The Network Structure of International Trade

Thomas Chaney

American Economic Review 2014, 104(11): 3600–3634
http://dx.doi.org/10.1257/aer.104.11.3600

The Network Structure of International Trade†

By THOMAS CHANEY*

Motivated by empirical evidence I uncover on the dynamics of French firms' exports, I offer a novel theory of trade frictions. Firms export only into markets where they have a contact. They search directly for new trading partners, but also use their existing network of contacts to search remotely for new partners. I characterize the dynamic formation of an international network of exporters in this model. Structurally, I estimate this model on French data and confirm its predictions regarding the distribution of the number of foreign markets accessed by exporters and the geographic distribution of exports. (*JEL* D85, F11, F14, L24)

This paper proposes a new theory of the frictions associated with international trade, and more generally the frictions that affect the ability of firms to trade with each other. Samuelson (1954) and later Krugman (1980) recognized the key importance that trade frictions play not only in shaping the patterns of international trade, but also in determining relative factor prices between countries, and ultimately comparative development. Despite the central role they play in trade models, trade frictions remain largely unexplained, and we only have a very crude formalization of those frictions. Samuelson (1954), Krugman (1980) and most of the trade literature assume "iceberg"-type trade costs, a simple proportional cost. Melitz (2003); Helpman, Melitz, and Rubinstein (2008); and Chaney (2008) recognize the importance of the extensive margin of trade in determining firm level and aggregate flows, and introduce a fixed cost in addition to the usual iceberg cost. Arkolakis (2010) further endogenizes this fixed cost and allows firms to choose from a menu of fixed costs. Yet this simple combination of a fixed and a variable cost is too crude to capture many facts about firm-level exports. Whereas Bernard et al. (2003) or Melitz (2003) assume that differences in the ability of firms to enter foreign markets are entirely driven by heterogeneous productivities, Armenter and Koren (forthcoming) point out that productivity differences can only account for a fraction of the

* Toulouse School of Economics, 21 Allee de Brienne, 31000 Toulouse, France (e-mail: thomas.chaney@gmail.com). I am grateful to Enghin Atalay, Sylvain Chassang, Xavier Gabaix, Sam Kortum, Pierre-Louis Lions, Bob Lucas, Marc Melitz, Roger Myerson, David Sraer, Nancy Stokey, and seminar participants at Chicago (Math and Econ), Columbia, Harvard, MIT, the NBER Summer Institute, NYU, Princeton, Sciences Po (Paris), Toronto, the Toulouse School of Economics, UBC Vancouver, UQAM, UW Milwaukee, Wharton, and Yale for helpful discussions, and NSF grant SES-1061622 for financial support. I am indebted to Ferdinando Monte and Enghin Atalay for their superb research assistance. I declare that I have received in the last three years more than US$10,000 from the National Science Foundation, in the form of a research grant (SES-1061622) for the particular research topic addressed in this paper. Beyond this grant, I have no relevant or material financial interests that relate to the research in this paper. First draft: July 2010.
† Go to http://dx.doi.org/10.1257/aer.104.11.3600 to visit the article page for additional materials and author disclosure statement(s).

exposure to international markets. Similarly, Eaton, Kortum, and Kramarz (2011) show that a large amount of idiosyncratic noise has to be added to the simple combination of fixed and variable costs of the Melitz model in order to empirically match firm level exports from France.

The main contribution of this paper is to develop a theory of trade frictions based on the notion of informational frictions. This theory is motivated by new stylized facts I uncover on the dynamics of firm-level exports in France. The second contribution of this paper is to build a dynamic model of trade frictions. While there are strong patterns in the dynamics of firm-level trade, most existing trade models are static in essence.[1] By adding structure to the export dynamics, I generate predictions linking the cross section and the time series of international trade. The third contribution of this paper is explicitly to account for the geography of trade. Geography, measured as the physical distance between countries, plays a crucial role in explaining the empirical patterns of international trade, yet it is absent from most trade models.[2] I show how to introduce geographic space into a theoretical model of firm-level trade, and provide precise empirical evidence in support of the model. I focus primarily on the physical distance between locations. The reason is both that this is a measure that is easy to calculate and that this is the measure that is empirically most relevant to explain trade flows. Putting together those three contributions—the notion that information is a key friction to trade, its corollary that the diffusion of information will follow an intrinsically dynamic process, and the fact that geography matters for trade in a specific way—this paper offers a very different perspective on international trade compared to traditional models.

Before describing the related literature, I will spell out quickly the main intuition from the model, as well as the main predictions that I bring to the data.

Potential exporters meet foreign trading partners in two distinct ways. First, a firm searches directly for foreign partners, which I model as a geographically biased random search. Second, once a firm has acquired a network of foreign contacts in various foreign locations, it can search remotely for new trading partners from these locations. Those two assumptions are motivated by novel empirical evidence on the dynamics of firms exports I uncover using data on French firms from 1986 to 1992. The more countries a firm exports to, the more likely it is to enter new market subsequently. Moreover, *where* a firm exports to affects which specific markets it will enter in the future: if a French firm exports to country a in year t, it is then more likely to enter in year $t + 1$ a country b geographically close to a, even if b is not close to France. The possibility to use existing contacts to find new ones gives an advantage to firms with many contacts. This generates a fat-tailed distribution for the number of foreign contacts across firms. The empirical distribution of the number of foreign contacts is well described by the theory.

A more elaborate contribution of this paper accounts for geographic space. Remote search allows say a French exporter that has a acquired a contact in Japan

[1] Dixit (1989); Krugman (1987); and Young (1991) are among a few early and notable exceptions, as are a few recent papers mentioned below in this introduction.

[2] There are a few important exceptions in the economic geography literature (for instance, Fujita, Krugman, and Mori 1999). However, this literature is primarily theoretical, and rarely goes beyond testing a few stylized empirical facts, if any. Desmet and Rossi-Hansberg (2010) identify some of the challenges of introducing space in an equilibrium model, and show some empirical evidence. See Allen and Arkolakis (forthcoming) for a recent contribution.

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. By acquiring more foreign contacts, firms expand into more remote countries and, as a result, export over longer distances. Empirically, the geographic distance of exports increases with the number of foreign contacts as the theory predicts.

This is a theory of a network. Therefore, a shock that hits anywhere will be transmitted throughout the network, with an intensity that depends on the structure of the network. The data confirms this prediction. For instance, I show that for a French firm which already exports to a, the probability that it begins exporting to b will be higher following an increase in the trade volume between a and b, all else equal.

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

There is a nascent literature in international trade and macroeconomics on the role that informational barriers and informational networks play in facilitating or hampering transactions, and in transmitting shocks. In a seminal paper, Rauch (1999) conjectures that informational barriers play an important role. He offers a classification of traded goods between differentiated and homogeneous goods, and shows that geographic proximity is more important for trade in differentiated goods. He argues that this is evidence for the importance of informational barriers. While the Rauch classification has been used widely in international trade, the notion that informational networks are important in overcoming informational barriers has remained relatively underexplored. I offer a formal treatment of the network that allows information to diffuse, and show evidence of this network using firm-level trade data. Rauch and Trindade (2002) show that the presence of ethnic Chinese networks facilitates bilateral trade, and particularly so for trade in differentiated goods. They argue that these findings are evidence for the importance of informational barriers, and that social networks mitigate those barriers. Rauch (2001) offers a survey of the literature on networks in international trade. In the context of intranational trade, Combes, Lafourcade, and Mayer (2005) show that social and business networks facilitate trade between regions within France, where they use migrations and multiplant firms to infer a measure of social and business linkages. Using Spanish data, Garmendia et al. (2012) show that social and business networks have a stronger impact on the extensive margin than on the intensive margin of trade, a prediction that holds in my model. Burchardi and Hassan (2013) show that West German regions which have closer social ties with East Germany inherited from the tumultuous history of refugees relocations after WWII experienced faster growth and engaged in more investment into East Germany after the German reunification. 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 special cases studied so far.

On a somewhat related topic, Hidalgo et al. (2007) show that the product mix of goods manufactured and exported by countries can be described as a network, and that countries move toward more connected sectors as they grow. Acemoglu et al. (2012) describe the input-output linkages between sectors in the United States as a network, and show how idiosyncratic shocks to individual sectors have a nonnegligible impact on aggregate volatility. The results I present on the transmission of aggregate trade shocks on firm exports suggest that similar forces may be at play in trade.

This paper is also related to a recent literature which emphasizes the role of trade intermediaries in overcoming informational barriers. Casella and Rauch (2002) offer a formal model of trade with informational barriers. They assume that there are only two types of agents: some are perfectly informed about the quality of foreign goods, while the others are uniformed. The informed agents may chose to act as intermediaries for international trade. I offer a more nuanced model where firms gradually learn about foreign markets, so that there is close to a continuum of firms with a differential access to information about foreign markets. Antràs and Costinot (2011) 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 (2011) 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 network can partially overcome these barriers. The network I describe can be thought of as a formal treatment of how intermediaries connect importers and exporters.

This paper is complementary to models of international trade with heterogeneous firms such as Bernard et al. (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 driven entirely by some exogenous productivity differences, and by the configuration of exogenous parameters which govern the accessibility of different foreign markets. These models replicate successfully 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 (2011). While successful at explaining the intensive margin of firm-level trade, 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 mostly silent about the intensive margin of trade. In that sense, this model is complementary to the existing models of trade with heterogeneous firms.

This paper is also complementary to a recent literature on the dynamics of exports or more generally expansion at the firm level. Albornoz et al. (2012) and Defever, Heid, and Larch (2010) both present simple models of learning about a firm's potential in a foreign market. They show evidence of the sequential entry into foreign markets of Argentine and Chinese exporters, respectively, meaning that where a firm already exports influences where it enters next. Morales, Sheu, and Zahler (2013) use a moment inequality estimation procedure to estimate a similar model of sequential export choice, and document that exports tend to be history dependent. They stress the importance of what they call *extended gravity*, which is the fact that if a firm exports to a particular country, it is subsequently more likely to export to other similar countries. This corresponds to the notion of remote search in my model. In the case study of a single firm, Jia (2008) and Holmes (2011) study the geographic expansion of Wal-mart in the United States. Both stress the importance of local complementarities. New Wal-mart outlets tend to benefit from the proximity of its existing retail centers. Local complementarities are similar to the notion of remote search in my paper, and the expansion of this single firm is similar to the expansion of exporters in my model. My paper is complementary to those papers, in the sense that I incorporate these observations formally into a theoretical model of the dynamics of entry of firms. I show how to analyze the properties of this model in a tractable way. And I show formally how the dynamics of firm-level exports shape both the cross-sectional distribution of exports as well as the time series of exports at the firm level. By going further into solving a theoretical model, I extract empirical predictions which are easier to test.

Finally, this paper is indirectly related to the literature on social networks. While there is no explicit notion of social ties in my model, the formal treatment of firm linkages resembles the analysis of the social network literature. Jackson and Rogers (2007) propose a tractable way to combine the features of a random network and a preferential network. The notions of direct and remote search in my model are similar to their notions of random and preferential attachment. The main theoretical innovation of my model is to embed this general network into an arbitrary space. For the purpose of this paper, I assume that this space corresponds to the physical geographic space. It could alternatively correspond to any other space that describes some of the attributes of the agents connected through that network.[3] Bramoulle et al. (2012) consider a model with a finite number of types that are biased against each other. They show that over time, agents increase the diversity of their contacts, in the sense that they get connected with different types. They derive conditions under which an agent's initial bias asymptotically vanishes. As the notion of a bias between types is similar to the notion of geographic distance between firms in my model, their results are comparable to the gradual geographic expansion of exports in my model. The technique used in those papers for finitely many types is complementary to the approach for infinitely many types I use: while I can model a large number of types, I have to impose an assumption of symmetry that these authors relax. Those more general assumptions however limit them to results with only two types, or to only monotonicity and asymptotic results with more than two types. I also offer an empirical application of a network model to a dataset much larger than has typically been used in the social network literature.

I present reduced-form evidence on the dynamics of firms exports in Section I, build a theory motivated by this evidence in Section II, and structurally estimate the theory in Section III.

