# The dynamics of Mexican firms' export portfolio: a network analysis

Juan de Lucio (High Council of Spanish Chambers of Commerce)

Raúl Mínguez (High Council of Spanish Chambers of Commerce)

Asier Minondo (Deusto Business School)

Francisco Requena (University of Valencia)

**[VERY PRELIMINARY VERSION]**

**[COMMENTS WELCOME]**

## Abstract

We examine market entry choices of Mexican exporters, using a firm level data on merchandise trade over the period 2000-2009. We focus our enquiry not on the broad question of what determines a firm's ability to export, but on the subsequent question: given that a firm has the ability to export, what determines the choices they make about where to export? We find strong evidence that firms tend to enter new markets which are geographically close and culturally related to their prior export destinations. The contribution of the paper is to use a "revealed" measure of closeness between markets based on network analysis, which we argue to be superior to other gravity-type measures of geographic and cultural proximity. First, it captures all the extended gravity measures in one indicator. Second, it can be calculated at sector level so we can test whether path-dependence across destinations is sector-specific.

JEL Code: F1

Keywords: export market, network analysis, modularity, extended gravity

Acknowledgements: We thanks financial support from the Spanish Ministry of Science and Innovation (project ECO 2011-27619/ECON) and Generalitat Valenciana (project Prometeo/2011/098).

## Introduction

The international literature provides broad support for the assumption that sunk costs influence firms' export decisions. However, until recently firm-level research in this area has tended to treat export status as a binary variable - firms are either exporting or they are not. Hence empirical studies of entry into exporting have focused on the initial entry decision, particularly on identifying the firm-specific characteristics which set exporting firms apart from non-exporters. We focus our enquiry on a subsequent question: Given that a firm has the ability to export, what determines the choices they make about where to export?

Focusing on the behavior of already-exporting firms is essential for understanding the processes by which aggregate export value increases over time. Recent studies show that changes in the product-country mix of existing exporters account for the largest percent of net export growth in the long run, well above the share associated with net export entry and exit or net growth in existing relationships (see Bernard et al, 2009, De Lucio et al, 2012). For the case of Mexico, this paper shows that the changes in product-country mix of "new" regular exporters account for a large percent of net export growth (21 percent) over the period 2003-2009.

The literature points to the importance of sunk costs in determining firms' initial export entry decisions. At least theoretically, this argument seems equally persuasive for subsequent entries. Every geographic market provides new challenges for firms, including setting up distribution networks, and coming to grips with foreign consumer preferences and government regulations. However, firms may become more adept at handling these challenges over time, building up both general exporting competencies and market-specific knowledge. To identify the existence of relationship-specific sunk export costs, we look at whether firms' past export destination portfolio influences the choices they make about entry into new destinations (i.e. once a firm has exported to one country, where is it more likely to export next?)

We combine firm's own history of international engagement and variables measuring geographic and cultural proximity between unexplored and explored countries, i.e. the so-called extended gravity (Morales et al, 2011) or marginal distance (Lawless, 2010) variables. The novel feature of this paper is to use network analysis to construct a "revealed" indicator of proximity between pairs of countries. The advantage of this indicator is that it includes not only all proxies used for closeness in the previous literature but also any unobservable characteristic between pairs of countries that may affect how "close" pairs of countries are taking into account the entire network of destinations that firms have.

The analysis in this paper is related with different strands of literature that analyze firms´ exports dynamics. Recently, several theoretical papers have investigated the role of uncertainty and learning in export markets on the selection of export markets and the timing of entry. Albornoz et al (2012) and Nguyen (2012) develop alternative multi-market export models based on the idea that a firm's foreign demands are uncertain and correlated across markets. When faced with multiple destinations to which they can export, many firms will choose to sequentially export in order to slowly learn more about its chances for success in untested markets. Experimentation becomes an optimal strategy leading to path-dependence in firms´ export destinations. These models also predict that new exporters are more likely to add new export destinations and to exit from export markets than more established exporters. Chaney (2011) proposes a model of international network formation where firms obtain information about new potential partners from their current trading partners. The network formation game yields an equilibrium where firms' export destinations are path-dependent.[1]

Our paper is also related to the concept of "the geographic spread of trade", term originally proposed by Evenett and Venables (2002). They showed that geographic and linguistic proximity to an existing market was a consistently significant factor in determining expansion into new markets for sector-level exports from developing countries, implying a role for learning from existing export experiences. Using firm data, Morales et al (2011), Lawless (2011) and

[1] See Freund and Pierola (2010) and Eaton et al. (2010) for theoretical models that describe firms´ export dynamics within one market.

