are not crucial for the main theoretical predictions of

---

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

[27]Δ_0^2 is not a unit-free parameter.

[28]See Appendix A on page 39 for a precise description of the numerical simulation of the model.

[29]Atalay (2011) derives an analytical solution for the Jackson and Rogers (2007) model without a mean-field approximation, and shows analytically that this approximation is precise. He does not however consider the more general model embedded in geographic space presented above.
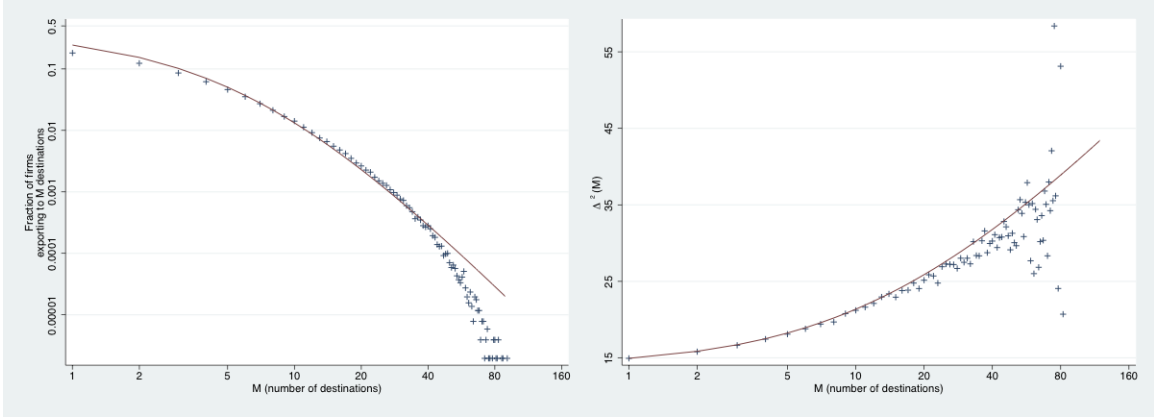
Figure 1: Numerical simulations of Proposition 2 and 3.

Notes: The simulation covers 100 periods. The solid red line in the left panel corresponds to the theoretical prediction from Propositions 2, and in the right panel to the theoretical prediction from Proposition 3. The blue plus signs corresponds to the results from the numerical simulation.

the model. The simulated cross-sectional distribution of the number of contacts is close to the theoretical prediction in Proposition 2. The geographic dispersion of contacts closely matches the theoretical prediction of Proposition 3 for most firms, but idiosyncratic noise increases as we consider firms with very many contacts. This is mostly due to the high variance of the realizations of the simulation among the small sample of well connected firms. This pattern of the simulated data resembles the actual data on firm level exports presented in Section 2.3.

1.4 Discussion

Interpretation of the model: 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 detail 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 markets. 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

13

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 that the discount in the quality of information is the same for all foreign markets.[30] 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 for that exporter to enter yet another market. The geographic dispersion of a firm's exports increases as it enters more markets.

Welfare analysis: Given the simplicity of the structure of the network that emerges, I can describe the welfare implications of this model in a variety of economic settings.[31] 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 as 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 welfare gains of those consumers that have access to many goods is not enough to dominate the welfare losses of those consumers that have access to few goods. An identical 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.

Relation to the existing firm-level trade literature: Existing international trade models with heterogeneous firms, such as Bernard, Eaton, Jensen and Kortum (2003) or Melitz (2003) and its extension in Chaney (2008) do not offer specific predictions regarding the distribution of the number of countries reached by different firms. By comparison, the network model I develop offers a parsimonious theory for the extensive margin of international trade.

---

[30]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.

[31]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.

In the original Melitz (2003) model, all trade barriers are symmetric, and any exporter exports to all foreign markets. This is obviously an artifact of the counter-factual assumption that all trade barriers are perfectly symmetric. In Chaney (2008), I offer 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 the Mth most accessible market, it will necessarily export to all markets $M' \leq M$. 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 distance between 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 make the counter factual assumption 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. I develop these arguments formally in the Appendix, and provide a calibration of the Melitz/Chaney model on the same data I use in Section 2.[32]

By adding to the Melitz/Chaney model firm-destination specific idiosyncratic shocks to the entry cost and demand faced by each firm, Eaton, Kortum and Kramarz (2010) can a priori replicate any pattern of entry in the data. Calibrating their model to the data, they need to assume a very large amount of idiosyncratic noise[33], so that the productivity thresholds are essentially randomly distributed. With the assumption of this additional noise, the fraction of firms that export to exactly M markets inherits the Pareto distribution of productivities across firms, which matches the data well. This distribution is directly assumed, and the fact that the model lines up with the data comes from the assumption of a large amount of idiosyncratic noise and of Pareto distributed productivity shocks, and not from the underlying Melitz/Chaney model.[34] In contrast, the network model I develop offers a theory of the distribution of entry into foreign

---

[32]See Appendix B.3 on page 47.

[33]The ratio of the relevant combination of fixed entry cost and local demand shock varies by a factor of 1 to 13 between the 25th and 75th percentiles.

[34]Similarly, Armenter and Koren (2010) estimate from the data the distribution of the number of shipments (the distribution of the number of "balls") from the data, and then generate predictions for the occurrence of zeroes in the trade data. By contrast, instead of assuming this distribution to match the data, my model offers a theory that generates such a distribution.

15

contacts, without the need for ad hoc assumptions on firms' productivity distribution.

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 has a low productivity, may still export to many foreign countries, if this exporter is lucky enough to face unproductive foreign competitors. 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 either 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 not monotone. The fraction of firms exporting to exactly M markets can be made arbitrarily small or large by simply varying bilateral trade barriers.

Finally, if trade barriers increase with distance, and if there is no systematic correlation between country size and distance from France, both the Melitz/Chaney model and Bernard, Eaton, Jensen and Kortum (2003) 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. Even if a large amount of noise is added as in Eaton, Kortum and Kramarz (2010), the very strong tendency of firms in the Melitz/Chaney model to first enter close by markets implies that exports are far more geographically concentrated than in the data. For instance, among firms that export to a single foreign market, the average squared distance (in thousands of km) between France and that country is 18 in the data, 16 in my calibrated model, but only 2 in the calibrated Eaton, Kortum and Kramarz (2010) model.[35][36]

---

[35]For a more intuitive interpretation of these numbers, the average distance is 3,500 km in the data versus 900 km in the calibrated Eaton, Kortum and Kramarz (2010) model.

[36]Of course, as in the Melitz/Chaney model, as firms eventually enter all countries, the difference between the

To summarize, while existing firm level trade models directly make ad hoc assumptions to match the extensive margin of trade, I develop a parsimonious model that endogenizes these assumptions. On the other hand, my model is silent about the determinants of the intensive margin of trade, or about the relation between a firm's exposure to international trade and its size in different markets, while those models make precise predictions about those. I that sense, the proposed network model is complementary to the existing firm level trade literature.

I have developed in this section a parsimonious 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 the 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 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. In doing so, I link formally the time-series and the cross-section of firm level exports.

---

data and the model shrinks. Among firms that export to the maximum observed number of countries (98), the average squared distance is 38 in the data, 43 in my calibrated model, versus 36 in the Eaton, Kortum and Kramarz (2010) model. Those numbers however are less precisely estimated, as few firms export to that many markets.

2.1 Data

To bring the model to the data, I use two sources of data.[37] First, I use firm level export data for French exporters, over the period 1986-1992. The data used come 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 to 1992, 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, collected from the Penn World Tables.[38] The distance between two countries is the population weighted geodesic distances between the main cities in both countries, which come from the CEPII.[39] Finally, I use data on aggregate bilateral trade flows between countries, which are collected from the NBER.[40]

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 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 is one main complication that arises when bringing this prediction to the data: the data on firm level exports only provide information on the number of countries a firm exports to, not the number of individual foreign firms it exports to. To circumvent this problem, I will use a simple statistical correction, following the guidance of the theoretical model. The model predicts the cross sectional distribution of the number of foreign contacts. Assuming that those contacts randomly fall into foreign countries, the model also predicts the cross sectional distribution of the

---

[37]See Appendix C.1 for a detailed description of the data.

[38]See the description of the data in `http://pwt.econ.upenn.edu/`.

[39]See the description of the data in `http://www.cepii.fr/distance/noticedist_en.pdf`.

[40]See the description of the data in Feenstra et al. (2004).

number of distinct foreign markets reached by French exporters. For the parameters estimated from the data, I find for instance that firms that export to a single foreign market have on average 1.002 foreign contacts (with a standard deviation of .0476), while firms that export to 162, the maximum observed in the data, have on average 303 foreign contacts (with a s.d. of 19.5). This predicted distribution of the number of foreign markets reached only depends on the parameters of the model, r, the ratio of random to network-based meetings, and m, the total number of meetings initiated by new firms, as well as the total number of foreign countries in the data, 210.[41]

The estimation of the parameters (r, m) is as follows. For any combination of those parameters, assuming that each French firm independently draws a number of foreign contacts from the distribution $F(M)$ in Proposition 2, and assuming that these contacts randomly fall into 210 distinct foreign countries, there is a certain likelihood of observing a given cross sectional distribution of the number of foreign countries reached by French exporters. I estimate through Maximum Likelihood the parameters (r, m) that match the data best.

The results are presented in column (1) of Table 1. 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 network-based meetings. Newborn firms form approximately $m_r = 2.4$ contacts at random versus $m_n = 1.4$ network-based contacts.[42] Figure 2 plots the empirical density of the number of foreign markets served by French exporters and the theoretical prediction. The fact that both random and network-based meetings coexist explains the curvature of the empirical density in a log-log scale.