I. Reduced-Form Evidence on Trade Dynamics

In this section, I present reduced-form evidence that individual firms follow a history-dependent process when expanding into foreign markets. In particular, I show that a firm which exports to more countries is more likely to enter new markets subsequently. More interestingly, *where* a firm currently exports affects which new markets it enters subsequently: if a firm exports to country c' at time t, it is subsequently more likely to enter any country c that is closely connected to country

[3] 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.

c', either in the sense that it is geographically close to c', or that it trades a lot with c'. This reduced-form evidence motivates the theory presented in the next section.

Data Sources.—I use two sources of data.[4] First, I use firm-level export data for French exporters, over the period 1986–1992. The data come from the same source as the data used by Eaton, Kortum, and Kramarz (2011). For each firm and each year, I use information on the set of countries to which a firm exports. There are between 115,000 (in 1988) and 122,000 exporters (in 1987) in my sample (121,581 in 1992). Those firms export to a total of 103 different foreign countries for which I have additional information on size and location. French firms export on average to between 3.49 (in 1991) and 3.62 (in 1986) different foreign countries (3.50 in 1992).

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 each other, and aggregate bilateral trade between country pairs. The size of a country is measured as nominal gross domestic product (GDP), collected from the Penn World Tables. The distance between two countries is the population-weighted geodesic distances between the main cities in both countries, collected from the CEPII. Finally, I use data on aggregate bilateral trade flows between countries, collected from the NBER.

Regression Specification.—Formally, I estimate a Probit regression of different specifications of the following equation:

$$\begin{aligned}(1)\qquad &\Pr(export_{i,c,t+1} > 0\,|\,observables)\\ &= \Phi\Bigg(\alpha \sum_{c'} \mathbf{1}[export_{i,c',t} > 0] \qquad + \beta_1 g(Dist_{France,c})\\ &\quad + \beta_2 \frac{\sum_{c'} \mathbf{1}[export_{i,c',t} > 0]\, g(Dist_{c',c})}{\sum_{c'} \mathbf{1}[export_{i,c',t} > 0]}\\ &\quad + \beta_3 \frac{\sum_{c' \neq Fr} g(Dist_{c',c})}{N_{c' \neq France}} + \gamma_1 \sum_{c'} \frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t}}\\ &\quad + \gamma_2 \sum_{c'} \mathbf{1}[export_{i,c',t} > 0]\, \frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t}}\\ &\quad + \delta\, \mathbf{1}[export_{i,c,t} > 0] + Controls_{c,t}\Bigg),\end{aligned}$$

where Φ is the c.d.f. of the standard normal distribution; $\mathbf{1}[export_{i,c,t+1} > 0]$ takes the value 1 if firm i exports to country c at time t and 0 otherwise; $Dist_{c',c}$ is the distance between countries c' and c; $N_{c' \neq France}$ is the number of countries excluding

[4] Further details about the data sources are provided in Appendix A.

France in my sample; and $\frac{\Delta\, Exports_{c',c,t}}{Exports_{c',c,t}}$ is the growth of aggregate exports from country c' to country c between years t and $t+1$. The downward-sloping function g governs how the proximity between countries c' and c is related to the geographic distance between them. I consider the following two specifications for the function g:[5]

$$g(Dist_{c',c}) \;=\; \begin{cases} 1/Dist_{c',c} \\ e^{-Dist_{c',c}/3.5} \end{cases}. \tag{2}$$

Coefficients Interpretation.—The coefficient α controls for impact of the number of countries a firm exports to on the likelihood it enters new markets subsequently. $\alpha > 0$ would mean that the more markets a firm exports to today, the more likely it is to enter new markets in the future. The coefficient β_1 controls for the direct impact of proximity on trade: for both specifications of the function g in equation (2), the term $g(Dist_{France,c})$ is larger for a country c that is geographically closer to France. $\beta_1 > 0$ would mean that proximity has a beneficial effect on entry, in the sense that a firm is more likely to enter close-by markets than remote ones. The coefficient β_2 controls for the *indirect* impact of proximity on trade: the term $\frac{\sum_{c'}\mathbf{1}[export_{i,c',t} > 0]\, g(Dist_{c',c})}{\sum_{c'}1[export_{i,c',t} > 0]}$ measures the average proximity between the countries toward which firm i already exports in year t and country c. $\beta_2 > 0$ would mean that if a firm exports to countries which are close to c, it is subsequently more likely to enter that country c. The coefficients γ_1 and γ_2 are analogous to β_1 and β_2, except that the proximity between two countries is not measured by their physical distance, but by how much trade between them increases. $\gamma_1 > 0$ would mean that the faster a country's imports grow, the more likely it is that any firm enters that country. $\gamma_2 > 0$ would mean that if a firm already exports to countries whose exports to c grow, it is subsequently more likely to enter that country. Finally, the coefficient δ controls for the export status of firm i in the previous year, and the possibility that a firm loses foreign contacts. I expect α, β_1, β_2, γ_1, $\gamma_2 > 0$ and $0 < \delta \leq 1$.

I control for country size, since firms are mechanically more likely to export to a large country than to a small one. I also add controls for the sector in which a firm operates, as firms in different sectors may be more or less likely to export to any particular country. Removing the sector fixed effects does not affect the results materially. Replacing the flexible $g(Dist)$ function by country fixed effects does not affect the results materially either. Finally, it is likely 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 milder, and all else equal, it will be easier to access c. In order not to bias the estimated direct impact of distance (β_1), the coefficient β_3, expected to be negative, controls for this remoteness measure.

Results.—Table 1 shows the marginal effects from the Probit estimation of different specifications of equation (1). Standard errors are clustered at the firm level.

[5] I take the number 3.5 in $e^{-x/3.5}$ from the SMM (simulated method of moments) estimate of the theory presented in the next section.

TABLE 1—NUMBER AND LOCATION OF CONTACTS AND TRADE BETWEEN THIRD COUNTRIES PREDICT ENTRY

Dependent variable: $\mathbf{1}[export_{i,c,t+1} > 0]$	dy/dx						
		— $g(x) = 1/x$ —			— $g(x) = e^{-x/3.5}$ —		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\sum_{c'} \mathbf{1}[export_{i,c',t} > 0]$	0.0048	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016
	(0.00003)	(0.00001)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
$g(Dist_{France,c})$		0.1255	0.1437	0.1310	0.0936	0.1437	0.0969
		(0.0007)	(0.0005)	(0.0007)	(0.0005)	(0.0005)	(0.0005)
$\frac{\sum_{c'}\mathbf{1}[export_{i,c',t} > 0]\, g(Dist_{c',c})}{\sum_{c'}\mathbf{1}[export_{i,c',t} > 0]}$		0.0333		0.0281	0.0456		0.0433
		(0.0006)		(0.0007)	(0.0005)		(0.0005)
$\frac{\sum_{c' \neq Fr} g(Dist_{c',c})}{\lvert\lvert c' \neq France \rvert\rvert}$		−0.0773		−0.0752	−0.0334		−0.0283
		(0.0037)		(0.0037)	(0.0010)		(0.0011)
$\sum_{c'} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t}}$			0.0028	0.0028		0.0034	0.0050
			(0.0002)	(0.0002)		(0.0001)	(0.0002)
$\sum_{c'}\mathbf{1}[export_{i,c',t} > 0] \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t}}$			0.0034	0.0033		0.0029	0.0027
			(0.0001)	(0.0001)		(0.0002)	(0.0001)
$GDP_{c,t}$		0.009	0.009	0.009	0.009	0.010	0.010
		(0.00004)	(0.00004)	(0.00004)	(0.00005)	(0.00005)	(0.00004)
$\mathbf{1}[export_{i,c,t} > 0]$		0.4196	0.4403	0.4220	0.4002	0.4403	0.4023
		(0.0013)	(0.0013)	(0.0014)	(0.0013)	(0.0013)	(0.0014)
Sector fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	— 21,884,616 in (1); 21,603,426 in (2) and (5); 20,857,435 otherwise —						
Firms	— 35,412 in (1); 34,957 otherwise —						
Years	— 6 —						
Destinations	— 103 —						
Pseudo-R^2	0.1262	0.5529	0.5494	0.5499	0.5560	0.5493	0.5528

Notes: This table shows the marginal effects for the PROBIT estimation of equation (1) 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 $t+1$. The description of the explanatory variables is given along with equation (1). 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 percent level of significance.

In every specification, all coefficients are statistically significant (at the 1 percent confidence level), and of the expected signs.

A firm which currently exports to more countries $\left(\sum_{c'}\mathbf{1}[export_{i,c',t} > 0] \text{ larger}\right)$ is more likely to enter yet another country. The increment in the probability of entering any country from having an extra export destination ($\alpha \approx 0.0048$ in column 1 of Table 1) is of the same order of magnitude as the unconditional probability of entering any country in my sample (0.0047). This effect is statistically and economically significant.

More interestingly, the actual existing portfolio of a firm's export destinations shapes its future expansion into new markets. For instance, firm i is likely to enter country c if c is geographically close to countries where firm i already exports ($\beta_2 \approx 0.0333$ in column 2), or if c has experienced an increase in its imports from those countries ($\gamma_2 \approx 0.0034$ in column 3). This is true even after controlling for the fact that firm i is likely to export to any country that is close to France ($\beta_1 \approx 0.1255$ in column 2), or to any country experiencing an increase in imports, because it has been growing fast, for instance ($\gamma_1 \approx 0.0028$ in column 3).

Of special interest is the size of the coefficient δ. δ measures the persistence of a firm's exports to a particular country. Across the various specifications of equation (1), δ is around 40 percent. This implies that every year, a firm has a 60 percent chance of exiting a country where it is currently exporting. This large number implies a high degree of churning in exports, an observation that resonates with the findings in Eaton et al. (2010). The estimated δ is of course the marginal effect across a heterogeneous set of exporters, so it may hide a large amount of heterogeneity.

To conclude, I find reduced-form evidence that individual firms follow a history-dependent process which governs their gradual entry into foreign markets. I isolate two stylized facts. First, the more countries a firm exports to today, the more likely it is to enter yet other countries in the future. Second, *where* a firm exports today affects where that firm will export in the future: all else equal, if a firm exports to countries that are close to country c, it is more likely to enter that country c in the future. Motivated by those stylized facts, I now build a theoretical model of firm-level export dynamics.

II. A Dynamic Model of Exports

In this section, I develop a model of the sequential entry of firms into foreign markets that incorporates the stylized facts uncovered earlier in the paper. I show that a model which features a history-dependent process for exporting generates strong predictions not only for the time series of exports, but also for the cross-section of exports.

A. *Setup*

Space.—$\mathcal{S}$ is a discrete set of locations. I will consider several alternatives for the set $\mathcal{S}$. I start with a presentation of the theory without imposing any restriction on $\mathcal{S}$. I then fully solve the model for the special case $\mathcal{S} = \mathbb{Z}$. I use this special case to illustrate the key forces of the model. I finally turn back to a more general setup where $\mathcal{S} \neq \mathbb{Z}$. Using numerical simulations, I show that the results derived in the special case $\mathcal{S} = \mathbb{Z}$ offer a good approximation of what happens in more general cases, and provide a useful guidance for the structural estimation of the model.

Firms.—In each location $x \in \mathcal{S}$, there is a finite set of firms. Those firms sell their output to consumers in various locations. Time is discrete, and the number of firms in each location grows at a constant rate γ.

Search Frictions.—In the absence of any frictions, all firms would sell to all consumers in every location. I assume instead that firms face the following matching frictions.[6] Every period, a firm acquires new consumers in two distinct ways. First, the firm searches for new consumers locally, meaning that the search originates from where the firm itself is located. This first direct search corresponds to β_1, $\gamma_1 > 0$ in

[6] I develop in the online Appendix a simple extension of the Krugman (1980) model which endogenizes those assumptions.

the reduced-form evidence presented in Table 1. Second, the firm uses its existing network of consumers to search remotely, meaning that the search originates from where the existing consumers are located. This second remote search corresponds to $\beta_2, \gamma_2 > 0$ in the reduced form evidence presented in Table 1. It captures the idea of local externalities as in the case of the geographic expansion of Wal-mart in Jia (2008) and Holmes (2011), or in the case of Chilean exporters in Morales, Sheu, and Zahler (2013). It may either correspond to the technological constraint on the expansion of a distribution network as in Holmes (2011); to the cost of customizing a product for local tastes and requirements as in Morales, Sheu, and Zahler (2013); or more generally to the notion that exporting entails some amount of traveling and communicating with business partners, so that a firm which exports to a location y will acquire some knowledge about y and its surrounding locations.