Defever et al (2011) show that entry to an export market is strongly related to export experience within the same region.

Using firm-level export data for Chilean manufacturing firms in the chemicals sector, Morales et al. (2011) show that the startup costs of accessing a new country are significantly determined by the countries to which a firm had previously exported. In other words, firms tend to choose new export destinations that are similar (geographically, culturally or economically) to destinations the firm is already exporting. Those "extended gravity" forces imply that the entry pattern of exporting firms is path-dependent.[2]

Lawless (2011) proposes a firm-level gravity model to examine the role played by the distance between various potential export markets. Using panel data on Irish exporters from 2000 to 2007, she finds that exporting experience in related markets is found to have a positive effect on entry and to reduce the probability of exit. When a "extended distance" variable is introduced— i.e. the distance from the closest existing market to the new market—its effect is so strong that it overrides the effect of distance from the home market, the standard measure of trade cost in the gravity model.

In a recent related paper, Defever et al. (2011) propose a simple spatial model to measure the importance of "extended gravity" variables on the decision of Chinese textile firms about spreading geographically their export portfolio. They use a pseudo-natural experiment to control for the timing of entry and exit into export destinations: the end of the restrictions on MFA products in 2005 and the possibility to start exporting to 25 EU countries, USA and Canada. Controlling for firm-product and destination specific effects and accounting for possible multiple new export destinations they find that the probability to export to a country increases by 15 to 38 percent for each prior export destination with a geographical or cultural link with this country.

The remainder of the paper is organized as follows. Section 2 provides the conceptual model for the analysis drawing based on reviews the empirical

[2] Albornoz et al (2012) and Morales et al. (2011) takes into account the arising option value of waiting to enter an additional export market after a first export decision. Taking into account the value of waiting considerably complicates the firm's problem and gives rise to a dynamic discrete choice problem.

literature. Section 3 outlines the data used in the paper, discussing the construction of the explanatory variables, while Section 4 presents the estimation approach. Sections 5 and 6 discuss the main empirical results and robustness checks, respectively. Section 7 concludes.

## 2. A simple model of destination-specific export participation by a firm

In this section we present a simple model of export participation into specific foreign markets by rational, profit-maximizing firms. Consider a firm that produces one good in the local market and sells part of the production abroad. There are several alternative markets and the firm has to decide which markets to export to. At any period, the exporting firm has the choice of entering into a number of markets if they did not export to those markets in the previous period.

Let $\pi_{igt}$ be firm *i's* profits from exporting to market g in year *t*. We assume that the expected profits of exporting to country g by firm i is a linear function of factors affecting the destination choice,

$$\pi_{igt} = I_{-g,t-1}\alpha + z_g\gamma + x_i\beta_g + \theta_g + \varepsilon_{ig}$$

where $\mathrm{I}_{-\mathrm{g,t}-1}$ measures the "mass" of information about destination g that firm i might obtain from previous exporting experience in other destinations. The variable $z_g$ is a vector of observable attributes of the destination-country, the variable $x_i$ is a vector of firm-specific characteristics, $\theta_{\mathrm{g}}$ is a vector of fixed-effects or destination-specific constants and $\varepsilon_{\mathrm{ig}}$ is a random term denoting the unobservable (by the researcher) unique profit advantage to the firm i from selling in the country g.

Let $\boldsymbol{y_{igt}} =(y_{i1t}, y_{i2t},\ldots,y_{iGt})$ denote the vector of firm *i's* current participation into export markets (g$=1\ldots G$). The variable $y_{igt}$ is a binary variable that indicates whether a firm exports to market *g* in period *t* ($y_{igt}=1$) or not ($y_{igt}=0$). This paper focuses on the decision of an exporter about expanding the portfolio of destinations **G**, given that it is already exporting. In the presence of learning-by-exporting, it is likely that the decision on exporting a new destination depends on firms' previous portfolio of destinations. In particular, our interest lies in the quantification of the effect of "proximity" between countries on the probability of entering a new export destination, $I_{-g,t-1}$. In particular we want to examine two types of measures: those based on gravity-type indicators and those based on network analysis.