For robustness, I also estimate the parameters (r, m) through various alternative specifications. The parameter estimates vary little across the various specifications or estimation methods. In column (2) of Table 1, I estimate the parameters through Maximum Likelihood under the simplifying assumption that a firm that exports to exactly M distinct foreign countries has exactly M different foreign contacts. In column (3), I estimate through a Non Linear Least Square procedure the parameters under the same simplifying assumption that one country corresponds to one contact.[43] Finally, in column (4), I estimate through Maximum Likelihood a simpler benchmark

---

[41]See Appendix C.2 on page 56 for a formal derivation of the distribution of the number of distinct countries where French firms export when countries and not contacts are observed.

[42]$m_n = \frac{m}{1+r} \approx \frac{3.83}{2.65} \approx 1.4$; $m_r = m - m_n \approx 3.8 - 1.4 \approx 2.4$.

[43]Specifically, I estimate through NLLS the parameters (α, r, m) in the following equation,

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

Table 1: Empirical fit of Proposition2

	(1) (MLE w. correction)	(2) (MLE)	(3) (NLLS)	(4) (MLE)
r	1.65	1.82	1.58	$+\infty$
	(.022)	(.025)	(.14)	
m	3.83	3.71	4.86	3.84
	(.019)	(.017)	(.44)	(.011)
Adj. $R^2/\log(\text{lik.})$	-282,098	-282,022	0.98	-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. All coefficients are statistically different from zero at the 1% level of significance.

model where contacts are only acquired at random, again under the same simplifying assumption that one country corresponds to one contact. I propose a likelihood-ratio test of my model against this simpler benchmark, and reject the simpler model unambiguously. I offer further robustness checks in the Appendix.[44]

The fact that the estimated parameters vary little across the different specifications is not surprising. Eaton, Eslava, Krizan, Kugler and Tybout (2010) are able to identify separately each U.S. contact of Colombian exporter. They show that 80% of Colombian firms that export to the U.S. have a single contact (buyer) there. In the context of within country trade, using detailed information on the input-output linkages between individual US firms, Atalay, Hortaçsu, Roberts and Syverson (2010) report that the average number of U.S. suppliers of U.S. firms is only marginally above 1. Assuming that firms have exactly one single contact per market seems to be a good approximation of the data.

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.

---

where ϵ is a normally distributed error term. This corresponds to the log-linearization of Proposition 2.

[44]See Table 4 on page 60 in Appendix C.3.
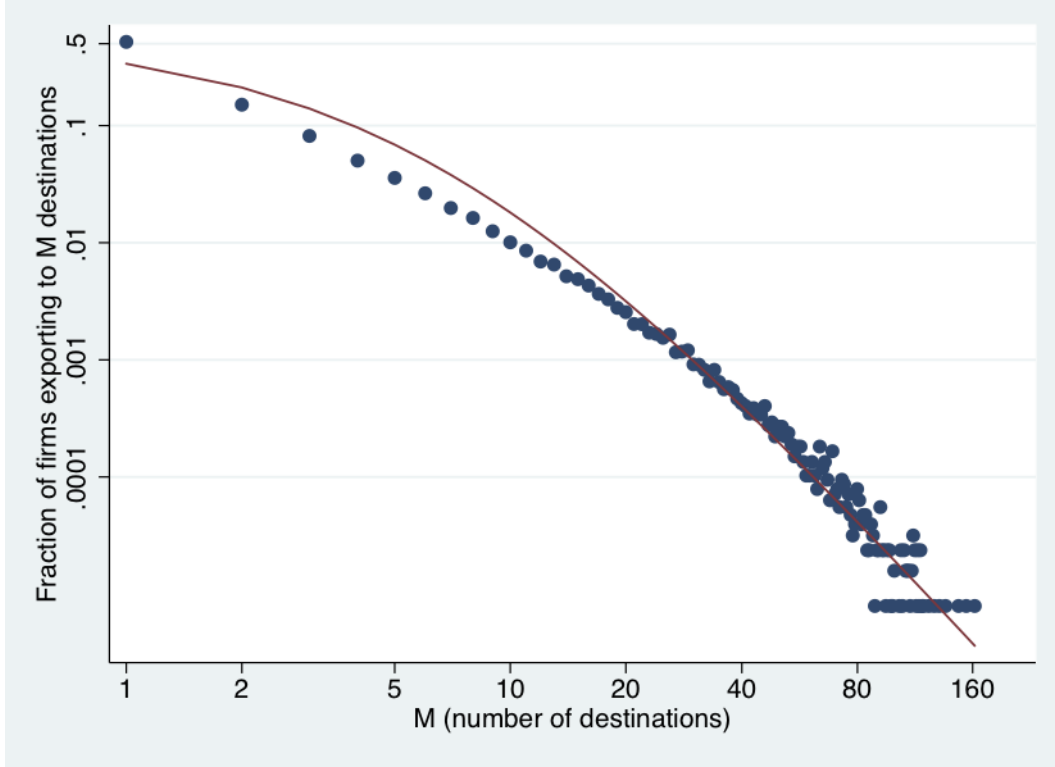
Figure 2: Empirical fit of Proposition 2, $f(M)$ versus M.

Notes: $f(M)$ is the fraction of firms exporting to M destinations; dots: data, all French exporters in 1992; line: theory. $r = 1.65$ (.022) and $m = 3.83$ (.019) are estimated through Maximum Likelihood.

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 the specific way in which 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, $\Delta^2(M)$.[45] As a reminder, the theory predicts the following relationship between the geographic dispersion of exports and the number of markets a firm is able to enter,

$$\Delta^2(M) = \left(1 + \frac{r \times m}{M}\right) \ln\left(1 + \frac{M}{r \times m}\right) \times \Delta_0^2$$

Using the cross sectional distribution of the number of export destinations across firms, I estimated in the previous section that $r \times m \approx 6.32$. I only need to calibrate Δ_0^2, which is not a unit-free

---

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

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

Notes: $\Delta^2(M)$ is the average squared distance to a firm's export destinations, among firms exporting to M destinations; dots: data, all French exporters in 1992; line: theory. $r \times m = 6.32$ is taken from the est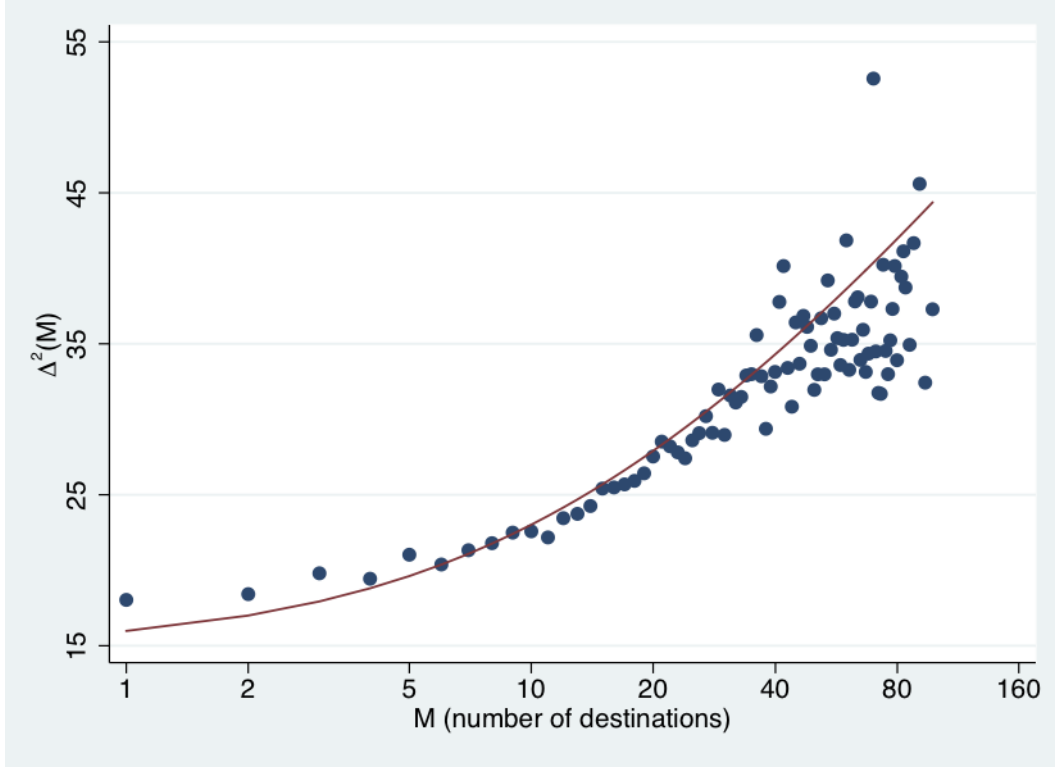imation of Proposition 2, and $\Delta_0^2 = 14.860$ (.109) is estimated through non linear least squares, each point weighted by the square root of the number of observations used to compute $\Delta^2(M)$.

measure, to bring the theoretical prediction to the data.[46] I use a non linear least square estimation of the previous equation, and recover $\Delta_0^2 \approx 14.860$ (.109), with an R^2 of 87%.[47]

Figure 3 plots the geographic dispersion of exports, $\Delta^2(M)$, as a function of the total number of foreign countries entered, both in the data and in the theory. Note that I only calibrate the intercept of this relationship. I have no other degrees of freedom that would allow me to calibrate the shape of this relationship. This shape 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.