Note that this is a model of the extensive margin of trade only. To fix ideas, think of the firm as an intermediate input producer, and its consumers as other downstream firms, potentially in other locations. I model explicitly how this firm over time sells to more consumers in more locations, but I do not model how much it sells to each of them. Superimposing a model for the intensive margin of sales is left for future research.

Before describing the dynamic acquisition of consumers formally, it is useful to introduce a few notations. Consider firm i of age t in a location which I arbitrarily call the origin. It has a network of consumers in various locations. The total number of consumers of firm i is $m_{i,t}$, distributed in various locations. I call $f_{i,t}(x)$ the number of consumers firm i has in location x,

$$f_{i,t} : \mathcal{S} \to \mathbb{N} \text{ with} \sum_{x \in \mathcal{S}} f_{i,t}(x) \equiv m_{i,t},$$

so that $\sum_{x \in \mathcal{A}} f_{i,t}(x)$ is the number of consumers firm i of age t has in the subset $\mathcal{A} \subset \mathcal{S}$. The function $f_{i,t}$ specifies both the number and the location of all the consumers of the firm. $f_{i,t}$ is not a probability distribution, as it sums up to $m_{i,t}$ and not 1.

The distribution of consumers $f_{i,t}$ evolves as follows.

First, firm i searches locally for consumers from where it is located (the location arbitrarily called the origin). Each period, it finds $\widetilde{\gamma\mu}$ new consumers where $\widetilde{\gamma\mu}$ is a positive integer-valued random variable of mean $\gamma\mu$. γ is the (constant) growth rate of the population of firms, and $\mu > 0$ is a parameter.[7] The location $x \in \mathcal{S}$ of each of these consumers is drawn randomly according to a function g, where $g(0, x)$ denotes the probability that a search originating from the origin (arbitrarily called 0) identifies a customer in location x. I expect that the function $g(0, x)$ depends on the distance between the origin (0) of the search and the destination (x), and the size of the destination x, but I will only impose such conditions later in Sections IIC and IID and when I bring the model to the data in Section III.

Second, given that firm i already has consumers in various locations, it searches for new consumers remotely from these locations. For each existing consumer in location $y \in \mathcal{S}$, the firm meets $\widetilde{\gamma\mu\pi}$ new consumers where $\widetilde{\gamma\mu\pi}$ is a positive integer

[7] Expressing the number of randomly drawn new consumers as a multiple of the population growth rate is merely a normalization, which will simplify the exposition of the main results.

valued random variable of mean $\gamma\mu\pi$. $\pi \geq 0$ is a parameter. The geographic location of these consumers is independently and randomly drawn according to the same function g, where $g(y, x)$ is the probability that a search originating from y identifies a customer in x. Remote search works exactly as local search, except that (i) it is shifted from the origin to location y, and (ii) the efficiency of this remote search is scaled by a constant factor π, which measures the relative importance of remote versus local search.

Without loss of generality, neither does a firm lose consumers, nor do firms die. Adding a random death process to either contacts or firms does not change any of the results below, beyond some simple rescaling of the parameters.[8]

Firm Level Dynamics.—The dynamic evolution of the network of consumers described above can be summarized in the following difference equation for $f_{i,t}$:

$$f_{i,t+1}(x) - f_{i,t}(x) = \sum_{k_0=1}^{\widetilde{\gamma\mu}_i} \mathbf{1}[\tilde{x}_{i,k_0} = x] + \sum_{y \in \mathcal{S}} f_{i,t}(y) \sum_{k_y=1}^{\widetilde{\gamma\mu\pi}_{i,y}} \mathbf{1}[\tilde{x}_{i,k_y} = x], \tag{3}$$

with the initial condition $f_{i,0}(x) = 0$, $\forall x \in \mathcal{S}$. $\mathbf{1}[\cdot]$ is the indicator function. $\widetilde{\gamma\mu}_i$ and the $\widetilde{\gamma\mu\pi}_{i,y}$ are independent draws from the random variables $\widetilde{\gamma\mu}$ and $\widetilde{\gamma\mu\pi}$, respectively. The $\tilde{x}$s are independent realizations from the probability distribution g, which determine the geographic location of each new contact. I give these draws some arbitrary index: for instance, $\Pr(\mathbf{1}[\tilde{x}_{i,k_y} = x]) = g(y, x)$ is the probability that a remote search from y identifies a consumer in x. The change in the number of consumers in location x from time t to time $t + 1$ can be decomposed in two terms. The first term corresponds to the local search for new contacts. Any k_0 of the $\widetilde{\gamma\mu}_i$ new contacts is located in x only if $\tilde{x}_{i,k_0} = x$. The second term corresponds to the remote search for new contacts. For each existing contact firm i has in location y (there are $f_{i,t}(y)$ of them), any k_y of the $\widetilde{\gamma\mu\pi}_{i,y}$ new contacts acquired from y is located in x only if $\tilde{x}_{i,k_y} = x$. Since the remote search can be intermediated via any location $y \in \mathcal{S}$, the new consumers found in x via y have to be summed over all possible remote location $y \in \mathcal{S}$.

The same parameters $(\gamma\mu, \gamma\mu\pi)$ in equation (3) govern the dynamic evolution of the network of contacts of any firm. This does not mean of course that any two firms will follow the same path ex post, as the luck of the draw will shape each individual firm's network differently. In particular, the second term in equation (3) implies a strong history dependence in firms export dynamics.

Aggregate Dynamics.—Averaging across a large number of firms within a cohort however, the randomness of each draw disappears, and I can derive a simple expression for the recursive evolution of population averages. Consider all the firms of age t located in the origin, and call N the number of such firms. I define $f_t^N(x)$ as the average number of contacts in location $x \in \mathcal{S}$ within this cohort, and $f_t(x)$ the limit of this population average when N gets large,

[8] See Atalay et al. (2011) for a related model, without geography, that features firm deaths.

$$f_t^N(x) \equiv \frac{\sum_{i=1}^N f_{i,t}(x)}{N} \in \mathbb{Q} \quad \text{and} \quad f_t(x) \equiv \lim_{N\to\infty} f_t^N(x) \in \mathbb{R}^+.$$

I show in Appendix B how to use the law of large numbers and the fact that all random shocks are i.i.d. in order to derive the following difference equation for the dynamics of f_t, the network of consumers of an entire cohort when the population is large:

$$(4) \qquad f_{t+1}(x) - f_t(x) = \gamma\mu g(0, x) + \gamma\mu\pi \sum_{y\in\mathcal{S}} f_t(y)\, g(y, x)\,,$$

with the initial condition $f_0(x) = 0$, $\forall x \in \mathcal{S}$. Note that as the population is large, all uncertainty has been removed in the aggregate. This does not mean that all firms within a cohort are identical, but that those differences are summarized by a stable function. This stable behavior for the population average obtains despite the inherent randomness of the small sample of consumers of any individual firm, but also despite the fact that as time goes on, the network of consumers of individual firms within the same cohort diverges.

The recursive definition of f_t in equation (4) is complex. In Section IIB, I present an analytical solution for the distribution of the total number of consumers (m_t) within the population. This solution holds for any set $\mathcal{S}$ and function g. Solving for other moments of f_t requires me to take a stand on $\mathcal{S}$ and on the function g. In Section IIC, I present an analytical solution for other moments of f_t in the special case where $\mathcal{S} = \mathbb{Z}$ and $g(y, x)$ only depends on the distance $|x - y|$. In Section IID, I conjecture using numerical simulations that as long as $g(y, x)$ only depends on the distance $||x - y||$ and on the size of the destination location x, the solution for the special case $\mathcal{S} = \mathbb{Z}$ is a good approximation of the general case where $\mathcal{S} \neq \mathbb{Z}$.

B. *The Number of Consumers*

Geography plays no role in the number of a firm's consumers. The geographically biased distribution g affects the location of consumers, but not the total number of them. Summing equation (4) over $\mathcal{S}$, I get a difference equation for the average number of consumers of firms within a cohort,

$$(5) \qquad m_{t+1} - m_t = \gamma\mu + \gamma\mu\pi m_t.$$

This process does not depend on any of the properties of the distribution g or of the set $\mathcal{S}$.

This recursive equation for the number of consumers resembles the model of acquisition of a network of "friends" in Jackson and Rogers (2007), which itself is an extension of the Steindl (1965) model of the firm size distribution.[9] In particular, the same mean-field approximation as in Jackson and Rogers can be used to solve

[9] The model is different from Jackson and Roger's in that firms do not explicitly learn about new contacts from the contacts of their existing contacts. For more elaborate models of the dynamic evolution of size, see for instance Gabaix (1999); Luttmer (2007); or Rossi-Hansberg and Wright (2007).

for the invariant distribution of the number of contacts among all cohorts for the cases where $\gamma\mu$ and $\gamma\mu\pi$ are nonintegers.[10] Numerical simulations in Section IID suggest that such a mean-field approximation is precise.

The following proposition characterizes the invariant distribution of the total number of consumers for any set $\mathcal{S}$ and function g.

PROPOSITION 1: *Under the mean-field approximation that the number of a firm's contacts evolves as the population average, the fraction of firms with fewer than m consumers is*

$$F(m) = 1 - \left(\frac{1}{1 + \pi m}\right)^{\frac{\ln(1+\gamma)}{\ln(1+\gamma\mu\pi)}}.$$

PROOF:

See Appendix B.

Let me briefly describe the properties of the cross-sectional distribution of the number of consumers, and provide some intuition. The upper tail of the distribution asymptotes to a scale-free Pareto distribution (with exponent $-\frac{\ln(1+\gamma)}{\ln(1+\gamma\mu\pi)} \approx -\frac{1}{\mu\pi}$ for γ small), whereas the lower tail is close to an exponential distribution (with rate parameter $\frac{\ln(1+\gamma)}{\gamma\mu} \approx \frac{1}{\mu}$ for γ small).

For m large (i.e., for firms that already have many consumers), the local search component $(\gamma\mu)$ becomes negligible, and only the remote search component remains. Each existing consumer allows the firm to gain a constant number $(\gamma\mu\pi)$ of new consumers. The growth rate of the number of consumers is approximately constant $(m_t \propto_{t\to\infty} e^{\gamma\mu\pi t})$. With the population growing at a rate γ, this leads to a Pareto distribution with an exponent $-1/(\mu\pi)$ in the upper tail.

For m small on the other hand (i.e., for firms with few consumers), the remote search component becomes negligible, and only the local search component remains. Each period, an approximately constant number $(\gamma\mu)$ of new consumers are added $(m_t \propto_{t\to 0} \gamma\mu t)$. With the population growing at a rate γ, this leads to an exponential distribution with parameter $1/\mu$ in the lower tail.

For intermediate values of m, the cross-sectional distribution of the number of consumers is a mixture of the above exponential and Pareto distributions. Plotting the counter-cumulative distribution $1 - F(m)$ in a log-log scale, the right end would asymptote a straight line (the Pareto upper tail), while the left end would exhibit some degree of concavity (the exponential lower tail). The slope of the upper tail, the range over which the distribution is concave, and how concave it is, all depend on the parameters μ and π.

[10] Under a mean-field approximation, the dynamics of the number of consumers of all individual firms within a cohort is assumed to be the same as the average within the cohort, where the number of a firm's consumers is understood as an expected number, generically not an integer. This assumption is perfectly suitable for the discrete time setup of my model, and does not require to take a continuous time limit.

C. *The Geography of Consumers when* $\mathcal{S} = \mathbb{Z}$

In the special case where $\mathcal{S} = \mathbb{Z}$, $g(y, x)$ only depends on the distance $|x - y|$ and the function $g(|\cdot|)$ has a finite second moment, the recursive characterization of the distribution of consumers f_t in equation (4) always admits an analytical solution.[11] In the interest of concision, I relegate this explicit solution to Lemma 1 in the Appendix. This solution allows me to derive closed-form solutions for any moment of the geographic distribution of consumers. I focus my attention on one particular moment, the average (squared) distance from a firm's consumers. I present the results in the special case $\mathcal{S} = \mathbb{Z}$ to sharpen the reader's intuition of the mechanics of the model. Those results do not necessarily hold exactly for other choices of $\mathcal{S}$ and g. I offer suggestive evidence in the next section that the solution derived in this special case is a good approximation of the general case where $\mathcal{S} \neq \mathbb{Z}$.