## 3. Data and construction of variables

To bring our model to the data, we use transaction level customs data on the universe of Mexican exporters over the period 2000-2009.[3] In order to investigate the importance of past export portfolio on the decision about where to export next we select new exporting firms since 2003 that once start to export will carry on until 2009. For this sample of firms we know the complete past export destination portfolio. There are 6026 new exporters since 2003 that did not stop exporting.

Table 1 displays information on the number of trading firms, number of transactions and value (in million US$) in the 2000–2009 period. The number of regular exporting firms for the entire period is 5697, representing 20% of total firms, 46% of all transactions and 69% of all value of exports in 2009. The number of firms that start exporting after 2002 and do not stop is 6026 in 2009, representing 21% of total firms, 23% of all transactions and 21% of all value of exports in 2009. The weight of new successful exporters in Mexican exports is high and confirms the importance of the extensive margin as a source of growth of exports in the long run.

*[Insert Table 1 here]*

Our sample contains 6025 new regular exporters: 757 started to export in 2003, 948 in 2004, 1110 in 2005, 1283 in 2006 and 1928 in 2007. In Table 2

[3] Our database covers all Mexican firms' export transactions per country and HS classification 6-digit product for the 2000-2009 period. The firm identification code reported in the database changes from 2007 onwards. For the year 2007 we have data with the old firm classification and with the new firm classification. Matching firm-level country and HS 6-digit specific records we can establish a correspondence between the old firm classification and the new firm classification for firms that exported in 2007. For the rest of firms that exported in 2008 and/or 2009, we cannot know whether they are new exporters or they are firms that also exported in the 2000-2006 period.

we can examine the dynamics of the export destination portfolio of the new regular exporters. Considering year-to-year changes, the largest percentage of firms (about 60 percent) does not modify the export destination portfolio. Every year some firms only enter into new destinations (15 in 2007-08) depending on the period), others opt for only exiting (9 percent in 2007/08) and, finally, others decide to enter and exit simultaneously (17 percent in 2007/08). Compared to the sample of regular exporters over the entire period, the dynamics in the export destination portfolio are quite similar.

*[Insert Table 2 here]*

Table 3 presents the year-to-year changes in the geographical area of the destinations. Panel A covers the 5967 regular exporters and Panel B covers the 6025 new exporters. We observe high persistence (diagonal vector) for both the regular exporters and new exporters that do not stop exporting. The finding can be taken as evidence of path-dependence in the export destination portfolio.

*[Insert Table 3 here]*

### Endogenous variable

Our dependent variable is a firm-specific vector of export indicators **yit** = (yi1t;...; yijt; ...; yiGt) which indicates whether a firm exports to a specific destination g in year t. Notice that (1) the total number of potential destinations changes every year for each firm and (2) the number of potential destinations is constrained to those countries that the firm has never been before. In Table 5, we aggregate our dependent variable at the firm-year level and count the number of new destinations. 56 percent of the firm-year pairs report only one new destination, while 1.12 percent report 11 or more new destinations.

*[Insert Table 4 here]*

**"Revealed" connectedness between two foreign markets based on network analysis**

It is illustrative to represent the web of destinations of Mexican exporters through a network. In this network export destinations are nodes. Two nodes are connected by an edge if there is at least a Mexican firm that exports to both nodes. The weight of the edge is the number of firms that export to both destinations. To construct the network we use data for the year 2002. We only consider data from regular exporters, defined as firms that export in 2000, 2001 and 2002. The sample contains 14948 firms. As we do not include self-loops, we only use data on regular exporters that export at least to two different destinations. The reduced sample contains 6271 firms.