---

[46] As discussed in Section 1.4, I do not need to rely on any specific assumption on the shape of the distribution g, except for symmetry and finite variance. The evolution of the geographic dispersion of exports, $\Delta^2(M)$, as a function of M only depends on the single moment Δ_0^2, whichever the shape of distribution g is. The multiplication by Δ_0^2 is only needed to match the initial conditions and to scale the units, as $\Delta^2(M)$ is not a unit-free measure.

[47] 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 59 for a series of robustness checks using different years and different empirical measures of $\Delta^2(M)$.

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. The proposed explanation is that through network-based contacts, exporters are able to reach deeper and deeper into geographic space. They use their faraway contacts as remote hubs to access even more distant markets. Tthe exact shape of the relationship between geographic dispersion and the 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, giving support to the theory.

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

2.4   Matching trade dynamics

In this section, I directly test in a reduced form some of the main predictions of the theoretical model regarding the time-series of entry of individual firms into foreign markets.

First, firms with more foreign contacts are more likely to enter an additional market.

Second, 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 as a proxy for the intensity of communication between those 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 increase from $t-1$ to t, then firm i is more likely to enter country c at time t, everything else being equal. I implicitly assume that if aggregate exports from c' to c increase, some firms in c' acquire new contacts in c.

Third, 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 three predictions using a Probit regression of different specifications of the following equation,

$$\mathbb{I}\{export_{i,c,t} > 0\} = \alpha \times \{\text{N. contacts}_{i,t-1}\}$$
$$+ \beta_1 \times \sum_{c' \in C_{i,t-1}} \frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}} + \beta_2 \times \sum_{c' \neq Fr} \frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}}$$
$$+ \gamma_1 \times \frac{1}{|C_{i,t-1}|} \sum_{c' \in C_{i,t-1}} \ln Dist_{c',c} + \gamma_2 \times \ln Dist_{Fr,c} + \gamma_3 \times \frac{1}{|\{c' : c' \neq Fr\}|} \sum_{c' \neq Fr} \ln Dist_{c',c}$$
$$+ \delta \times \mathbb{I}\{export_{i,c,t-1} > 0\} + \text{Controls}_{c,t} + \epsilon_{i,c,t} \quad (2)$$

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 predicts that firms with more existing contacts are more likely acquire new contacts. I expect, $\alpha > 0$.

The model predicts that firms benefit from the contacts of their contacts. I expect that $\beta_1 > 0$. Note that it is possible that some fast growing country c may see an increase in its imports from the entire 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 of country c, and expect $\beta_2 > 0$.

The model predicts that firms benefit from the location of their contacts. I expect $\gamma_1 < 0$. At the same time, distance impedes link formation directly, so that I expect $\gamma_2 < 0$. 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 $\gamma_3 > 0$.

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)
$\frac{1}{|C_{i,t-1}|}\sum_{c'\in C_{i,t-1}} \ln Dist_{c',c}$					-0.1743	-0.1381
					(0.00264)	(0.00276)
$\ln Dist_{Fr,c}$		-0.6056	-0.3802	-0.3812	-0.3689	-0.3965
		(0.00158)	(0.00109)	(0.00111)	(0.00274)	(0.00277)
$\frac{1}{|\{c':c'\neq Fr\}|}\sum_{c'\neq Fr} \ln Dist_{c',c}$					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)
$\mathbb{I}\{export_{i,c,t-1}>0\}$			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 (2) 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 (2) on page 24. Standards errors are clustered at the firm level. All coefficients are statistically different from zero at the 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)
$\frac{1}{|C_{i,t-1}|}\sum_{c'\in C_{i,t-1}} \ln Dist_{c',c}$					-0.0069	-0.0055
					(0.00011)	(0.00011)
$\ln Dist_{Fr,c}$		-0.0287	-0.0152	-0.0153	-0.0145	-0.0157
		(0.00013)	(0.00006)	(0.00006)	(0.00010)	(0.00011)
$\frac{1}{|\{c':c'\neq Fr\}|}\sum_{c'\neq Fr} \ln Dist_{c',c}$					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)
$\mathbb{I}\{export_{i,c,t-1} > 0\}$			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 (2) 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. All coefficients are statistically different from zero at the 1% level of significance.

Table 2 shows the results of the Probit estimation of different specifications of Equation (2), 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 quantitatively 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 also equal to .46%.[48]

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 contacts must have been created between those countries. However, even with this caveat in mind, the results are both 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 very 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 close to $r \approx 1.65$ estimated from the cross-sectional distribution of the number of foreign contacts across firms in Section 2.2.

To confirm the results above, I run a series of robustness checks in the Appendix.[49] For instance, using various lags of exports, I show that the data would not be consistent with a model where exporters enter countries in which they have acquired a distribution network, or sell to consumers that have a particular taste for their good, and where either the cost of a distribution network or tastes are spatially correlated.

Taking the specific functional form of the model even more seriously, I can structurally estimate

---

[48]From Equation (1) in Section 1.1, 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 adding one contact increases the growth in the number of contacts by $\gamma m_r \left(1 + \frac{M_{i,t-1}+1}{r \times m} - 1 - \frac{M_{i,t-1}}{r \times m}\right) = \frac{\gamma m_r}{r \times m}$. The average probability of entering a new country in the sample is 4.7%, the average number of contacts in the sample is 3.9, and $r \times m \approx 6.32$ from estimating Proposition 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 3.9 to 4.9 contacts is given by $\frac{\gamma m_r}{r \times m} \approx \frac{4.7\%}{6.32+3.9} \approx .46\%$.

[49]See Appendix C.3 and the results in Table 6 on page 63 for a series of robustness checks on the time-series of firm level exports.

the law of motion for the number of contacts implied by Equation (1). Integrating the in-degree distribution f_τ over $\mathbb{R}$, I get the following law of motion for M_τ, the number of foreign contacts of a firm of age τ,

$$M_{\tau+1} - M_\tau = \frac{\gamma}{1+r} M_\tau + \gamma m_r$$

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 close to $r \times m \approx 6.32$ estimated from the cross-sectional distribution of the number of foreign contacts across firms in Section 2.2.[50]

I have presented in this section direct evidence in support of the assumptions of the proposed theoretical model based on the time-series dimension of firm level French exports. This evidence is consistent with the cross-sectional evidence presented earlier. The theoretical model ties together qualitatively and quantitatively the cross-sectional and time-series of firm-level exports.

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 supported by the data on firm level exports from France. Firms acquire about 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

---

[50]Depending on the specifications, the time-series estimate of $r \times m$ ranges between 2 (without any controls) and 5.68. See Appendix C.3 and the results in Table 8 on page 66 for a series of robustness checks on this structural time-series estimation.

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. A large shock to this dynamic system would generate non trivial transitional dynamics. For example, a large disruption of trade linkages, or the rapid growth of a large country may have a long lasting impact on the world geography of trade, since (re)building 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

AHN, JaeBin, Amit K. KHANDELWAL and Shang-Jin WEI. 2010. "The Role of Intermediaries in Facilitating Trade," *Journal of International Econoimcs*, forthcoming.

ANTRÀS, Pol and Arnaud COSTINOT. 2010. "Intermediated Trade," *Quarterly Journal of Economics*, forthcoming.

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. "On the Network Structure of Production," University of Chicago, *mimeo*.

ATALAY, Enghin. 2011. "An Exact Expression for the In-degree Distribution of the Jackson and Rogers (2007) Model," University of Chicago, *mimeo*.

BARABÁSI, Albert-László and Réka ALBERT. 1999. "Emergence of Scaling in Random Networks," *Science*, 286: 509-12.

BERNARD, Andrew B., Jonathan EATON, J. Bradford JENSEN, and Samuel KORTUM. 2003. "Plants and Productivity in International Trade," *American Economic Review*, 93(4): 1268-90.

BRAMOULLÉ, Yann and Brian W. ROGERS. 2009. "Diversity and Popularity in Social Networks," CIRPEE Working Paper No. 09-03.

BURCHARDI, Konrad B. and Tarek A. HASSAN. 2010. "The Economic Impact of Social Ties: Evidence from German Reunification." Chicago Booth, *mimeo*.

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

COMBES, Pierre-Philippe, Miren LAFOURCADE and Thierry MAYER. 2005. "The Trade-Creating Effects of Business and Social Networks: Evidence from France," *Journal of International Economics*, 66(1):1-29.

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*.

ERDÖS, Paul and Alfréd RÉNYI. 1959. "On Random Graphs," *Publicationes Mathematicae*, 6:290-

7.

Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo. 2004. "World Trade Flows: 1962-2000," NBER WP 11040.

Gabaix, Xavier. 1999. "Zipf Law for Cities: an Explanation," *Quarterly Journal of Economics*, 114(3): 739-67.

Hidalgo, César A., B. Klinger, Albert-László Barabási, and Ricardo Hausmann. 2007. "The Product Space Conditions the Development of Nations" *Science*, 317(5837): 482-487.

Jackson, Matthew O. 2010. "An Overview of Social Networks and Economic Applications," *The Handbook of Social Economics*, edited by J. Benhabib, A. Bisin, and M.O. Jackson, North Holland Press.

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.

Luttmer, Erzo G. J. 2007. "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 122(3): 1103-44.

Mayer, Thierry, and Soledad Zignago. 2006. "Notes on CEPII's Distances Measures," *mimeo*.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, 27: 414-44

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

Rauch, James E. and Vitor Trindade. 2002. "Ethnic Chinese Networks in International Trade," *Journal of International Economics*, 84(1): 116-30.

Steindl, Josef. 1965. "Random Processes and the Growth of Firms," Charles Griffin, London.