Define $\Delta(m)$ as the average (squared) distance from a firm's consumers among firms with m consumers, for a firm located in the origin $x = 0$. I call f_m the distribution of consumers among firms with m consumers, and $g_m = f_m/m$ the corresponding probability distribution. This average (squared) distance corresponds to the second moment of $g_m(|\cdot|)$,

$$\Delta(m) \equiv \sum_{x \in \mathbb{Z}} x^2 \, g_m(|x|) \, dx. \tag{6}$$

The next proposition characterizes the average (squared) distance from a firm's consumers among firms with m consumers, as a function of m.

PROPOSITION 2: *Under the mean-field approximation that the number of a firm's contacts evolves as the population average, the average (squared) distance from a firm's consumers,* $\Delta(m)$, *increases with the number of consumers* m,

$$\Delta(m) = \frac{\gamma\mu\pi}{(1 + \gamma\mu\pi) \ \ln(1 + \gamma\mu\pi)} \left(1 + \frac{1}{\pi m}\right) \ln(1 + \pi m) \, \Delta_g \,,$$

with $\Delta_g \equiv \sum_{x \in S} x^2 g(|x|)$ *the second moment of* $g(|\cdot|)$.

PROOF:

See Appendix B.

First note that the first term on the right-hand side becomes arbitrarily close to 1 for γ small,[12] which may help the reader get a more intuitive understanding of this proposition.

[11] f_t also admits an analytical solution when $\mathcal{S} = \mathbb{R}$, with the summation signs in equation (4) replaced by integral signs. It even admits a closed-form solution for the special cases where $\mathcal{S} = \mathbb{R}$ and g is either a Gaussian or a Cauchy distribution. See Lemma 1 in Appendix B for the presentation and derivation of those results.

[12] Using l'Hopital's rule, $\lim_{\gamma \to 0} \frac{\gamma\mu\pi}{(1 + \gamma\mu\pi) \ \ln(1 + \gamma\mu\pi)} = 1$.

Over time, not only does a firm acquire more consumers, but the geographic distance from these consumers increases. This result is entirely due to remote search, and can be understood as follows. Each time a firm gains one more consumer, it searches remotely from where this consumer is located. On average, existing consumers are some distance away from the firm, and remote searches bring new consumers who are themselves some distance away from those existing consumers. So each new wave of remote searches brings new consumers who tend to be further and further away.

Formally from the difference equation (4), the location of the new consumers acquired via remote search is the sum of the signed distance of the existing consumers and the signed distance of the remote search: for each consumer at a signed distance y from the origin, if the remote search delivers a new consumer who is herself at a signed distance $(x - y)$ from y, the new consumer will be at a signed distance $y + (x - y) = x$ from the origin. In other words, the remote search process is equivalent to taking the sum of two random variables, the first one being the variable that describes the location of existing consumers, and the second one being the remote search. The fact that the variance of the sum of two independent random variables is the sum of their variances explains why the average (squared) distance from a firm's consumers increases over time. This can be seen formally in equation (4). The term $\sum_{y \in \mathbb{Z}} f_t(y)\, g(|x - y|)$ is the convolution product of the functions $f_t(\cdot)$ and $g(|\cdot|)$. In probability theory, the convolution product is used to study the sum of random variables: the probability distribution of the sum of two random variables is the convolution product of their respective probability distributions. This is the essence of the proof of Proposition 2, where I show how to use Fourier transforms to manipulate convolution products.

From the reasoning above, and as can readily be seen in Proposition 2, the fact that the average (squared) distance from a firm's consumers increases with the number of consumers is only driven by the remote search process. Absent this remote search $(\pi \to 0)$ the average (squared) distance from consumers would be constant: $\Delta(m) = \Delta_g$, $\forall m$. Without remote search, a firm accumulates over time more and more consumers from a series of independent waves of local searches. All waves of new consumers brought by this local search have the exact same geographic distribution. Large firms sell to more consumers, but they have the same geographic distribution of consumers as small firms. If remote search is present $(\pi > 0)$, the average (squared) distance of sales increases with the number of consumers. Initially, for m small, the majority of new consumers come from local searches, and $\Delta(m)$ is relatively insensitive to $m : \partial\Delta(m)/\partial m|_{m=0} = 0$. As the number of consumers gets large (i.e., m large), the average (squared) distance of exports increases with the number of consumers in a log-linear way: $\Delta(m) \underset{m \to \infty}{\approx} \text{constant} + \Delta_g \ln(m)$.

Note that Proposition 2 holds for any arbitrary $g(|\cdot|)$ with a finite second moment. This is true despite the fact that the geographic distribution of new consumers depends in a complex nonlinear fashion on the entire distribution g, and hence on all the moments of this distribution. This result, while striking at first, can easily be understood as follows. The average (squared) distance from a firm's consumers after t periods corresponds to the second moment of $g_t = f_t/m_t$. The second moment of a distribution can be derived from the second derivative of the characteristic function of that distribution evaluated at zero. The proof of Proposition 2 in the Appendix shows how to transform the difference equation (4) so as to express the characteristic function of g_t in terms of the characteristic function of g. Then, for the same reason that the nth derivative of a composition of functions depends only on the first n derivatives of these functions, the second moment of g_t (the second derivative of its characteristic function) does not depend on any moment of g higher than 2 (any derivative of its characteristic function of order above 2).

Using the same analytical tools, potentially I can describe all the moments of the distribution of consumers f_t. While I will not describe all these moments, I make two observations which help understand the process of acquiring consumers.

First, the process of acquiring a network of consumers exhibits a strong history dependence. I have characterized above, for a large population, the distribution of the location of a firm's consumers, across cohorts with different numbers of consumers. This population average hides a lot of idiosyncracies even among firms with the same number of consumers.[13] If a firm initially happens to gain consumers in one particular location, it is subsequently more likely to keep gaining consumers in the vicinity of that location. So over time, the distribution of consumers of two initially identical firms will tend to diverge, each following its own history-dependent path. In equation (3), the location of a firm's new consumers at time $t + 1$ depends on the location of its existing consumers at time t. This property of the model conforms with the reduced-form empirical evidence presented in the previous section.

Second, over time, firms' consumers are not only further away, but also more geographically dispersed. This result is also due to the history dependence of the search process, but this time within and not between firms. Each existing consumer allows a firm to acquire new ones who will tend to be geographically concentrated around this existing consumer. Consumers tend to be clustered around each other. Time brings new clusters of consumers who tend to be increasingly far apart from each other. I formally state and prove this proposition in the online Appendix.

D. *The Number and Geography of Consumers when $\mathcal{S} \neq \mathbb{Z}$*

I have presented in the previous section a full characterization of the geographic distribution of any firm's consumers f_t for the special case where $\mathcal{S} = \mathbb{Z}$ and $g(y, x)$ only depends on the distance $|x - y|$. This exact solution can easily be extended to $\mathbb{R}$, $\mathbb{Z}^n$, or $\mathbb{R}^n$, but not to more complex sets. In particular, having in mind an empirical test of the theory using real world data, I have not been able to characterize f_t for the case where the set $\mathcal{S}$ is a discrete set of locations on a sphere.

To investigate the properties of the model when $\mathcal{S} \neq \mathbb{Z}$, I resort to numerical simulations.[14] Those simulations serve two purposes. First, the results presented in the previous section hold exactly for $\mathbb{Z}$, but not necessarily for other sets. Second, even in the special case where $\mathcal{S} = \mathbb{Z}$, I only characterize population averages for the limit when the population is large. The behavior of the model with a finite number of firms may differ from the large population limit.

[13] The distribution f_t is not a sufficient statistic to directly calculate some other population averages of interest, such as the average clustering of consumers of individual firms.

[14] Detailed instructions for replicating the simulations are given in the online Appendix.

I simulate the model numerically for three special cases for the set $\mathcal{S}$ and the function g:

- $\mathcal{S}_{circle}$ is a set of 8,766 equidistant locations along the circumference of a circle,

$$g(y, x) \;=\; \alpha_y e^{-||x-y||/\lambda} \;,$$

where $\lambda = 3.5$, $||x - y||$ is the length of the shorter arc between x and y, and α_y is a simple constant that ensures that probabilities sum to 1.

- $\mathcal{S}_{sphere}$ is a set of 8,766 approximately equidistant locations on the surface of a sphere,

$$g(y, x) \;=\; \alpha_y e^{-||x-y||/\lambda} \;,$$

where $\lambda = 3.5$, $||x - y||$ is the great circle distance between x and y, and α_y is a simple constant which ensures that probabilities sum to 1.

- $\mathcal{S}_{cities}$ is the set of the largest 8,766 actual cities in all countries in 2012 (with population above 50,000),

$$g(y, x) \;=\; \alpha_y \, Pop_x \, e^{-||x-y||/\lambda} \;,$$

where $\lambda = 3.5$, Pop_x is the population of city x, $||x - y||$ the great circle distance between x and y, and α_y is a simple constant that ensures that probabilities sum to 1. This means that after controlling for distance, it is twice as likely to find a contact in a city twice as large. The cardinality of the set $\mathcal{S}$ (8,766) is dictated by the number of cities in my dataset.[15]

Figure 1 compares the theoretical prediction from Proposition 1 (solid line) to a simulation of model with a finite number of firms (plus signs). The number of consumers does not depend on the properties of the set $\mathcal{S}$ or the distribution g, so that the results from the simulations are identical for all three choices for $\mathcal{S}$. I simulate 250 successive cohorts with a population of 20 firms in the first cohort. I chose parameters $(\gamma, \mu, \pi) = (0.02, 1{,}1)$. Both $\gamma\mu$, the number of directly searched contacts per period, and $\gamma\mu\pi$, the number of remotely searched contacts per period for each existing contact are nonintegers. Strictly speaking, a firm cannot receive a number $\gamma\mu$ or $\gamma\mu\pi$ of contacts in a period as the mean-field approximation in the proof of Proposition 1 assumes. They only do so *on average*. For instance, each period 1 in 50 firms receives a single direct contact, and the remaining 49 receive none. This integer constraint introduces some amount of residual noise. Figure 1 shows that this residual noise is relatively small.

[15] The dataset of city sizes and geographic coordinates is downloaded from http://download.geonames.org/. I am grateful to Thierry Mayer for providing me with references to this dataset.

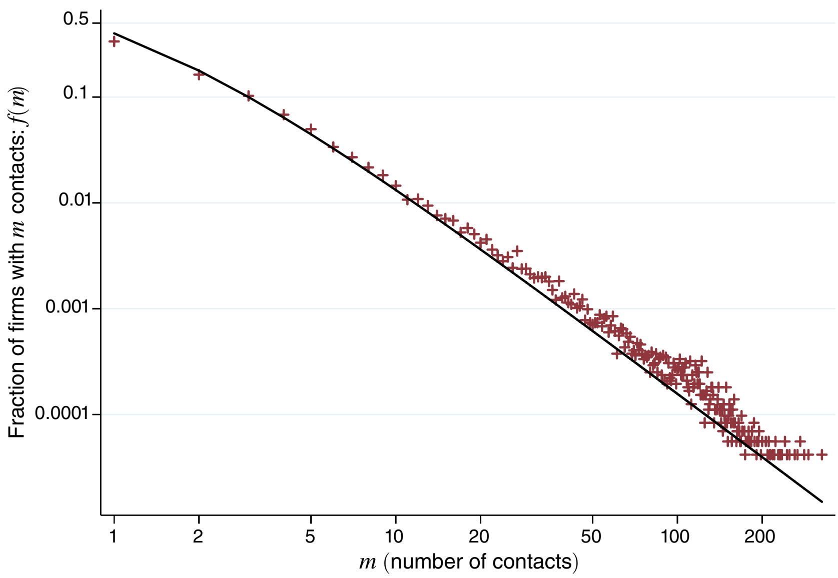

FIGURE 1. THE DISTRIBUTION OF THE NUMBER OF CONTACTS: THEORY VERSUS SIMULATION

Notes: Fraction of firms which export to m different contacts. Plus signs: 71,864 simulated firms in 250 cohorts. Solid line: theory under a mean-field approximation. Parameters: $\gamma = 0.02$, $\pi = 1$, and $\mu = 1$.