Figure 1 presents the network of Mexican firms' destinations for the year 2002. The network is highly dense, with 6271 edges between the 150 destinations of Mexican regular exporters.[4] The size of the node is correlated with the number of Mexican firms that export to that destination, and the size of edge is correlated with the weight of the edge. We can see that the most important destinations for regular Mexican exporters in the year 2002 were the US (12675 firms), Guatemala (1778 firms), Canada (1338 firms), Costa Rica (1318 firms) and El Salvador (1044 firms). The edges with the higher weight were Canada-US with 1206 firms exporting to both destinations, Guatemala-US with 1145 firms, Costa Rica-US with 915 firms, Costa Rica-Guatemala with 775 firms, and Colombia-US with 753 firms. All the nodes are connected in the network; this means that there is no destination where all exporters to that destination only exported to that destination. Each node has an average degree of 42; that is, as average, the total number of different destinations of all firms that export to a destination is 42. As expected, the destination with the highest degree is the US: 148 edges. There was only one destination that was not

[4] The network has 0.561 density. If the network had all possible edges density would equal 1.

covered by firms that exported to the US in 2002: Senegal. The US is followed by Chile (143 edges), Colombia (140 edges), Guatemala (140 edges), and Canada (134 edges). The preeminent place that these nodes occupy in the network is confirmed by other centrality measures, such as betweeness (how often a node appears on shortest path between nodes), and closeness (the average distance between the node and all other nodes). However, when we calculate eigenvector centrality, where a node's centrality depends on the centrality of its neighbors, Germany becomes one of the five most important destinations.

To explain the dynamics of Mexican firms' export portfolio, it is very important to determine whether there are distinctive communities within the network. If a Mexican firm starts to export to a country belonging to a community, we would expect the firm to expand its export portfolio within the community rather than outside the community. In a network, a community arises when the number of edges between some nodes is higher than the average. The community might arise due to the reasons that are captured in a standard gravity equation (such as speaking the same language or belonging to the same free trade area), or for reasons that are not captured in the gravity equation (for example, the existence of a large distributor that for historical reasons is specialized in a group of countries).

To identify the communities within our network, we first reduce the number of nodes in the network, focusing on the 26 most important destinations of Mexican regular exporters in the year 2002, which represent 27% of all exports. The reduction in the number of destinations also reduces the number of edges to 171. The community detection mechanism is based on the maximization of a modularity index (Newman, 2006). The modularity index compares the number of edges between two destinations with the number of edges that we would expect if edges were distributed randomly, conditional on the given degrees of nodes. Once modularity is calculated for each pair or nodes, nodes start to be grouped successively into a community until a total (community) modularity measure is maximized. Analytically:

$$Q = \frac{1}{2m} \sum_{ij} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \qquad (1)$$

where Q is total (community) modularity, m is the number of edges in the network, Aij is the number of edges between node i and node j, ki and kj are the degree of nodes i and j respectively, and δ(ci,cj) is a delta function that takes the value of 1 if i and j belong to the same community.

Figure 2 presents the different communities that are identified in the network of Mexican exporters' destination after applying the modularity maximization algorithm.5 We can identify seven different destination clusters. The first community (blue) is composed by large developed countries, such as the US, Canada, European Union Countries (Germany, Italy, Spain, France, Great Britain and the Netherlands), Japan and Australia. The second group (red) is composed by four South American countries: Argentina, Brazil, Venezuela and Colombia. The third group (dark blue) is composed by two Andean countries: Ecuador and Peru. The fourth group (green) encompasses three Central American countries: Panama, Honduras and Nicaragua; note, that other two Central American countries (El Salvador and Costa Rica) are classified in another community (orange). The sixth group (violet) is composed by two Caribbean Islands (Puerto Rico and the Dominican Republic). The seventh group (ochre) mixes different Latin American countries, such as Guatemala, Chile and Cuba. We can see that extended gravity variables, such as geography and income level have a role in the emergence of communities. For example, in the community formed by El Salvador and Costa Rica both countries are located in Central America; in the community formed by Puerto Rico and Dominican Republic both countries are islands; or in the community formed by European Union countries, the US, Canada and Japan the income level is high. However, we also observe that there are communities that seem to emerge for other reasons in addition to those incorporated in a extended gravity equation. For example, El Salvador, Costa Rica, Panama, Honduras and Nicaragua are all located in Central America; however, they are separated in two communities.

[5] We use the modularity detection algorithm incorporated in Gephi.

Ecuador and Peru are adjacent to Colombia; however, this latter country joins another community formed by Venezuela, Brazil and Argentina.