Technical Appendix for:

The Network Structure of International Trade

by Thomas CHANEY, University of Chicago

A Mathematical proofs and numerical simulations

A.1 Mathematical proofs

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

Proposition 1 (reminded) *The in-degree distribution of a firm of age τ, $f_{\tau,t}$, is given by,*

$$f_\tau = \frac{m_r}{m_n}(m_r + m_n)\left(\left(\delta + \gamma\frac{m_n}{m_r + m_n}g\right)^{*\tau} - \delta\right)$$

where δ is the Dirac delta function, $*$ is the convolution product,[51] γ is the growth rate of the population, g is the geographic distribution of random contacts for newborn firms, and m_r and m_n are respectively the number of random and network-based meetings of newly born firms.

f_τ admits closed form solutions in the special cases where g is a Gaussian, Laplace, or Cauchy distribution.

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,

$$\mathcal{F}[f_{\tau+1}] = \gamma m_r \mathcal{F}[g] + \gamma\frac{m_n}{m_r + m_n}\mathcal{F}[g] \times \mathcal{F}[f_\tau] + \mathcal{F}[f_\tau]$$

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

$$\mathcal{F}[f_\tau] = \gamma m_r \mathcal{F}[g] + \left(1 + \gamma\frac{m_n}{m_r + m_n}\mathcal{F}[g]\right) \times \mathcal{F}[f_\tau]$$

for $\tau > 0$, with the initial condition at $\tau = 0$,

$$\mathcal{F}[f_0] = M_0 \mathcal{F}[g]$$

---

[51] *The exponent $*\tau$ stands for a distribution convoluted with itself τ times.*

This recurrence admits a simple solution,

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

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,

$$f_\tau = \left(\delta + \gamma\frac{m_n}{m_r + m_n}g\right)^{*\tau} * \left(\frac{m_r}{m_n}\left(m_r + m_n\right)\delta + M_0 g\right) - \frac{m_r}{m_n}\left(m_r + m_n\right)\delta$$

Taking $M_0 = 0$, I get the proposed expression for f_τ.

Note two interesting special cases where f_τ admits a closed form solution, the Normal and the Cauchy distributions. These distributions are such that the sum of two normally (Cauchy) distributed random variables is itself normally (Cauchy) distributed. Using the above formula, I get

$$\begin{aligned}
f_\tau &= \frac{m_r}{m_n}\left(m_r + m_n\right)\left(\left(\delta + \gamma\frac{m_n}{m_r + m_n}g\right)^{*\tau} - \delta\right) \\
&= \frac{m_r}{m_n}\left(m_r + m_n\right)\left(\sum_{s=0}^{\tau}\binom{N}{n}\delta^{*(\tau-s)} * \left(\gamma\frac{m_n}{m_r + m_n}g\right)^{*s} - \delta\right) \\
&= \frac{m_r}{m_n}\left(m_r + m_n\right)\left(\sum_{s=0}^{\tau}\left(\gamma\frac{m_n}{m_r + m_n}\right)^s\binom{N}{n}g^{*s} - \delta\right) \\
&= \frac{m_r}{m_n}\left(m_r + m_n\right)\sum_{s=1}^{\tau}\left(\gamma\frac{m_n}{m_r + m_n}\right)^s\binom{N}{n}g^{*s}
\end{aligned}$$

Normal distribution: using the fact that if $g = \phi_{\sigma^2}$ where ϕ_{σ^2} is the p.d.f. of Normal distribution with mean zero and variance σ^2, then $g^{*s} = \phi_{s\sigma^2}$, I get,

$$f_\tau = \frac{m_r}{m_n}\left(m_r + m_n\right)\sum_{s=1}^{\tau}\left(\gamma\frac{m_n}{m_r + m_n}\right)^s\binom{N}{n}\phi_{s\sigma^2}$$

Cauchy distribution: using the fact that if $g = \psi_\gamma$ where ψ_γ is the p.d.f. of Cauchy distribution centered around zero and with scale parameter γ, then $g^{*s} = \psi_{s\gamma}$, I get,

$$f_\tau = \frac{m_r}{m_n}\left(m_r + m_n\right)\sum_{s=1}^{\tau}\left(\gamma\frac{m_n}{m_r + m_n}\right)^s\binom{N}{n}\psi_{s\gamma}$$

Note that in the special case where distance represents an insurmountable barrier to the formation of trade linkages, which corresponds formally the case $g = \delta$, 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,

$$g = \delta \Rightarrow f_\tau = \left\{\left(1 + \gamma\frac{m_n}{m_r + m_n}\right)^\tau \times \left(\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0\right) - \frac{m_r}{m_n}\left(m_r + m_n\right)\right\} \times \delta$$

■

Lemma 1 *The total mass of contacts a firm of age τ, i.e. the in-degree of this firm, M_τ, is given by,*

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

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.

Proof. The in-degree of firm of age τ, M_τ, is defined as the integral over $\mathbb{R}$ of its in-degree distribution, f_τ,

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

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

$$M_\tau = \int_{\mathbb{R}} \left\{ \left(\delta(x) + \gamma \frac{m_n}{m_r + m_n} g(x) \right)^{*\tau} * \left(\frac{m_r}{m_n} (m_r + m_n) \delta(x) + M_0 g(x) \right) - \frac{m_r}{m_n}(m_r + m_n)\delta(x) \right\} dx$$

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

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

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

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

Taking $M_0 = 0$, I get the proposed expression for M_τ. ■

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 any time, in any give location, and therefore in the union of any given set of locations, the fraction of firms with more than M contacts is simply the fraction of firms older than $\tau(M)$, defined as $M_{\tau(M)} = M$. Using the expression for the number of contacts of a firm of age τ from Lemma 1, and leaving aside considerations of integer constraints (τ only takes integer values[52]), I can solve for $\tau(M)$,

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

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

Given the exponential growth of the population, the fraction of firms older than $\tau(M)$ is given at any time t by,

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

Using the previous expression for $\tau(M)$, I get the fraction of firms older than $\tau(M)$,

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

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 older than $\tau(M)$ is the counter-cumulative distribution of in-degrees, $\frac{N_{t-\tau(M)}}{N_t} = 1 - F_t(M)$, I get the proposed cumulative distribution function for the number of contacts across 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 $F(M)$.[53]

---

[52]Note that I am not making any continuous approximation of the discrete model. The proposed formulas are exactly correct when τ is an integer. Those formulas simply extrapolate to non integer values for τ. Being a very serious person, I never use these extrapolations.

[53]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. If there were an initial period $t = 0$ for history, whereas the function $F(M)$ is time invariant, the location of the steps for the M's would evolve through time. All the formulas above hold exactly at any point in time, without any continuous approximation.

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's 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. $\blacksquare$

Lemma 2 *The geographic dispersion of a firm's contacts increases as a firm ages. Formally, Δ_τ^2, the average squared distance from the contacts of a firm of age τ is given by,*

$$\Delta_\tau^2 = \frac{\left(1 + \gamma\frac{m_n}{m_r+m_n}\right)^\tau \left(\frac{m_r}{m_n}(m_r + m_n) + M_0\right)}{\left(1 + \gamma\frac{m_n}{m_r+m_n}\right)^\tau \left(\frac{m_r}{m_n}(m_r + m_n) + M_0\right) - \frac{m_r}{m_n}(m_r + m_n)}$$

$$\left[\tau\frac{\gamma\frac{m_n}{m_r+m_n}}{1 + \gamma\frac{m_n}{m_r+m_n}} + \frac{M_0}{\frac{m_r}{m_n}(m_r + m_n) + M_0}\right]\Delta_0^2$$

where γ is the growth rate of the population, $r = m_r/m_n$ is the ratio of random versus network-based meetings, and $\Delta_0^2 \equiv \int_\mathbb{R} x^2 g(x)\, dx$ is the average squared distance of random contacts.

Proof. Plugging the expression for the in-degree distribution from Proposition 1 into the definition of the average squared distance of a firm's contacts, Δ_τ^2, I get,

$$M_\tau\Delta_\tau^2 = \int_\mathbb{R} x^2 \left\{\left(\delta(x) + \gamma\frac{m_n}{m_r+m_n}g(x)\right)^{*\tau} * \left(\frac{m_r}{m_n}(m_r + m_n)\delta(x) + M_0 g(x)\right) - \frac{m_r}{m_n}(m_r + m_n)\delta(x)\right\} dx$$

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

$$M_\tau\Delta_\tau^2 = \left(1 + \gamma\frac{m_n}{m_r+m_n}\right)^\tau \left(\frac{m_r}{m_n}(m_r + m_n) + M_0\right)$$

$$\times \int_\mathbb{R} x^2 \left\{\left(\frac{\delta(x) + \gamma\frac{m_n}{m_r+m_n}g(x)}{1 + \gamma\frac{m_n}{m_r+m_n}}\right)^{*\tau} * \left(\frac{\frac{m_r}{m_n}(m_r + m_n)\delta(x) + M_0 g(x)}{\frac{m_r}{m_n}(m_r + m_n) + M_0}\right) - \frac{m_r}{m_n}(m_r + m_n)\delta(x)\right\} dx$$

$$- \int_\mathbb{R} x^2 \frac{m_r}{m_n}(m_r + m_n)\delta(x)\, dx$$

36

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

$$\int_{\mathbb{R}} x^2 \left(A\delta\left(x\right) + Bg\left(x\right) \right) dx = B\Delta_0^2 \text{ for any } \left(A, B\right) \in \mathbb{R}^2$$