Figure 2 compares the theoretical prediction from Proposition 2 (solid lines), which holds in the special case $\mathcal{S} = \mathbb{Z}$, to numerical simulations of the model (plus signs) under three alternative choices for the set $\mathcal{S}$: 8,766 equidistant locations on a circle ($\mathcal{S}_{circle}$), 8,766 equidistant locations on a sphere ($\mathcal{S}_{sphere}$), and the set of the actual 8,766 largest cities in the world in 2012 ($\mathcal{S}_{cities}$). The numerical simulations suggest that despite the fact that the sets of $\mathcal{S}_{circle}$, $\mathcal{S}_{sphere}$, and $\mathcal{S}_{cities}$ differ from $\mathbb{Z}$, Proposition 2 provides a fair approximation of the geographic distribution of contacts even when $\mathcal{S} \neq \mathbb{Z}$. In particular, the average (squared) distance increases with m, the firm's number of contacts, as the theory predicts. How fast this average (squared) distance increases with m only depends on the parameters π, which governs the relative importance of remote versus direct search. The differences in units on the vertical axis across the three alternative choices for $\mathcal{S}$ are only due to differences in scale across the different sets.

Taken together, the results of the simulations for different sets $\mathcal{S}$ presented in Figures 1 and 2 suggest that Propositions 1 and 2 provide a fair approximation for alternative cases where $\mathcal{S} \neq \mathbb{Z}$. This is true in particular for the set $\mathcal{S}_{cities}$ that matches the actual geography of the world.

I will now formally bring the model predictions to the data.

III. Structural Estimation

In this section, I bring the key testable predictions from the theoretical model to the data on French exporters. The data are the same as those described in Section I. I structurally test the first two main aggregate predictions of the model regarding

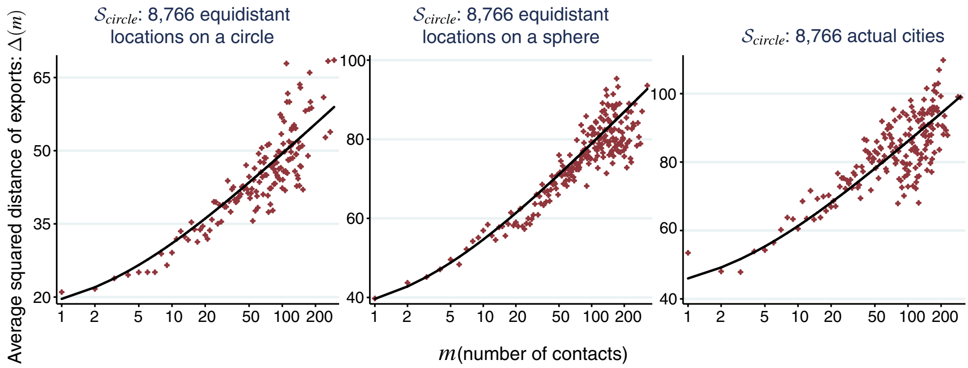

Figure 2. The Geography of Exports: Theory versus Three Simulations when $\mathcal{S} \neq \mathbb{Z}$.

Notes: Average squared distance from a firm's contacts, among firms with m contacts. Plus signs: 71,864 simulated firms in 250 cohorts for three different choices of the set $\mathcal{S}$. Left panel: $\mathcal{S}_{circle}$ is a set of 8,766 equidistant locations on a circle. Center panel: $\mathcal{S}_{sphere}$ is a set of 8,766 equidistant locations on a sphere. Right panel: $\mathcal{S}_{cities}$ is the set of the 8,766 largest cities in the world in 2012, with their actual sizes and coordinates. Solid line: theory, $\mathcal{S} = \mathbb{Z}$ and large population limit. Parameters: $\gamma = 0.02$, $\pi = 1$, $\mu = 1$, and $g(\| x - y \|) = \frac{1}{\lambda} e^{-\|x-y\|/\lambda}$ with $\lambda = 3.5$.

the fraction of firms exporting to various countries, as well as the distance of those exports, derived from Propositions 1 and 2. I use a simulated method of moments to bridge the gap between a micro-model (about firms selling to *contacts*) and macro-data (about firms exporting to *countries*).

A. *A Simulated Method of Moments Estimation*

The theory makes two separate yet equally important predictions. The first, Proposition 1, has to do with the fractions of firms that have many versus few consumers; the second, Proposition 2, has to do with the geographic location of those consumers. The theory predicts that the distribution of the number of consumers across firms is a mixture of an exponential and a Pareto distribution. The only two parameters which govern this distribution are μ, the number of new consumers acquired each period via local search (expressed in multiples of the firm population growth rate γ), and π, the efficiency of remote search relative to local search. The higher $\mu\pi$ is, the fatter the upper tail of the distribution of the number of contacts. In addition, the theory predicts that as a firm acquires more consumers, the average (squared) distance from those consumers increases. The higher is π, the efficiency of remote search relative to direct search, the faster the increase.

There is unfortunately one main complication which arises when bringing those predictions to the data: while the model makes predictions regarding the patterns of firm-level exports toward *consumers* (the number and location of those), the data only tells us about firm-level exports toward *countries*. Data on firm-level exports toward countries contain an informative but noisy signal about the underlying unobserved exports toward consumers. It is informative in the sense that if a firm exports to a given country, it has at least one consumer there. It is however noisy because the firm may have one or many consumers there. I have not been able to analytically characterize, directly from the model, the distribution of the number of countries a firm exports to in such a complex world. To circumvent this complication, I instead follow the guidance of the theory, in particular the simulations in Section IID, and estimate the model using the simulated method of moments (SMM). This SMM procedure closely follows that in Eaton, Kortum, and Kramarz (2011).

The model describes how firms acquire consumers gradually in various locations of the world. It is fully specified by equation (3) and by the actual geography of the world. From a simulated model where firms sell to notional consumers located in countries, I can observe whether or not a given firm exports to a given country, as I would do in the data.

Parametrization.—Equation (3) is fully parametrized by the probability distribution g, the geography of the world, and the parameters γ, μ, and π. I assume that g is an exponential function of distance, governed by a single parameter, $\lambda > 0$, scaled by the size (measured as GDP_x) of the destination country,

$$g(y, x) = \alpha_{\lambda, y} \, GDP_x e^{-||x-y||/\lambda} , \tag{7}$$

where $\alpha_{\lambda,y}$ is a simple scaling constant that ensures that the probabilities sum to 1.[16] The parameter λ dictates how geographically dispersed new contacts are, with the higher λ the higher g's variance.[17] Finally, I assume $\gamma = 0.02$ per period. As a small γ does not affect the theoretical moments, I do not include it in the set of parameters to be estimated. For the geography of the world, I simply read from the data the actual sizes and distances between countries. I am left with a vector of only three parameters to estimate,

$$\boldsymbol{\Theta} = (\mu, \pi, \lambda).$$

Simulation Algorithm.—For a given set of parameters $\boldsymbol{\Theta}$, I simulate successive artificial cohorts of French firms which sell to consumers in various countries. To ensure that as I search for the best $\boldsymbol{\Theta}$ the simulated data does not vary from purely idiosyncratic sources, I store all the realizations of the random elements of equation (3) for each contact of each firm at each time. This is not trivial. For instance, the higher μ or $\mu\pi$, the more consumers firms have, so that even the size of the simulated dataset varies with the choice of μ and π. Moreover, as the process is history dependent, where the consumers of a firm are located at a point in time affects where the future consumers of that firm will be located, so that changing the parameter λ that governs the function g alters entire branches of this dynamic tree. I solve these difficulties as follows.

Step 1 (**firms**)**:** The oldest generation has 20 firms. The number of firms of age t is $20(1 + 0.02)^t$, rounded up to the nearest integer. I simulate 360 successive cohorts,

[16] $\alpha_{\lambda,y} = 1/\sum_x GDP_x e^{-||x-y||/\lambda}$.

[17] This functional form is coincidentally the same as Comin, Dmitriev, and Rossi-Hansberg (2012) use for the geographical diffusion of technologies. Their median estimate for λ in the case of the diffusion of technology over space is 0.85 (mean = 0.4), compared to 3.5 for my estimate in the case of firm-level exports. This means that geographic barriers represent a hurdle for technology adoption about four times larger than for firms' exports.

for a total of 1,271,509 firms. Whether a firm is an exporter will of course depend on the parameters.

Step 2 (**potential consumers**)**:** I choose "large" initial values for μ and $\mu\pi$—i.e., much larger than the optimal values: $\mu = 1$ and $\pi = 1.35$. For each firm of age t, I keep all the consumers this firm had in period $t - 1$. I add to those a directly searched consumer with probability $\gamma\mu = 0.02$. For each existing consumer, I add an extra remotely searched consumer with probability $\gamma\mu\pi = 0.027$.

Step 3 (**stored information**)**:** For each link l between a French firm and a consumer I store four numbers: l', u_1, u_2, and u_3. l' is the name of the link that preceded l if l is the outcome of remote search, and 0 if l is the outcome of direct search. If l is the outcome of direct search, u_1 is a randomly generated uniform $[0, 1]$ number and $u_2 = 0$. If l is the outcome of remote search, u_1 is equal to the u_1' of the preceding link, and u_2 is the maximum of a randomly generated uniform $[0, 1]$ number and the u_2' of the preceding link l'. Storing the *maximum* of all the randomly generated random numbers is what will allow me appropriately to "chop off" entire chains of links when I search over alternative μs and πs corresponding to smaller datasets. Finally, u_3 is a randomly generated uniform $[0, 1]$ number that will determine the destination of the link.

Step 4 (**actual consumers**)**:** For given μ and π different from 1 and 1.35, a particular link between a firm and a customer exists if both $\gamma\mu > u_1$ and $\gamma\mu\pi > u_2$. Note that since all descendants of a random meeting share the same u_1, if this ancestor random meeting does not exist for a particular choice of μ, then none of the downstream links will exist either. By the same token, since u_2 is the maximum over an entire chain of links, if any upstream link does not exist for a particular choice of μ and π, then none of the downstream links that could have emanated from it will exist either.

Step 5 (**geographic location**)**:** For a given link's source country, which is the destination country? For each origin country c, I assume that the probability that a link originating from that country falls into country c' is given by $g(c, c')$. I then arbitrarily order all countries. The answer to the above question is that the link will fall in country c' if and only if $\sum_{n=1}^{c'} g(c, n) \leq u_3 < \sum_{n=1}^{c'+1} g(c, n)$.

I run steps 1, 2, and 3 once and for all. Steps 4 and 5 are run iteratively while searching for the parameters that best fit the data. For a given set of parameters $\boldsymbol{\Theta}$, at the end of step 5, I have generated an artificial set of French exporters, exporting to various sets of countries.

Moments.—I match 120 moments. The first 70 moments are the fraction of firms exporting to 1 country, 2 countries, …, 69 countries, and 70 or more countries.[18] They are the analog to the probability distribution $f(M) = F(M + 1) - F(M)$ in the theory, where M counts countries instead of consumers. The remaining

[18] The reason for stopping at 70 countries has to do with the construction of the weight matrix, described below.

50 moments are the average (squared) distance of exports, among firms that export to 1 country, 2 countries, …, and 50 countries.[19] They are the analog to the $\Delta(M)$s in the theory, where M counts countries instead of consumers. Those moments are constructed as follows.

To compute $\Delta(M)$, the average (squared) distance of exports among firms that export to exactly M countries, I need to measure both the geographic distance between France and other countries, as well as the distribution which governs the location of export destinations among firms that export to M countries. I use the empirical distribution of exports among firms that export to M countries and a simple linear correction for country size to define the moment $\Delta(M)$,

$$\Delta(M) = \frac{\sum_{i\in\mathcal{E}(M),c} (Dist_{France,c})^2 \left(\frac{1}{GDP_c}\right) \mathbf{1}[export_{i,c} > 0]}{\sum_{i\in\mathcal{E}(M),c} \left(\frac{1}{GDP_c}\right) \mathbf{1}[export_{i,c} > 0]}, \tag{8}$$

where $\mathcal{E}(M)$ is the set of firms that export to M countries.