This analysis points out that variables included in an extended gravity equation might not exhaust the forces that might explain why exporters follow certain trajectories when expanding their destinations portfolio. Hence, a modularity measure derived from the network analysis can complement the extended gravity, and gravity, variables in explaining the dynamics of exporter's destination portfolio. In particular, we want to use the network analysis to derive an index, denominated as density, on how close a new export destination is from the rest of destinations served by a Mexican exporter. The expectation is that the larger the density around a new destination the higher the probability that the Mexican firm will start exporting to that destination. To derive this index, first we use the term in brackets in equation (1) to calculate the modularity of all Mexican exporters' destinations pairs. Note that this modularity will take a positive value if the weight of the edge between two destinations is larger than the one expected if edges were distributed at random, preserving the degree of nodes; in contrast, it will take a negative value if the weight of the edge between two destinations is lower than expected. It is important to emphasize that this modularity index measures the degree of closeness between two destinations above an expected value. Hence, this modularity index is not affected by other factors (e.g. economic size or income) that might make any destination attractive to Mexican firms; in this sense, the modularity measure is closer to extended gravity variables, than to standard gravity variables.

Figure 3 presents the histogram of modularity. Most of destination pairs have a modularity index close to zero: the weight of edges is similar to the one expected from a random distribution of edges, preserving the degree of nodes. The average modularity is 0.69 and the standard deviation 19. There are also some destination-pairs where the weight of edges is much larger than expected, and destination pairs where the weight of edges is much lower than expected. In the first group, the destination pairs with highest modularity are

US-Canada, Costa Rica-Guatemala, Guatemala-US and Germany-USA. For example, with regards the first destination pair, the modularity index indicates that the number of exporters than export both to the US and Canada is much larger than the one we would expect based on the number of exporters that have also US and Canada as destinations. This result denotes that US and Canada share some characteristics, such as language or belonging to the same regional agreement, that reduce the cost of entering in any of two markets once the exporter is present in the other market. In contrast, the destinations with the largest negative modularity are: Guatemala-UK, El Salvador-UK, Guatemala-Japan, Costa Rica-Japan and Costa Rica-Germany. For example, with regards to the first pair, the modularity denotes that Guatemala and the UK do not share characteristics and hence, few firms that decide to enter the Guatemalan market tend to extend their export activities in the UK.

*[Insert Figure 3 here]*

Using destination-pair modularity indexes, we calculate the density around a new destination for a specific exporter as the sum of the modularity indexes between the new destination and the destinations where the exporter is already present. Analytically,

$$d_{fi} = \sum_{j} m_{ij} t_{fj} \qquad (2)$$

where dfi is the density of firm f around the new destination i, mij is the modularity between destination i and destination j, and tfj takes the value on 1 if firm f already exports to destination j, and zero otherwise. If firm f already exports to destinations that are close to the new destination i, firm f will have a high probability to start exporting to the new destination i. In contrast, if firm f exports to destinations that do not share characteristics with the new destination i, the probability to start exporting to new destination i will be low.

We calculate the destination-pair modularity indexes for all exporters (all Network) and for 16 groups of firms classified according to their sector activity (See Table A3 for the list of sectors).

$$I_{-g,t-1} = \{NETWORK_{-g,t-1}^{all\ sectors}\}$$
$$I_{-g,t-1} = \{NETWORK_{-g,t-1}^{own\ sector}, NETWORK_{-g,t-1}^{rest\ sectors}\}$$

A detailed description of the variables and summary statistics for all variables can be found in Table A.1 in the Appendix.

**Connectedness between two foreign markets based on gravity-type indicators ("extended gravity")**

The firm's profitability will be correlated in markets which are geographically or culturally proximate to its previous export destinations. We make use of the past export portfolio decisions made by the firm in order to construct our measures of "extended gravity".

$$I_{-g,t-1} = \{N_{-g,t-1}^{distance}, N_{-g,t-1}^{border}, N_{-g,t-1}^{language}\}$$

The variable $N_{-g,t-1}^{distance}$ characterizes the countries' geographical relationship to prior export destinations of the firm. It is defined as the number of prior export destinations of firm whose capital city is less than certain number of kilometres away from the capital city of the destination under consideration. In the empirical section we investigate different distances (500, 1000, 1500, ...). We can also proxy for the geographical links between countries using a common border dummy variable, $N_{-g,t-1}^{border}$, which capture the number of prior export destinations of a firm with a common land-border with each new possible export destination. In addition to the information matrices based on geography, we also consider cultural closeness measures such as common language between export destinations. Specifically, the variable $N_{-g,t-1}^{language}$ includes the number of all destinations to which a firm exported in previous periods that share a common language with each potential new export destination. Data for the construction of the gravity and extended gravity measures comes from CEPII (www.cepii.fr/anglaisgraph/bdd/distances.htm). A detailed description of the variables and summary statistics for all variables can be found in Table A.1 in the Appendix.