From this intermediate result, I get,

$$\int_{\mathbb{R}} x^2 \frac{\delta\left(x\right) + \gamma \frac{m_n}{m_r+m_n} g\left(x\right)}{1 + \gamma \frac{m_n}{m_r+m_n}} dx = \frac{\gamma \frac{m_n}{m_r+m_n}}{1 + \gamma \frac{m_n}{m_r+m_n}} \Delta_0^2$$

$$\int_{\mathbb{R}} \frac{\frac{m_r}{m_n}\left(m_r + m_n\right)\delta\left(x\right) + M_0 g\left(x\right)}{\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0} dx = \frac{M_0}{\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0} \Delta_0^2$$

$$\text{and } \int_{\mathbb{R}} x^2 \frac{m_r}{m_n}\left(m_r + m_n\right)\delta\left(x\right) dx = 0$$

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,

$$M_\tau \Delta_\tau^2 = \left(1 + \gamma \frac{m_n}{m_r + m_n}\right)^\tau \left(\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0\right) \left[\tau \frac{\gamma \frac{m_n}{m_r+m_n}}{1 + \gamma \frac{m_n}{m_r+m_n}} + \frac{M_0}{\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0}\right] \Delta_0^2$$

Plugging in the expression for the total number of contacts of a firm from Lemma 1, I get the proposed expression for Δ_τ^2,

$$\Delta_\tau^2 = \frac{\left(1 + \gamma \frac{m_n}{m_r+m_n}\right)^\tau \left(\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0\right)}{\left(1 + \gamma \frac{m_n}{m_r+m_n}\right)^\tau \left(\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0\right) - \frac{m_r}{m_n}\left(m_r + m_n\right)}$$

$$\left[\tau \frac{\gamma \frac{m_n}{m_r+m_n}}{1 + \gamma \frac{m_n}{m_r+m_n}} + \frac{M_0}{\frac{m_r}{m_n}\left(m_r + m_n\right) + M_0}\right] \Delta_0^2$$

■

Proposition 3 (reminded) *The geographic dispersion of a firm's contacts increases with the firm's number of contacts. For a population growth rate γ small, the average squared distance from a firm's contacts for a firm with M contacts, $\Delta^2\left(M\right)$, is given by,*

$$\Delta^2\left(M\right) = \left(1 + \frac{r \times m}{M}\right) \ln\left(1 + \frac{M}{r \times m}\right) \times \Delta_0^2$$

where $\Delta_0^2 \equiv \int_{\mathbb{R}} x^2 g\left(x\right) dx$ is the average squared distance of random contacts, $r = m_r / m_n$ is the ratio of random versus network based meetings, and $m = \left(m_r + m_n\right)$ is the total number of contacts made by newly born firms.

Proof. From the proof of Proposition 2 I get a relation between τ and M_τ,

$$\tau = \ln\left(\frac{M_\tau + \frac{m_r}{m_n}(m_r + m_n)}{M_0 + \frac{m_r}{m_n}(m_r + m_n)}\right) \bigg/ \ln\left(1 + \gamma\frac{m_n}{m_r + m_n}\right)$$

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

$$\Delta_\tau^2 = \frac{M_\tau + \frac{m_r}{m_n}(m_r + m_n)}{M_\tau}\left[\tau\frac{\gamma\frac{m_n}{m_r+m_n}}{1 + \gamma\frac{m_n}{m_r+m_n}} + \frac{M_0}{\frac{m_r}{m_n}(m_r + m_n) + M_0}\right]\Delta_0^2$$

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

$$\Delta^2(M) = \frac{M + \frac{m_r}{m_n}(m_r + m_n)}{M}$$

$$\times \left[\frac{\gamma\frac{m_n}{m_r+m_n}}{\left(1 + \gamma\frac{m_n}{m_r+m_n}\right)\ln\left(1 + \gamma\frac{m_n}{m_r+m_n}\right)}\ln\left(\frac{M + \frac{m_r}{m_n}(m_r + m_n)}{M_0 + \frac{m_r}{m_n}(m_r + m_n)}\right) + \frac{M_0}{\frac{m_r}{m_n}(m_r + m_n) + M_0}\right]\Delta_0^2$$

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

$$\Delta^2(M) = \frac{M + \frac{m_r}{m_n}(m_r + m_n)}{M}\left(\ln\left(\frac{M + \frac{m_r}{m_n}(m_r + m_n)}{M_0 + \frac{m_r}{m_n}(m_r + m_n)}\right) + \frac{M_0}{M_0 + \frac{m_r}{m_n}(m_r + m_n)}\right)\Delta_0^2$$

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

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

$$f_{out} = m_r\sum_{k=0}^{\infty}\left(\frac{m_n}{m_r + m_n}\right)^k g^{*(k+1)} = \mathcal{F}^{-1}\left[\frac{m_r\mathcal{F}[g]}{1 - \frac{m_n}{m_r+m_n}\mathcal{F}[g]}\right]$$

Proof. Consider a newborn firm located in the origin. This firm forms m_r contacts at random, distributed over $\mathbb{R}$ according to the p.d.f. g. These contacts therefore follow the distribution $m_r g$. This firm then forms m_n network-based contacts with the contacts met at random. The contacts met at random by that firm have themselves met a fraction $\frac{m_r}{m_r+m_n}$ of their own contacts at random, distributed according to g. These network-based contacts therefore follow the distribution $m_n\frac{m_r}{m_r+m_n}g * g = m_r\frac{m_n}{m_r+m_n}g^{*2}$... etc. Using simple algebra and Fourier transforms, I get the proposed formula. ∎

A.2 Numerical simulations

In this section, I describe the procedure used to generate the numerical simulation presented in Section 1.3.

First, I generate artificially a finite and non uniformly populated world.

The world is a segment with coordinates from -300 to +300. Locations, or "cities", are located two grid points apart. The density of population in each location x is proportional to the number $d(x)$, generated in the following way:

- I start with $d(0) = 1$ at the center of the world.

- Going from one grid point x to the grid point $x+1$ directly to the East, I set $d(x+1) = d(x)$ with probability .9.

- With probability .1, I set $d(x+1)$ as the realization of a Poisson(1) random variable.

- I use the same procedure to build the world West of the origin.

- Once the whole world is populated, I randomly add or substract population until the average of the $d(x)$'s is equal to 1.

Figure 4 shows the realization of one such procedure. As can be seen, this artificial world is made of "islands", or "continents" of different sizes, separated by empty spaces, "oceans". Moreover, in each "islands", there are densely populated cities as well as sparsely populated ones.

Second, I allocate firms to the different locations according to the following procedure. In year 1, there are $12 \times d(x)$ newborn firms in each location x. Population then grows in each location at a constant growth rate of 5% per year, for 100 years. Each period, I round up the number of newborn firms in each location to the nearest integer.

Third, firms form directed links with each other, according to the process described in Section 1.1. To circumvent integer constraints, I use the following procedure. Each year, 5 out of every 12 newborn firms in each location form 3 links with existing firms at random, and 2 network-based links; the remaining 7 out of every 12 newborn firms forms 2 links at random, and 1 network-based link. This procedure an average $m_r = 2.4167$ and $m_n = 1.4167$, which corresponds to $(m, r) \approx (3.83, 1.71)$ which is close to $(m, r) \approx (3.83, 1.65)$ estimated in the data. Finally, I assume that $g \sim N(0, 16)$. From this Normal distribution, the probability that a given random link is formed in one particular location is given by the mass of the Normal distribution on that
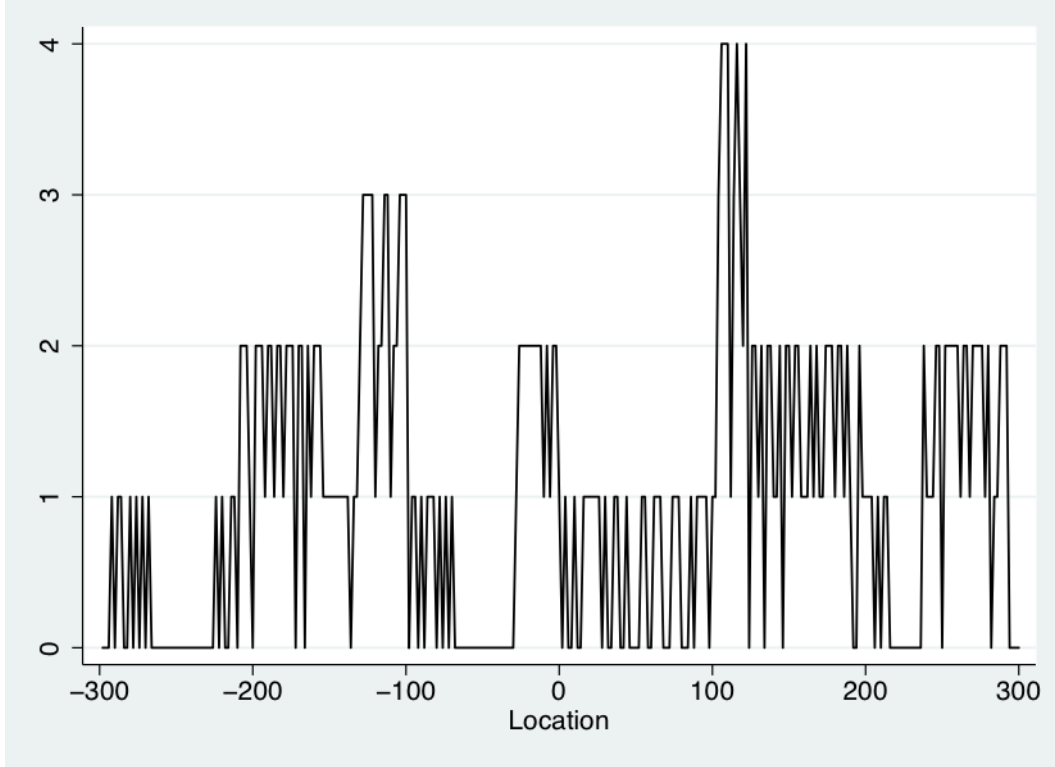
Figure 4: A non-uniform artificial world.

grid point multiplied by the population density $d(x)$ in that location. To compensate for the truncation of the upper and lower tails, all probabilities are scaled by a constant to ensure they sum up to one.