I now have a vector of 120 moments, 1 for the actual data, $\mathbf{k}$, and 1 for each artificial dataset generated using a candidate set of parameters $\mathbf{\Theta}$, $\hat{\mathbf{k}}(\mathbf{\Theta})$. For each candidate $\mathbf{\Theta}$, I have a vector of deviations between the actual and simulated moments,

$$\mathbf{y}(\mathbf{\Theta}) = \mathbf{k} - \hat{\mathbf{k}}(\mathbf{\Theta}).$$

Estimation Procedure.—Under the moment condition that $E[\mathbf{y}(\mathbf{\Theta}_0)] = 0$ for the true value of the parameters $\mathbf{\Theta}_0$, I search the set of parameters $\hat{\mathbf{\Theta}}$ that minimizes the weighted deviations between the actual and simulated moments,

$$\hat{\mathbf{\Theta}} = \arg\min_{\mathbf{\Theta}} \{\mathbf{y}(\mathbf{\Theta})'\mathbf{W}\mathbf{y}(\mathbf{\Theta})\},$$

where $\mathbf{W}$ is a weight matrix[20] that accounts for the fact that some moments in the data are more precisely estimated than others. To search for $\hat{\mathbf{\Theta}}$, I first use a simulated

[19] The reason for stopping at 50 countries has to do with the construction of the weight matrix as well.

[20] The weight matrix $\mathbf{W}$ is constructed as follows. From the data, I draw 1,000 samples (Eaton, Kortum, and Kramarz 2011 take 2,000 draws). For each sample, I draw with replacement from the data as many firms as there are exporters *and* non-exporters in the actual data (about 600,000). Since I draw with replacement, the same firm may be sampled more than once. For each sample b, I calculate the moments $\mathbf{k}^b$. The weight matrix $\mathbf{W}$ is the inverse of the empirical variance-covariance matrix of my 120 moments, $\mathbf{\Omega}$,

$$\mathbf{W} = \mathbf{\Omega}^{-1}, \text{with } \mathbf{\Omega} = \frac{1}{1{,}000}\sum_{b=1}^{1{,}000} (\mathbf{k} - \mathbf{k}^b)(\mathbf{k} - \mathbf{k}^b)'.$$

In the data, there are no firms selling to exactly 70 countries (as well as some other numbers above 70). If I had chosen as one of the moments to match "the fraction of firms that export to exactly 70 countries," there would have been no difference between that moment in the data and in any of the 1,000 random samples. That would have meant taking $\frac{0}{0}$ in the above formula for W. This justifies my choice for the first 70 moments (fraction of firms exporting to 1 country, 2 countries, …, 69 countries, and 70 or more countries). Furthermore, when drawing samples from the data, it sometimes happens that no firm exports to exactly 51 countries (as well as other numbers above 51), so that would have meant taking $\frac{0}{0}$ in equation (8) and I cannot calculate $\Delta(51)$. This justifies my choice of the remaining 50 moments (average (squared) distance of exports, $\Delta(M)$, only for $M = 1, \dots, 50$).

Table 2—Direct Search, Remote Search, and Geography (*SMM estimates*)

	(1986)	(1987)	(1988)	(1989)	(1990)	(1991)	(1992)
π	2.420	2.495	2.479	2.499	2.574	2.633	2.401
	(0.187)	(0.114)	(0.150)	(0.066)	(0.114)	(0.130)	(0.200)
μ	0.371	0.368	0.384	0.362	0.357	0.338	0.384
	(0.022)	(0.013)	(0.021)	(0.010)	(0.013)	(0.014)	(0.027)
Parameter for $g(\|\|x-y\|\|) = \frac{1}{\lambda}e^{-\|\|x-y\|\|/\lambda}$:							
λ	3.419	3.398	3.448	2.906	3.515	3.418	3.513
	(0.131)	(0.145)	(0.130)	(0.403)	(0.177)	(0.132)	(0.135)

Notes: This table presents the SMM estimates of μ, π, and λ. The parameters μ and π govern the acquisition of the number of new consumers, while the parameter λ governs the geographic location of those consumers. Data: all French exporters, 1986–1992. Bootstrapped standard errors are in parentheses. All coefficients are statistically different from zero at the 1 percent level of significance.

annealing algorithm[21] to find $\hat{\Theta}_{(1)}$. Then, starting from $\hat{\Theta}_{(1)}$, I run a simplex maximization algorithm to get $\hat{\Theta}$. This two-step approach is required because of the presence of many local minima, with a series of small "lakes" (local minima) separated by a series of "ridges." Standard errors are calculated by bootstrapping and account for both sampling and simulating errors.[22] I run the above estimation separately for each year from 1986 to 1992.

Results.—The estimated parameters are presented in Table 2. For the year 1992, the data suggest that $\mu = 0.38$ and $\pi = 2.4$. In other words, remote search is more than twice as important as direct search for a firm with a single existing contact. Of course, as firms acquire more contacts, remote search accounts for an increasing share of the firm's new contacts. For a firm with the sample mean number of 3.5 foreign contacts, 90 percent of new contacts come from remote search. For a firm with 20 foreign contacts (ninetieth percentile), remote search dominates, and accounts for 98 percent of new contacts. In the aggregate, remote search accounts for about 90 percent of all new contacts.[23]

Figure 3 plots together the actual data (dots) and the simulated data (plus signs) for the year 1992. It shows the fraction of firms that export to different number of

[21] For this algorithm, I use the MATLAB code from Joachim Vandekerckhove, available at http://www.mathworks.com/matlabcentral/fileexchange/10548/.

[22] For bootstrap b, I take a sample with replacement from the data as well as from the steps 1, 2, and 3 of the simulation. As for the construction of the weight matrix, the number of draws is the same as either in the actual data or in the simulated data. With these two samples, I can calculate $\mathbf{y}^b(\Theta) = \mathbf{k}^b - \hat{\mathbf{k}}^b(\Theta)$ for any Θ. I use the same maximization to find $\hat{\Theta}^b = \arg\min_{\Theta}\{\mathbf{y}^b(\Theta)'\mathbf{W}\mathbf{y}^b(\Theta)\}$. Note that as Eaton, Kortum, and Kramarz (2011), I do not recalculate a new weight matrix $\mathbf{W}^b$ for each bootstrap. I perform 20 such bootstraps (Eaton et al. 2011 use 25 bootstraps), and calculate the empirical variance-covariance matrix of the estimated parameters,

$$\mathbf{V}(\Theta) = \frac{1}{20}\sum_{b=1}^{20}(\hat{\Theta}^b - \hat{\Theta})(\hat{\Theta}^b - \hat{\Theta})'.$$

The bootstrapped standard errors are the square roots of the diagonal elements of $\mathbf{V}(\Theta)$.

[23] For this calculation, I use a continuous approximation to circumvent integer constraints. In any period, direct search delivers $\overline{m}_{direct} = \gamma\mu$ new contacts for all firms, irrespective of their number of existing contacts. Remote search brings $\gamma\mu\pi m$ new contacts for a firm with m existing contacts. Given that the contacts are distributed within the population according to the c.d.f. $F(m)$, the average number of new contacts brought by remote searches is $\overline{m}_{remote} = \int_0^{+\infty}\gamma\mu\pi m dF(m) = \gamma\mu(\mu\pi)/(1-\mu\pi)$. The fraction of new contacts delivered by remote search is therefore $\overline{m}_{remote}/(\overline{m}_{direct} + \overline{m}_{remote}) = \mu\pi \approx 0.93$.

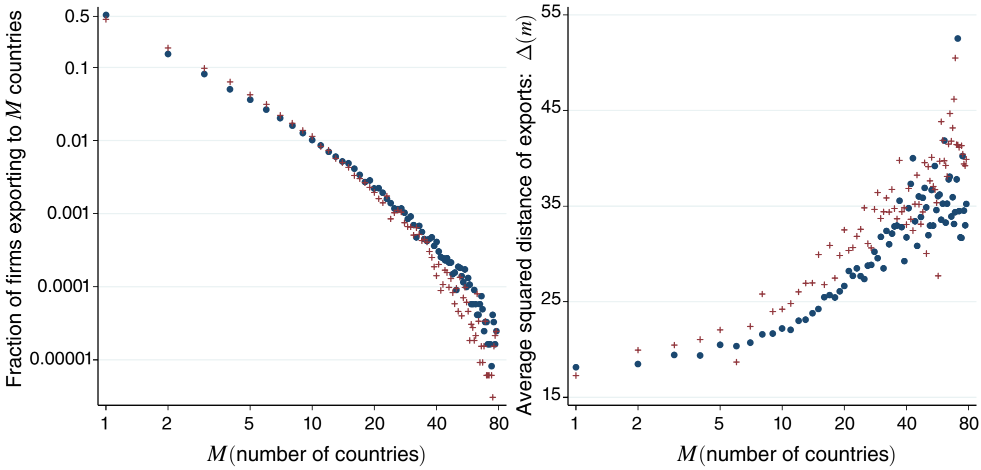

Figure 3. The Number and Geography of Exports (*SMM estimates*)

Notes: Left panel: fraction of firms that export to M different countries. Right panel: average squared distance to a firm's export destinations, among firms exporting to M destinations, as defined in equation (8); distances are calculated in thousands of kilometers. Dots: data, all French exporters in 1992. Plus signs: simulated data; $\pi = 2.401$ (0.200), $\mu = 0.384$ (0.027) and $\lambda = 3.513$ (0.135) are estimated by simulated method of moments.

countries (left panel), and the average (squared) distance of exports among firms that export to different number of countries (right panel). For both, the fit between the simulation and the data is good. The simulated data underestimates the fraction of firms that export to many countries, and overestimates the average (squared) distance of exports among firms that export to an intermediate number of countries.

The next section discusses the relation between my theory and existing trade models.

B. *Discussion*

Existing international trade models with heterogeneous firms, such as Bernard et al. (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 model I develop offers a parsimonious theory for the extensive margin of international trade.

For this discussion, I assume that in my model, the number of contacts of a firm, m, is proportional to the number of countries it exports to, M: $m \propto M$.

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 simplifying assumption that all trade barriers and country sizes 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 $M' \leq M$. The fraction of firms that export to exactly M markets is then the fraction of firms that have a productivity between the productivity thresholds for exporting to M and $M + 1$. 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 counterfactual 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.[24]

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 (2011) can a priori replicate any pattern of entry in the data. Calibrating their model to the data, they need to assume a large amount of idiosyncratic noise, with a ratio of the relevant combination of fixed entry costs and local demand shocks of 1 to 13 between the twenty-fifth and the seventy-fifth percentiles. So 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 assumed Pareto distribution of productivities across firms, which matches the data well. 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.[25] My model can be thought of as a micro-foundation for the distribution of the entry shocks that Eaton, Kortum, and Kramarz (2011) agnostically take as a purely random process.

In the stochastic model of Bernard et al. (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 that they analyze, there is no uncertainty either in the fraction of firms entering any given market, or in the distribution of the number of markets entered. This distribution depends on the specific trade barriers and country characteristics. Even under the convenient Frechet assumption for the 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 foreign markets ($f(M) = 0$ if $0 < M \leq M_{\max}$). In the other extreme of perfectly free trade, all firms which sell domestically also export to all countries in the world ($f(M) = 0$ if $M < M_{\max}$ and $f(M_{\max}) = 1$). So whereas the fraction

[24] I develop these arguments formally and provide a calibration of the Melitz/Chaney model in the online Appendix.

[25] Similarly, Armenter and Koren (2014) 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 (empty "bins"). By contrast, instead of assuming this distribution to match the data, my model offers a theory that generates such a distribution endogenously.

of firms that export to all foreign countries in the world $(M_{\max})$ is monotonically decreasing from 1 to 0 as trade barriers rise, the fraction of firms exporting to any other number of foreign countries $(M < M_{\max})$ 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 tend to increase with distance, and if as in the data there is no systematic correlation between country size and distance from France, both the Melitz/Chaney model and Bernard et al. (2003) would correctly predict that the distance of exports 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 (2011), 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, 17 in my simulated model, but only 2 in the calibrated Eaton, Kortum, and Kramarz (2011) model.[26]

To summarize, while existing firm-level trade models cannot match several facts about the extensive margin of trade, I develop a parsimonious model that matches those facts. On the other hand, my model is silent about the determinants of the intensive margin of trade, or about the relationship between a firm's exposure to international trade and its size in different markets, while those models make precise and factually correct predictions about those patterns. In that sense, the network model I develop is complementary to the existing firm-level trade models. The proposed theory can be thought of as a micro-foundation for the assumptions on export costs needed in those existing models to match the data on the extensive margin of trade.

IV. Conclusion

Motivated by reduced-form evidence on the dynamics of firms' exports I uncover, I propose a new model of trade frictions, which generates a dynamic entry of firms into geographically dispersed markets. Firms can sell only in locations where they have a contact. Firms search for trading partners directly, but they also use their existing network of contacts to search remotely for new partners. This dynamic model generates a stable spatial distribution of sales across firms. Bringing those predictions to data on French firm-level exports between 1986 and 1992 suggests that remote search is about twice as important as direct search for a firm with a single foreign contact, but it quickly dominates as a firm acquires more contacts. This explains both the fat upper tail and the thinner lower tail of the distribution of the number of foreign countries accessed by French firms. It jointly explains the fact that the average distance of exports increases at an accelerating pace with the number of foreign countries accessed.

This model and the empirical findings that support it suggest several extensions and generalizations. First, the emergence of a stable distribution of entrants into

[26] For a more intuitive interpretation of these numbers, the average distance is 3,500km in the actual and simulated data versus 900km in the calibrated Eaton, Kortum, and Kramarz (2011) model.

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 equation. This may provide an explanation for the stable role geographic distance plays in explaining aggregate bilateral trade flows. I propose such an explanation in Chaney (2013). Second, 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 welfare analysis. In particular, the welfare gains from trade in my model are unevenly distributed, because information about profitable foreign trade is unevenly distributed. Moreover, the measure itself of the welfare gains from trade in my model would be very different from conventional models. For instance, information on price dispersion would be of little help, as they would reflect to a large extent the uneven distribution of information. This point is further developed and tested by Allen (forthcoming), who uses a similar theoretical framework. Third, 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 (e.g., wars or economic crises such as the 2008 global Great Recession), the rapid growth of a large country (e.g., China), or the sudden decline of a set of firms (e.g., silicon wafer and semiconductor producers after the Fukushima disaster) may have a long-lasting impact on the world geography of trade, since (re)building contacts is a lengthy and history-dependent process. I leave these questions for future research.

Appendix

A. *Data Sources*

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 (2011). I only keep a 0/1 indicator for positive exports. I use data on all French exporters.[27] In addition, the customs data are matched with balance sheet information collected by the French fiscal authorities for all firms with a turnover of 1,000,000 French francs in services, or 3,000,000 French francs in manufacturing (US$1 $\approx$ 5 French francs in 1992). Virtually all exporters are in this dataset. I use this dataset to assign each firm to its primary two-digit industrial sector. Table A1 reports the list of two-digit sectors, as well as the distribution of all exporters in those sectors.

Distance data: I use data on bilateral distances between countries collected and constructed by the CEPII. The distance between two countries, measured in thousands of kilometers, is calculated as a population-weighted arithmetic average of the geodesic distances between the main cities in these countries. See Mayer and Zignago (2006).

Country size data: I use as a measure of a country's size its nominal GDP (in millions US$) in the current year. See http://pwt.econ.upenn.edu/.

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

TABLE A1—INDUSTRIAL SECTORS

Sector	N100 Industries	Firm-Year-Destination
Agriculture	0–3	387,589
Mining	4–14	217,433
Construction	15–15	335,059
Manufacturers	16–56	11,589,869
Transportation	68–75	179,941
Wholesale	57–59	5,960,095
Retail	60–64	1,883,046
F.I.R.E.	76, 78–81, 88–89	269,345
Services	65–67, 77, 82–87	1,041,124
Public Administration + Other	90–99	21,115
Total		21,884,616

Note: F.I.R.E. refers to Finance, Insurance, and Real Estate.

Bilateral trade flows: To proxy for the intensity of firm-level contacts between countries, I use data on the nominal value (in US$) of aggregate bilateral exports. See Feenstra et al. (2004).

B. *Mathematical Proofs*

PROPOSITION 3: *When the population N of a cohort of firms grows large, the distribution of consumers for the entire cohort evolves recursively according to*

$$f_{t+1}(x) - f_t(x) = \gamma\mu g(0,x) + \gamma\mu\pi \sum_{y\in\mathcal{S}} f_t(y)\, g(y,x).$$

PROOF:

The law of motion for the distribution of consumers for a single firm i is

$$f_{i,t+1}(x) - f_{i,t}(x) = \sum_{k=1}^{\widetilde{\gamma\mu}_i} \mathbf{1}[\tilde{x}_{i,k} = x] + \sum_{y\in\mathcal{S}} f_{i,t}(y) \sum_{k_y=1}^{\widetilde{\gamma\mu\pi}_{i,y}} \mathbf{1}[\tilde{x}_{i,k_y} = x]\,.$$

Averaging across N firms of age t, one gets the laws of motion for the population average,

$$\begin{aligned} f_{t+1}^N(x) - f_t^N(x) &= \frac{\sum_{i=1}^N (f_{i,t+1}(x) - f_{i,t}(x))}{N} \\ &= \frac{\sum_{i=1}^N \Big(\sum_{k=1}^{\widetilde{\gamma\mu}_i} \mathbf{1}[\tilde{x}_{i,k} = x] + \sum_{y\in\mathcal{S}} \frac{m_t f_{i,t}(y)}{m_t} \sum_{k_y=1}^{\widetilde{\gamma\mu\pi}_{i,y}} \mathbf{1}[\tilde{x}_{i,k_y} = x] \Big)}{N} \\ &= \frac{\sum_{i=1}^N \sum_{k=1}^{\widetilde{\gamma\mu}_i} \mathbf{1}[\tilde{x}_{i,k} = x]}{N} \\ &\quad + m_t \sum_{y\in\mathcal{S}} \frac{\sum_{i=1}^N \sum_{k_y=1}^{\widetilde{\gamma\mu\pi}_{i,y}} g_{i,t}(y)\, \mathbf{1}[\tilde{x}_{i,k_y} = x]}{N}, \end{aligned}$$

with $g_{i,t}(x) = f_{i,t}(x)/m_t$ the function describing not the number of contacts of firm i at time t in location x, but the fraction of i's contacts in x. The term $g_{i,t}(y)\,1[\tilde{x}_{i,k_y} = x]$ is the analog for a particular firm i of the joint probability distribution of the events "a random draw from i's direct search at t is in y *AND* a random new search originating from y is in x." For a large population, define $h_t(y, y, x)$ as the corresponding joint probability distribution of the events "a random draw from all firms' existing contacts at t is in y *AND* a random new search originating from y is in x." From the law of large numbers, the empirical frequencies converge to the true probabilities as the population grows large. Moreover, from the law of larger numbers, the number of terms in the sum $\sum_{i=1}^{N}\sum_{k=1}^{\widetilde{\gamma\mu_i}}$ converges to $N\gamma\mu$ and the number of terms in the sum $\sum_{i=1}^{N}\sum_{k_y=1}^{\widetilde{\gamma\mu\pi_{i,y}}}$ converges to $N\gamma\mu\pi$. For any (x, y),

$$\text{when } N \to \infty: \quad \begin{cases} \dfrac{\sum_{i=1}^{N}\sum_{k=1}^{\widetilde{\gamma\mu_i}} \mathbf{1}[\tilde{x}_{i,k} = x]}{N} & \overset{a.s.}{\to} \gamma\mu g(0,x) \\[2ex] \dfrac{\sum_{i=1}^{N}\sum_{k_y=1}^{\widetilde{\gamma\mu\pi_{i,y}}} g_{i,t}(y)\,\mathbf{1}[\tilde{x}_{i,k_y} = x]}{N} & \overset{a.s.}{\to} \gamma\mu\pi h_t(y,y,x) \end{cases}.$$

As the new draws from the g distribution are independent from the existing distribution of contacts $(g_{i,t})$, the joint distribution $h_t(\cdot,\cdot,\cdot)$ is simply the product of the distributions of each variable,

$$h_t(y, y, x) = g_t(y)\,g(y, x) \ \text{ with } g_t(y) = \lim_{N\to\infty}\frac{\sum_{i=1}^{N}\frac{f_{i,t}(y)}{m_t}}{N} = \frac{f_t(y)}{m_t}.$$

Plugging those limits into the law of motion for the population average, I get the proposed expression when the population of the cohort grows large,

$$\lim_{N\to\infty} f_{t+1}^{N}(x) - \lim_{N\to\infty} f_t^{N}(x) = \quad = \gamma\mu g(0, x) + \gamma\mu\pi\sum_{y\in\mathcal{S}} f_t(y)\,g(y, x).$$

PROPOSITION 1 (Reminded): *Under the mean-field approximation that the number of a firm's contacts evolves as the population average, the fraction of firms with fewer than m consumers is*[28]

$$F(m) = 1 - \left(\frac{1}{1+\pi m}\right)^{\frac{\ln(1+\gamma)}{\ln(1+\gamma\mu\pi)}}.$$

PROOF:

At any time, in any location, and therefore in the union of any set of locations (any country), the fraction of firms with more than m contacts is the same. Summing the number of contacts of a firm of age t over the entire set $\mathcal{S}$, I get the following simple

[28] Note that the ms only take a discrete set of values (corresponding to $m_1, m_2, \ldots$, etc.). The formula for $F(\cdot)$ is neither an approximation nor a limit. It holds exactly at any time t for those values $(m_1, m_2, \ldots, m_t)$.

difference equation for the evolution of the average number of contacts of firms within a cohort,

$$m_{t+1} - m_t = \gamma\mu + \gamma\mu\pi m_t,$$

with initial condition $m_0 = 0$. Note that when $\gamma\mu$ (respectively $\gamma\mu\pi$) is not an integer, the number of directly (respectively, remotely) searched contacts from one period to the next is equal to $\gamma\mu$ (respectively, $\gamma\mu\pi$) only *in expectation*. As in Jackson and Rogers (2007), when $\gamma\mu$ or $\gamma\mu\pi$ are nonintegers, I use a mean-field approximation and assume that each firm receives exactly the population average number of directly (resp., remotely) searched contacts. The numerical simulations in Section IID suggest that this approximation is precise. See Atalay (2013) for a formal derivation without a mean-field approximation. The above difference equation admits the solution,

$$m_t = \frac{1}{\pi}((1 + \gamma\mu\pi)^t - 1).$$

I invert this equation to get the age of a firm as a function of its number of contacts,

$$t(m) = \frac{\ln(1 + \pi m)}{\ln(1 + \gamma\mu\pi)}.$$

where t only takes integer values.[29] The fraction of firms with more than m contacts, $1 - F(m)$, is the fraction of firms older than $t(m)$. Given the exponential growth rate of the population, this fraction is $(1 + \gamma)^{-t(m)}$. Using the above expression for $t(m)$, I get a general expression for any γ,

$$\begin{aligned} 1 - F(m) &= (1 + \gamma)^{-t(m)} \\ &= (1 + \pi m)^{-\ln(1+\gamma)/\ln(1+\gamma\mu\pi)}, \end{aligned}$$

from which I derive the proposed expression,

$$F(m) = 1 - \left(\frac{1}{1 + \pi m}\right)^{\frac{\ln(1+\gamma)}{\ln(1+\gamma\mu\pi)}}.$$

LEMMA 1: *The distribution of consumers f_t is given by*

$$f_t = \frac{1}{\pi}((\delta + \gamma\mu\pi g)^{*t} - \delta),$$

[29] Note that I am not making any continuous approximation of the discrete model. The proposed formulas are exactly correct when t is an integer.

where δ is the Dirac delta function and the exponent $*t$ stands for a function convoluted t times with itself. f_t admits a closed-form solution in the special cases where g is a centered Gaussian or Cauchy distribution.

PROOF:

First note that the sum on the right-hand side of equation (4) is a convolution product, so that the difference equation can be written in a compact form as

$$f_{t+1} = \gamma\mu g + f_t + \gamma\mu\pi g * f_t,$$

where $*$ stands for the convolution product of two functions. I take the Fourier transform of this equation, where I denote $\hat{f}(\omega) \equiv \sum_{x\in\mathbb{Z}} f(x)\, e^{-i\omega x}$ the Fourier transform of f. Using the convolution theorem which states that the Fourier transform of the convolution of distributions is the product of their Fourier transforms, I get

$$\hat{f}_{t+1} = \gamma\mu\hat{g} + \hat{f}_t + \gamma\mu\pi\hat{g}\hat{f}_t,$$

with initial condition $\hat{f}_0 = 0$, where I denote by $\hat{f}$ the Fourier transform of the function f. This first-order linear recursive equation admits the following solution:

$$\hat{f}_t = \frac{1}{\pi}\big((1 + \gamma\mu\pi\hat{g})^t - 1\big).$$

Taking the inverse Fourier transform of this equation, I get the proposed expression for f_t.