I also simulate the model relaxing only the mean-field approximation, and the assumption of a continuous unbounded world, but keeping the assumption that the world is uniformly populated. Figure 5 shows that the theoretical predictions of Proposition 2 and 3 are even more precise when the world population is uniformly distributed over space. In unreported simulations, I show that the precision of those propositions is unaffected by changes in the underlying g distribution.[54]

---

[54]I simulate the model using for the g distribution various Normal distributions with different variances, as well as Laplace distributions with different variances. Changing the variance of the g distribution or the fatness of its tails does not affect the results substantially.

Figure 5: Numerical simulations of Proposition 2 and 3, uniform world.

Notes: The simulation covers 100 periods. The solid red line in the left panel corresponds to the theoretical prediction from Propositions 2, and in the right panel to the theoretical prediction from Proposition 3. The blue plus signs corresponds to the results from the numerical simulation.
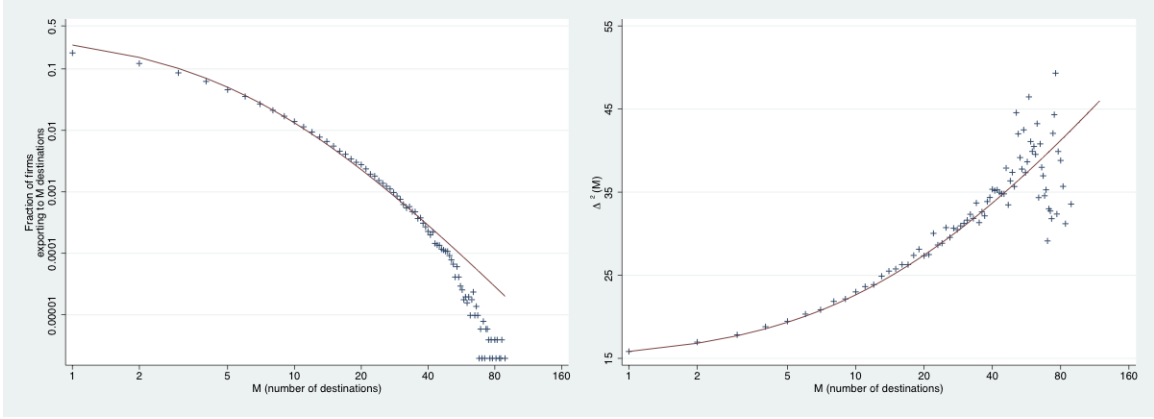
B   Additional economic assumptions and discussions

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,

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

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

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

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

41

above), and consumers can only buy from firms they know (through the same network). More precisely, each consumer has access to a masss 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.[55] 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 normalize to 1. Without loss of generality, I can also normalize $\beta = \frac{\sigma-1}{\sigma}$ so that $p = 1$.

Facing those prices, if consumer i 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 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.

---

[55]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.

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, where the welfare gain of the very connected agents dominates the welfare loss of the less connected agents.

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 predictions on the intensive margin of sales, i.e. the average sales per consumer. More precisely, in expectation, the consumers reached by any firm have access to the same number of goods, irrespective of the number of consumers 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 $G\left(p_H\right)$. Conditional on receiving a

45

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,[56]

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

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. g.

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 client is similar to the game above. The initial client makes a take-it-or-leave-it offer to each of her suppliers 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, and 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.

---

[56]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 $p_H^* = 2$.

Each period, a supplier will therefore meet a fraction $G(p_H^*)$ of the clients of each of her own clients. 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, firms 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 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 exporting 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 export 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 argument 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 markets, 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 $M = 1, \ldots, 6$:

49
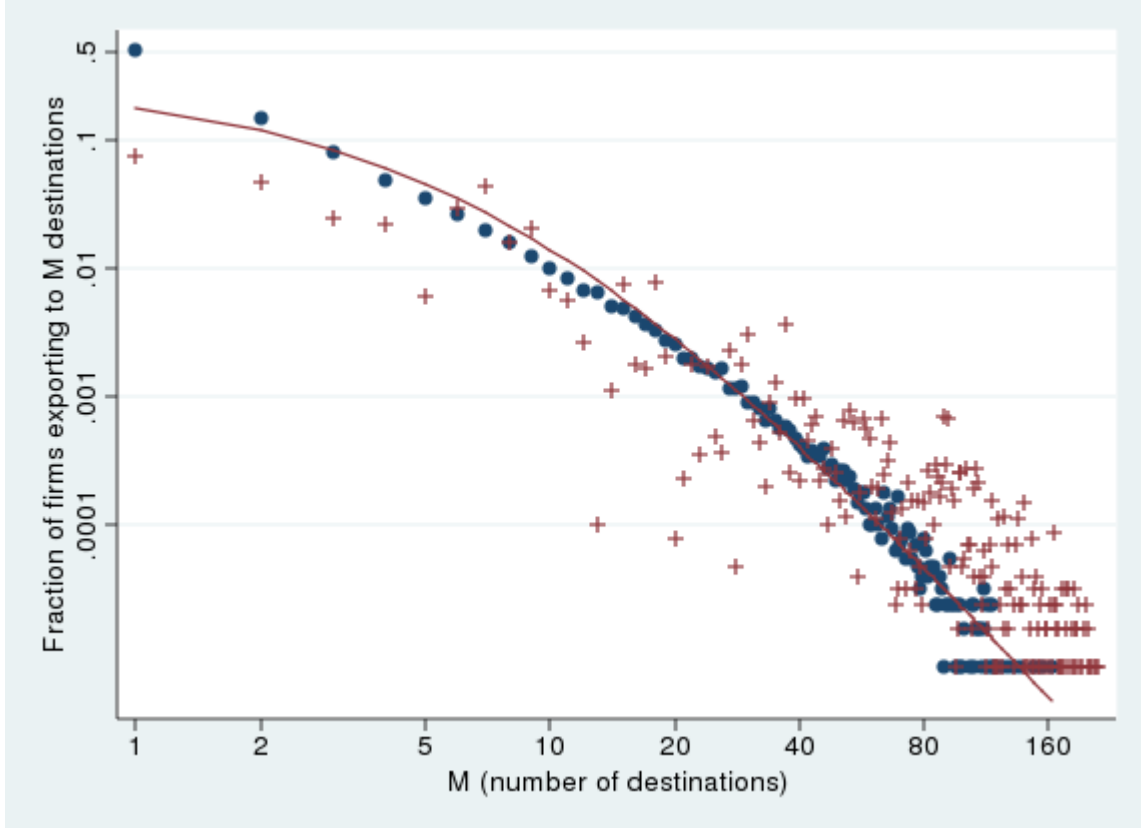
Figure 6: Network versus Melitz/Chaney model, $f(M)$ versus M.

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.

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

Evidently, the theoretical predictions of the Melitz/Chaney 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 Mth and $(M+1)^{th}$ market. As can be seen visually on Figure 6, 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. 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 predictions regarding the shape of the relationship between geographic dispersion and the number of markets accessed. To illustrate this point, I calibrate the geographic dispersion of exports, $\Delta^2(M)$, in the Melitz/Chaney model. I control for country size in the same way as I do when constructing the empirical measure of $\Delta^2(M)$ presented in Appendix C.2 on page 58. 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 $\frac{N_1}{GDP_1} > \frac{N_2}{GDP_2} > \frac{N_3}{GDP_3} > \dots$ 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,

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

As can be seen visually on Figure 7, 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 empirical fit of the Melitz/Chaney model would be substantially worse.[57]

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.

---

[57]Calibration available from the author upon request.
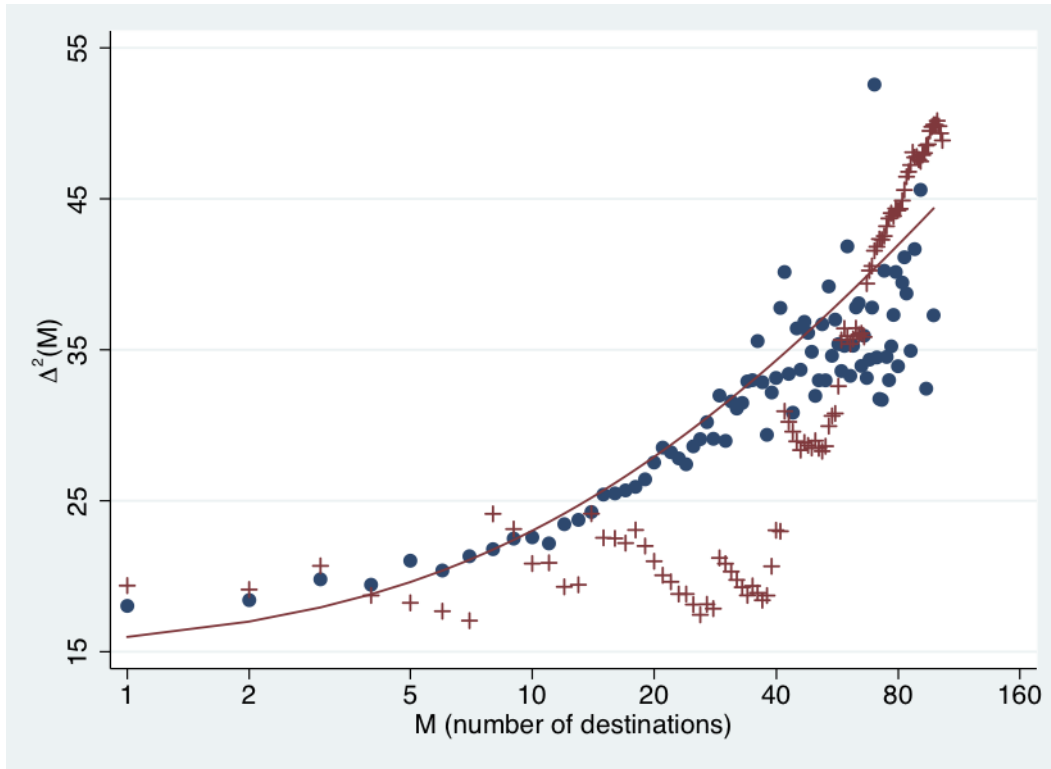
Figure 7: Network versus Melitz/Chaney model, $\Delta^2(M)$ versus M.