To derive a closed-form solution for the special cases where $g(|\cdot|)$ is a Gaussian or a Cauchy distribution,[30] I manipulate this expression and get

$$\begin{aligned} f_t &= \frac{1}{\pi}\big((\delta + \gamma\mu\pi g)^{*t} - \delta\big) \\ &= \frac{1}{\pi}\sum_{s=1}^{t} (\gamma\mu\pi)^s \binom{t}{s} g^{*s}. \end{aligned}$$

Note that the convolution of the p.d.f.s of t random variables is the p.d.f of their sum. As the sum of t Gaussian (respectively, Cauchy) distributed random variables is also a Gaussian (Cauchy), I derive closed form solutions. If $g(|\cdot|) = \phi_{\sigma^2}(\cdot)$ where ϕ_{σ^2} is the p.d.f. of a Gaussian distribution with mean zero and variance σ^2, then $g(|\cdot|)^{*s} = \phi_{s\sigma^2}(\cdot)$, and I get

$$f_t = \frac{1}{\pi}\sum_{s=1}^{t} (\gamma\mu\pi)^s \binom{t}{s} \phi_{s\sigma^2}.$$

[30] Note that there is a slight abuse of language, as I have so far worked with a discrete set of locations, and the corresponding discrete random variables, while the Normal and the Cauchy are continuous random variables. All the analysis presented in the paper is identical with a continuum of locations ($x \in \mathbb{R}$) instead of a discrete set ($x \in \mathbb{Z}$), with summation signs replaced by integral signs, but the discrete case is more natural given the empirical application to international trade.

If $g(|\cdot|) = \psi_\gamma(\cdot)$ where ψ_γ is the p.d.f. of a Cauchy distribution centered around zero and with scale parameter γ, then $g(|\cdot|)^{*s} = \psi_{s\gamma}(\cdot)$, and I get

$$f_t = \frac{1}{\pi} \sum_{s=1}^{t} (\gamma\mu\pi)^s \binom{t}{s} \psi_{s\gamma}.$$

PROPOSITION 2 (*Reminded*): *Under the mean-field approximation that the number of a firm's contacts evolves as the population average, the average (squared) distance from a firm's consumers,* $\Delta(m)$, *increases with the number of consumers* m,[31]

$$\Delta(m) = \frac{\gamma\mu\pi}{(1+\gamma\mu\pi)\ \ln(1+\gamma\mu\pi)} \left(1 + \frac{1}{\pi m}\right) \ln(1+\pi m)\, \Delta_g$$

with $\Delta_g \equiv \sum_{x\in\mathbb{Z}} x^2 g(|x|)\, dx$ *the second moment of* $g(|\cdot|)$.

PROOF:

From Lemma 1, I get an expression not only for the the distribution of contacts, $g_t = f_t/m_t$, but more interestingly for its Fourier transform, $\hat{g}_t = \hat{f}_t/m_t$,

$$\hat{g}_t = \frac{(1+\gamma\mu\pi\hat{g})^t - 1}{(1+\gamma\mu\pi)^t - 1}.$$

Note that if $g_t(|\cdot|)$ is the probability distribution of a random variable X_t, then its Fourier transform $\hat{g}_t$ is closely related to $\varphi_{g_t(|\cdot|)}(\omega) = E[e^{i\omega X_t}]$, the characteristic function of X_t,

$$\hat{g}_t(\omega) = \sum_{x\in\mathbb{Z}} g_t(|x|)\, e^{-i\omega x} = E[e^{-i\omega X_t}] = \varphi_{g_t(|\cdot|)}(-\omega)\ .$$

The various moments of $g_t(|\cdot|)$ are then simply given by the various derivatives of $\hat{g}_t$ evaluated at zero. The first two derivatives are

$$\hat{g}_t' = \frac{\gamma\mu\pi t \hat{g}'(1+\gamma\mu\pi\hat{g})^{t-1}}{(1+\alpha)^t - 1}$$

$$\hat{g}_t'' = \gamma\mu\pi t\ \frac{\hat{g}''(1+\gamma\mu\pi)^{t-1} + \gamma\mu\pi(t-1)\,\hat{g}'^2(1+\gamma\mu\pi\hat{g})^{t-2}}{(1+\gamma\mu\pi)^t - 1}.$$

[31] As for Proposition 1, note that the ms only take a discrete set of values (corresponding to $m_1, m_2, \ldots$, etc). The formula for $\Delta(\cdot)$ for those values is exact and neither an approximation nor a limit.

Note that since the distribution $g(|x|)$ is symmetric about zero, its first moment is zero, $\hat{g}'(0) = 0$. The average (squared) distance of exports for a firm of age t, Δ_t, is simply the second moment of g_t, given by the second derivative of $\hat{g}_t$ evaluated at zero.

$$\Delta_t \equiv \sum_{x \in \mathbb{Z}} x^2 g_t(|x|) = E[X_t^2] = \hat{g}_t''(0) = \frac{\gamma\mu\pi t(1 + \gamma\mu\pi)^{t-1}}{(1 + \gamma\mu\pi)^t - 1} \Delta_g ,$$

with $\Delta_g \equiv \sum_{x \in \mathbb{Z}} x^2 g(|x|)$ the second moment of $g(|\cdot|)$. Using the expression that relates a firm's age to the number of its contacts from the proof of Proposition 1, I get[32]

$$t(m) = \frac{\ln(1 + \pi m)}{\ln(1 + \gamma\mu\pi)}$$

$$(1 + \gamma\mu\pi)^{t(m)} - 1 = \pi m$$

$$(1 + \gamma\mu\pi)^{t(m)-1} = \frac{1 + \pi m}{1 + \gamma\mu\pi} .$$

Plugging those expressions into the expression for $\Delta_{t(m)} = \Delta(m)$, I derive the proposed expression

$$\Delta(m) = \frac{\gamma\mu\pi}{(1 + \gamma\mu\pi) \ln(1 + \gamma\mu\pi)} \left(1 + \frac{1}{\pi m}\right) \ln(1 + \pi m) \Delta_g .$$

REFERENCES

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica* 80 (5): 1977–2016.

Ahn, JaeBin, Amit K. Khandelwal, and Shang-Jin Wei. 2011. "The Role of Intermediaries in Facilitating Trade." *Journal of International Economics* 84 (1): 73–85.

Albornoz, Facundo, Hector F. Calvo Pardo, Gregory Corcos, and Emanuel Ornelas. 2012. "Sequential Exporting." *Journal of International Economics* 88 (1): 17–31.

Allen, Treb. Forthcoming. "Informational Frictions in Trade." *Econometrica*.

Allen, Treb, and Costas Arkolakis. Forthcoming. "Trade and the Topography of the Spatial Economy." *Quarterly Journal of Economics*.

Antràs, Pol, and Arnaud Costinot. 2011. "Intermediated Trade." *Quarterly Journal of Economics* 126 (3): 1319–74.

Arkolakis, Costas. 2010. "Market Penetration Costs and the New Consumers Margin in International Trade." *Journal of Political Economy* 118 (6): 1151–99.

[32] To invert m_t into $t(m)$ when $\gamma\mu$ and $\gamma\mu\pi$ are nonintegers, I rely on the same mean-field approximation as in the proof of Proposition 1. In the full model, not all firms of age t have exactly m_t consumers, so that $t(m)$ is not the age of any firm with m consumers. Under the mean-field approximation that all firms follow the same process as the average within their cohort, the equivalence between a firm's age and its number of contacts is restored.

Armenter, Roc, and Miklós Koren. Forthcoming. "Economies of Scale and the Size of Exporters." *Journal of the European Economic Association*.
Armenter, Roc, and Miklós Koren. 2014. "A Balls-and-Bins Model of Trade." *American Economic Review* 104 (7): 2127–51.
Atalay, Enghin. 2013. "Sources of Variation in Social Networks." *Games and Economic Behavior* 79: 106–31.
Atalay, Enghin, Ali Hortacsu, James W. Roberts, and Chad Syverson. 2011. "Network Structure of Production." *Proceedings of the National Academy of Sciences* 108 (13): 5199–5202.
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.
Bramoulle, Yann, Sergio Currarini, Matthew O. Jackson, Paolo Pin, and Brian W. Rogers. 2012. "Homophily and Long-Run Integration in Social Networks." *Journal of Economic Theory* 147 (5): 1754–86.
Burchardi, Konrad B., and Tarek A. Hassan. 2013. "The Economic Impact of Social Ties: Evidence from German Reunification." *Quarterly Journal of Economics* 128 (3): 1219–71.
Casella, Alessandra, and James E. Rauch. 2002. "Anonymous Market and Group Ties in International Trade." *Journal of International Economics* 58 (1): 19–47.
Chaney, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review* 98 (4): 1707–21.
Chaney, Thomas. 2013. "The Gravity Equation in International Trade: An Explanation." National Bureau of Economic Research Working Paper 19285.
Chaney, Thomas. 2014. "The Network Structure of International Trade: Dataset." *American Economic Review.* http://dx.doi.org/10.1257/aer.104.11.3600.
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.
Comin, Diego, Mikhail Dmitriev, and Esteban Rossi-Hansberg. 2012. "The Spatial Diffusion of Technology." National Bureau of Economic Research Working Paper 18534.
Defever, Fabrice, Benedikt Heid, and Mario Larch. 2010. "Spatial Exporters." Unpublished.
Desmet, Klaus, and Esteban Rossi-Hansberg. 2010. "On Spatial Dynamics." *Journal of Regional Science* 50 (1): 43–63.
Dixit, Avinash K. 1989. "Hysteresis, Import Penetration, and Exchange Rate Pass-Through." *Quarterly Journal of Economics* 104 (2): 205–28.
Eaton, Jonathan, Marcela Eslava, C.J. Krizan, Maurice Kugler, and James Tybout. 2010. "A Search and Learning Model of Export Dynamics." Unpublished.
Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2011. "An Anatomy of International Trade: Evidence from French Firms." *Econometrica* 79 (5): 1453–98.
Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo. 2004. "World Trade Flows: 1962-2000." National Bureau of Economic Research Working Paper 11040.
Fujita, Masahisa, Paul Krugman, and Tomoya Mori. 1999. "On the Evolution of Hierarchical Urban Systems." *European Economic Review* 43 (2): 209–51.
Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics* 114 (3): 739–67.
Garmendia, Aitor, Carlos Llano, Asier Minondo, and Francisco Requena. 2012. "Networks and the Disappearance of the Intranational Home Bias." *Economics Letters* 116 (2): 178–82.
Helpman, Elhanan, Marc Melitz, and Yona Rubinstein. 2008. "Estimating Trade Flows: Trading Partners and Trading Volumes." *Quarterly Journal of Economics* 123 (2): 441–87.
Hidalgo, Cesar A., Bailey Klinger, Albert-Laszlo Barabasi, and Ricardo Hausmann. 2007. "The Product Space Conditions the Development of Nations." *Science* 317 (5837): 482–87.
Holmes, Thomas J. 2011. "The Diffusion of Wal-Mart and Economies of Density." *Econometrica* 79 (1): 253–302.
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.
Jia, Panle. 2008. "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry." *Econometrica* 76 (6): 1263–1316.
Krugman, Paul. 1980. "Scale Economies, Product Differentiation, and the Patterns of Trade." *American Economic Review* 70 (5): 950–59.
Krugman, Paul. 1987. "The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies." *Journal of Development Economics* 27 (1-2): 41–55.

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." Unpublished.

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.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler. 2013. "Gravity and Extended Gravity: Estimating a Structural Model of Export Entry." Unpublished.

Rauch, James E. 1999. "Networks versus Markets in International Trade." *Journal of International Economics* 48 (1): 7–35.

Rauch, James E. 2001. "Business and Social Networks in International Trade." *Journal of Economic Literature* 39 (4): 1177–1203.

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

Rossi-Hansberg, Esteban, and Mark L. J. Wright. 2007. "Establishment Size Dynamics in the Aggregate Economy." *American Economic Review* 97 (5): 1639–66.

Samuelson, Paul A. 1954. "The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments." *Economic Journal* 64: 264–89.

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

Young, Alwyn. 1991. "Learning by Doing and the Dynamic Effects of International Trade." *Quarterly Journal of Economics* 106 (2): 369–405.