Notes: $\Delta^2(M)$ 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.

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)

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. $f(M)$: 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 $z_1 > \max\{\tau z_2, \tau z_3\}$ 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, $f(3)$, is given by,

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

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 productivity z_1 sells in exactly two markets if either $\{z_1/\tau > z_2$ and $z_1/\tau < z_3 < z_1\}$ or $\{z_1/\tau > z_3$ and $z_1/\tau < z_2 < z_1\}$ are true. The respective probabilities of each of those mutually exclusive events are,

$$\Pr\{z_1/\tau > z_2 \text{ and } z_1/\tau < z_3 < z_1\} = \frac{T_1}{T_1 + \tau^\theta T_2 + T_3} - \frac{T_1}{T_1 + \tau^\theta T_2 + \tau^\theta T_3}$$

$$\Pr\{z_1/\tau > z_3 \text{ and } z_1/\tau < z_2 < z_1\} = \frac{T_1}{T_1 + T_2 + \tau^\theta T_3} - \frac{T_1}{T_1 + \tau^\theta T_2 + \tau^\theta T_3}$$

The fraction of firms from 1 that export to exactly two markets, $f(2)$, 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 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 come from the French customs, and are described in greater detail 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).[58] 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.

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 detail at `http://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 correspond 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 detail in Feenstra et al. (2004).

---

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

C.2 Data construction

In this section, I describe and discuss in detail 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 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.[59]

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 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 nth country. There are n equiprobable such cases,

---

[59]In the model, contacts are spatially correlated, so that some configurations of contacts are more likely than others. Not accounting for this spatial correlation may induce a downward bias in the estimated number of contacts. Despite this, I keep the assumption of equiprobable configurations for tractability. Numerical simulations suggest that this bias is not systematically different for firms with many or few contacts, so that the estimation of the parameters does not suffer much from this simplifying assumption.

one for each nth 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 export to at most M different countries. But for any $M > 1$, the probability that two different contacts fall into the same country is positive, and 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 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,

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

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 $\widehat{\Delta^2(M)}$ 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, $\Delta^2(M)$. The construction of this measure is as follows.

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

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

where GDP_c is country c's GDP, $(Distance_{Fr,c})^2$ is the squared distance (in 1,000's of km's) between France and country c, $\mathbb{I}\{export_{i,c}>0\}$ 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.[60] $\widehat{\Delta^2(M)}$ is simply the empirical second moment of this distribution f_M. Note that I have a large number of observations to calculate each $\widehat{\Delta^2(M)}$. 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

---

[60] f_M is the distribution f_τ for τ such that $M_\tau = M$.

small countries.[61] 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 replicate the estimations in columns (1)-(3) in Table 1 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 Maximum Likelihood after controlling for the fact that I observe countries and not contacts. In addition, I estimate the parameters using Maximum Likelihood and Non Linear Least Squares under the simplifying that firms have a single contact per country where they export.

The results are presented in Table 4. All estimated parameters are statistically different from zero at the 1% level of significance.[62] 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: $\Delta^2(M)$ versus M

To check the robustness of the results presented in Section 2.3, I estimate the relevant parameter Δ_0^2 using different empirical measures of $\Delta^2(M)$ and different samples of firms.

The results are presented in Table 5. For each year between 1986 and 1992, I separately estimate the parameter Δ_0^2. To do so, I use the formula for $\Delta^2(M)$ derived in Proposition 3, imposing the parameter $r \times m$ which is estimated on the cross-sectional distribution of the number of foreign

---

[61]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.

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

Table 4: Empirical fit of Proposition 2, robustness checks

Year	(1986)	(1987)	(1988)	(1989)	(1990)	(1991)	(1992)
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)
$\log(\text{lik.})$	-285,711	-297,839	-268,454	-280,670	-270,464	-277,844	-282,098
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)
$\log(\text{lik.})$	-285,653	-297,774	-268,388	-280,602	-270,395	-277,770	-282,022
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

Notes: This table shows the results of the Non Linear Least Square and Maximum Likelihood estimations of the parameters (r, m) in Proposition 2 for French exporters in all years from 1986 to 1992. The top panel estimates by maximum likelihood the p.d.f. $\varphi(n)$ using the correction proposed in Appendix C.2 on page 56 for the fact that the number of countries, n, and not contacts is observed. The middle panel estimates by maximum likelihood the p.d.f. $f(M)$ under the simplifying assumption that the number of contacts of a firm, M, is equal to the number of foreign countries. The bottom panel uses the log of the fraction of firms that export to M markets as the dependent variable, and estimates (r, m) using a non linear least squares estimation of Proposition 2. Standard errors are reported in parentheses. All coefficients on this table are statistically different from zero at the 1% level of significance.

contacts in the top panel of Table 4. I estimate Δ_0^2 using non linear least squares, weighting each observation by the precision of the empirical estimate of $\Delta^2(M)$. I use two alternative empirical measures for $\Delta^2(M)$.

The top panel of Table 5 uses the same empirical measure of the second moment of the distance from a firm's export destinations, $\Delta^2(M)$ as in Section 2.3, where I correct for differences in GDP across countries. The formula for calculating the empirical counterpart of $\Delta^2(M)$ is as follows,

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

where $E(M)$ is the set of firms that export to M countries, and C is the set of all countries. 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 5, I do not correct for differences in GDP across countries. The formula for calculating $\Delta^2(M)$ is as follows,

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

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 Δ_0^2 substantially. The statistical significance of the estimated Δ_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 coefficients themselves do not differ much across both measures. The estimated coefficients are very stable across the different years, and across the two alternative measures of $\Delta^2(M)$. 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

To check the robustness of the results presented in Section 2.4, I estimate different specifications of Equation (3). Specifically, I run a Probit regression of different specifications of the following

Table 5: Empirical fit of Proposition 3, robustness checks

Year	(1986)	(1987)	(1988)	(1989)	(1990)	(1991)	(1992)
			$\Delta^2(M)$ corrected for GDP differences:				
Δ_0^2	14.434	14.325	14.602	14.850	15.01	14.819	14.860
	(.124)	(.121)	(.118)	(.117)	(.103)	(.085)	(.109)
Adj. R^2	0.859	0.888	0.872	0.881	0.887	0.924	0.871
			$\Delta^2(M)$ not corrected for GDP differences:				
Δ_0^2	14.437	14.178	14.412	14.730	14.439	14.346	14.384
	(.292)	(.318)	(.335)	(.311)	(.360)	(.401)	(.373)
Adj. R^2	0.820	0.792	0.784	0.801	0.766	0.736	0.753

Notes: This table shows the non linear least square estimate of Δ_0^2 from Proposition 3, imposing the parameter $r \times m$ estimated by MLE in the top panel of Table 4. The estimation is run separately for each year from 1986 to 1992. The top panel corrects the empirical measure of $\Delta^2(M)$ 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 $\Delta^2(M)$. Standard errors are in parentheses. All coefficients on this table are statistically different from zero at the 1% level of significance.

equation,

$$
\begin{aligned}
\mathbb{I}\{export_{i,c,t} > 0\} = {} & \alpha \times \{\text{N. contacts}_{i,t-1}\} \\
& + \beta_1 \times \sum_{c' \in C_{i,t-1}} \frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}} + \beta_2 \times \sum_{c' \in C_{i,t-2}} \frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-2}} + \dots \\
& + \gamma_1 \times \frac{1}{|C_{i,t-1}|} \sum_{c' \in C_{i,t-1}} \ln Dist_{c',c} + \gamma_2 \times \frac{1}{|C_{i,t-2}|} \sum_{c' \in C_{i,t-2}} \ln Dist_{c',c} + \dots \\
& + \mu \times \ln Dist_{Fr,c} + \nu \times \frac{1}{|\{c' : c' \neq Fr\}|} \sum_{c' \neq Fr} \ln Dist_{c',c} \\
& + \delta_1 \times \mathbb{I}\{export_{i,c,t-1} > 0\} + \delta_2 \times \mathbb{I}\{export_{i,c,t-2} > 0\} + \dots \\
& + \text{Controls} + \epsilon_{i,c,t} \quad (3)
\end{aligned}
$$

Table 6 shows the results of the Probit estimation of different specifications of Equation (3), and Table 7 shows the marginal effects of these regressions. With a few exceptions, all coefficients are statistically significant (at the 1% confidence level), and of the expected signs.

The results of these regressions allow me to reject an alternative model. In this alternative model, firms either have to acquire a local distribution network to enter foreign countries, or sell goods to consumers with locally differentiated tastes. Moreover, the cost of expanding a

Table 6: Time-series of exports, robustness checks (PROBIT)

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: $\mathbb{I}\left\{export_{i,c,t} > 0\right\}$				
N. contacts$_{i,t-1}$	.0343***	.0293***	.0275***	.0261***	.0259***
	(.0003)	(.0003)	(.0003)	(.0003)	(.0004)
$\sum_{c' \in C_{i,t-1}} \frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t-1}}$	.0355***	.0406***	.0560***	.0363***	.0359***
	(.0028)	(.0036)	(.0043)	(.0053)	(.0067)
$\sum_{c' \in C_{i,t-2}} \frac{\Delta\, Exports_{c',c,t-1}}{Exports_{c',c,t-2}}$		.0083***	.0001	-.0070***	-.0115***
		(.0011)	(.0015)	(.0019)	(.0027)
$\sum_{c' \in C_{i,t-3}} \frac{\Delta\, Exports_{c',c,t-2}}{Exports_{c',c,t-3}}$			.0198***	.0111***	.0059***
			(.0015)	(.0019)	(.0027)
$\sum_{c' \in C_{i,t-4}} \frac{\Delta\, Exports_{c',c,t-3}}{Exports_{c',c,t-4}}$				.0237***	.0183***
				(.0018)	(.0024)
$\sum_{c' \in C_{i,t-5}} \frac{\Delta\, Exports_{c',c,t-4}}{Exports_{c',c,t-5}}$					.0832***
					(.0025)
$\frac{\sum_{c' \in C_{i,t-1}} \ln Dist_{c',c}}{|C_{i,t-1}|}$	-.1477***	-.955***	-.0720***	-.0509***	-.0372***
	(.0032)	(.0041)	(.0053)	(.0067)	(.00086)
$\frac{\sum_{c' \in C_{i,t-2}} \ln Dist_{c',c}}{|C_{i,t-2}|}$		-.0740***	-.0582***	-.0470***	-.0432***
		(.0039)	(.0054)	(.0067)	(.00089)
$\frac{\sum_{c' \in C_{i,t-3}} \ln Dist_{c',c}}{|C_{i,t-3}|}$			-.0491***	-.0458***	-.0459***
			(.0050)	(.0067)	(.0089)
$\frac{\sum_{c' \in C_{i,t-4}} \ln Dist_{c',c}}{|C_{i,t-4}|}$				-.0437***	-.0364***
				(.0062)	(.0085)
$\frac{\sum_{c' \in C_{i,t-5}} \ln Dist_{c',c}}{|C_{i,t-5}|}$					.202***
					(.0083)
$\ln Dist_{Fr,c}$	-.3806***	-.2923***	-.2592***	-.2449***	-.2127***
	(.0031)	(.0032)	(.0037)	(.0045)	(.0055)
$\frac{1}{|\{c' : c' \neq Fr\}|} \sum_{c' \neq Fr} \ln Dist_{c',c}$	.4397***	.3621***	.3387***	.3393***	.2714***
	(.0075)	(.0078)	(.0091)	(.0113)	(.0140)
$\ln GDP_{c,t}$	.1296***	.1154***	.1124***	.1114***	.1146***
	(.0008)	(.0008)	(.0009)	(.0010)	(.0013)
Control for: $\mathbb{I}\left\{export_{i,c,t-k} > 0\right\}$ and $\ln\left(\frac{\text{Domestic Sales}_t}{\text{Domestic Sales}_{t=k}}\right)$	$k = 1$	$k = 1, 2$	$k = 1, ..., 3$	$k = 1, ...4$	$k = 1, ...5$
Constant	-3.244***	-3.0227***	-3.0218***	-3.0942***	-2.7457
	(.0543)	(.0544)	(.0616)	(.0748)	(.0915)
N. obs	16,565,725	12,255,332	8,898,165	5,950,983	3,689,058
N. clusters	34,588	32,589	29,745	23,149	20,488
Pseudo-R^2	.5608	.5935	.6073	.6136	.6171

63

Notes: This table shows the results of the PROBIT estimation of Equation (3) for a panel of all French exporters between 1986 and 1992. See the description of the variables in Section 2.4 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 7: Time-series of exports, robustness checks (PROBIT: marginal effects)

Dep. Var.: $\mathbb{I}\{export_{i,c,t} > 0\}$	dy/dx				
	(1)	(2)	(3)	(4)	(5)
N. contacts$_{i,t-1}$	.0016***	.0014***	.0013***	.0012***	.0012***
	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
$\sum_{c'\in C_{i,t-1}} \frac{\Delta\,Exports_{c',c,t}}{Exports_{c',c,t-1}}$	.0017***	.0019***	.0026***	.0016***	.0016***
	(.0001)	(.0002)	(.0002)	(.0002)	(.0003)
$\sum_{c'\in C_{i,t-2}} \frac{\Delta\,Exports_{c',c,t-1}}{Exports_{c',c,t-2}}$		.0004***	0	-.0003***	-.0005***
		(.0001)	(.0001)	(.0001)	(.0001)
$\sum_{c'\in C_{i,t-3}} \frac{\Delta\,Exports_{c',c,t-2}}{Exports_{c',c,t-3}}$			.0009***	.0005***	.0003***
			(.0001)	(.0001)	(.0001)
$\sum_{c'\in C_{i,t-4}} \frac{\Delta\,Exports_{c',c,t-3}}{Exports_{c',c,t-4}}$				.0011***	.0008***
				(.0001)	(.0001)
$\sum_{c'\in C_{i,t-5}} \frac{\Delta\,Exports_{c',c,t-4}}{Exports_{c',c,t-5}}$					.0038***
					(.0004)
$\frac{\sum_{c'\in C_{i,t-1}} \ln Dist_{c',c}}{|C_{i,t-1}|}$	-.0070***	-.0045***	-.0033***	-.0023***	-.0017***
	(.0002)	(.0002)	(.0002)	(.0003)	(.0004)
$\frac{\sum_{c'\in C_{i,t-2}} \ln Dist_{c',c}}{|C_{i,t-2}|}$		-.0035***	-.0027***	-.0021***	-.0020***
		(.0002)	(.0002)	(.0003)	(.0004)
$\frac{\sum_{c'\in C_{i,t-3}} \ln Dist_{c',c}}{|C_{i,t-3}|}$			-.0023***	-.0021***	-.0021***
			(.0002)	(.0003)	(.0004)
$\frac{\sum_{c'\in C_{i,t-4}} \ln Dist_{c',c}}{|C_{i,t-4}|}$				-.0020***	-.0017***
				(.0003)	(.0004)
$\frac{\sum_{c'\in C_{i,t-5}} \ln Dist_{c',c}}{|C_{i,t-5}|}$					.0009***
					(.0001)
$\ln Dist_{Fr,c}$	-.0181***	-.0136***	-.0120***	-.0110***	-.0097***
	(.0001)	(.0001)	(.0002)	(.0002)	(.0003)
$\frac{1}{|\{c':c'\neq Fr\}|} \sum_{c'\neq Fr} \ln Dist_{c',c}$	.0209***	.0169***	.0156***	.0153***	.0124***
	(.0003)	(.0004)	(.0004)	(.0005)	(.0006)
$\ln GDP_{c,t}$	.0062***	.0054***	.0052***	.0050***	.0052***
	(.0000)	(.0000)	(.0000)	(.0000)	(.0001)
Control for: $\mathbb{I}\{export_{i,c,t-k} > 0\}$ and $\ln\left(\frac{\text{Domestic Sales}_t}{\text{Domestic Sales}_{t=k}}\right)$	$k=1$	$k=1,2$	$k=1,...,3$	$k=1,...4$	$k=1,...5$

Notes: This table shows the marginal effects for the PROBIT estimation of Equation (3) presented in Table 6. 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.

distribution network or local tastes are spatially correlated. In such an alternative model, a firm would be more likely to enter at time t a market c that is geographically close to, or that trade a lot with any market c' in which this firms exported at time $t-1$. However, in such an alternative model, conditional on exporting to countries that are either close or trades a lot with c at time $t-1$, the increment in the probability of entering c given that a firm exports to countries that are close or trade a lot with c at times $t-2$, $t-3$,... should decrease rapidly. In my model on the other hand, any year a firm exports to country c', and conditional on still not exporting to c, it has the same probability of learning about contacts in c. So controlling for the export status in c over longer and longer lags, the increment in probability of entering c brought about by the fact that a firm exports to countries that are close or trade a lot with c should not systematically decrease over longer and longer lags.

As can be seen in Columns (2)-(5), there is no systematic tendency for the coefficients $\beta_1, \ldots, \beta_5$ and $\gamma_1, \ldots, \gamma_5$ on these lags to fall. In Column (5) where I use the maximum number of five lags allowed by the data, the coefficients on the lags of export growth, β's, fall and then increase, while the coefficients on the lags of distance, γ's, increase, except for the fifth lag that is insignificant.

It must be noted however that there is a lot of turnover in the extensive margin of trade at the firm level, so that the systematic entry and exit of firms into different markets may obscure some of these results.

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.

Table 8: Time-series of exports, structural estimation of Proposition 1

	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.

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 8. In the different specifications, I control for various measures of a firm's growth trajectory. 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 increasing 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 controls) to 5.68 (column (7)). The specification that 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 $r \times m \approx 6.32$ estimated from the cross-sectional distribution of the number of foreign contacts.

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