## 4. Econometric issues and estimation results

An exporter will choose to export to a particular country if it will earn the highest possible profit. Formally, the gth country is chosen by firm i as the destination of exports if (we omit the subscript time t),

$$\pi_{ig} = \max(\pi_{ik}, k = 1 \ldots K)$$

If the firm-specific random terms are independently distributed, each with an Type I extreme value distribution, McFadden (1974) showed that the probability of a firm i to choose a destination g is

$$P_{ig} = Pr(\pi_{ig} > \pi_{ik,}\ j \neq k) = \frac{\exp\,(I_{-g,t-1}\alpha + z_g\gamma + x_i\beta_g + \theta_g)}{\sum_k \exp\,(I_{-g,t-1}\alpha + z_g\gamma + x_i\beta_g + \theta_g)}$$

where $P_{ig}$ is the population relative frequency of exporting to destination g. The estimates are obtained by maximizing the likelihood function, $L = \prod_i \prod_g P_{ig}$. The model described above is known as a conditional logit model (CLM). It is important to emphasise that the CLM does not allow explanatory variables that are not directly related to the choices. In our case, it means that we cannot estimate a single parameter to capture the impact of firm-specific characteristics or source-location characteristics on the firm's probability to export a particular destination. We estimate the CLM with and without country-fixed effects. Train (1986) shows that the inclusion of choice specific fixed effect contributes to improve the specification of the CLM as it reduces the risk of violation of the IIA assumption. What is left is a possible correlation across export destinations induced by destination-firm specific effects. We capture this aspect with our explanatory variables. We cluster standard errors at the firm-level. This takes into account unobserved within-firm correlation across destinations.

Table 5 reports estimates of the conditional logit allowing for simultaneous exports to multiple destinations but without the full set of country-specific dummies. Since we do not include country-specific dummies we can estimate a gravity-type logit model. Specification (1) reports to estimated coefficients of three standard gravity variable (GDP as proxy by economic size, distance to Mexico and dummies for common land border or common language) for the choice between all possible new additional export destinations. All the coefficients have the expected sign and all but one (border) are statistically significant at 1 percent. The transformation of the coefficients into odd-ratios for the discrete variables provides a easy way to interpret economically the estimates: the probability of a firm to export to a country that speaks Spanish is 95% (=exp(0.66)) higher than to export to a country that does not speak Spanish. For the continuous variables we can calculate the change in the probability to export to a new destination for a proportional change in the continuous explanatory variable. The probability to export to a country rises 0.7 when the GDP increases 1% and falls by 0.16 when distance increases 1%.

Specification (2) includes all the extended gravity measures to evaluate the connectedness between foreign countries. The variable distance area controls for prior exports in countries located within a 1500 kilometer radius around the capital of the chosen destination. The odds ratio of 1.192 =exp(0.176) implies an average increase of 19 percent in the probability of choosing a new export destination when we increase the number of prior export destinations which are in a 1500 kilometers distance area by one. Analogously, if we increase the number of contiguous export destinations that share a common border by 1, the probability of choosing a new destination increases on average by 85 [=exp(0.62)] percent. Notice that the extended border effect must be interpreted as an additional effect to the extended distance effect. Finally, the probability of choosing a new destination increases on average by 10 [=exp(0.099)] percent if the number of export destinations in the past portfolio that share a common language increases by 1. In all cases, the estimated coefficients are significant on the 1 percent level.

Specifications (3) to (6) introduce our measure of connectedness between foreign markets based on the concept of modularity. The coefficient on connectedness based on network analysis is always positive and significant at the 1 percent level. In column (6) the coefficient on the indicator based on networks of firms operating in the same sector is greater than the coefficient of the indicator based on networks of firms operating in sectors that are different to the one the firm operates. We interpret the results as evidence that "closeness" between markets varies for different types of products.

[To be completed] Table 6 reports estimates of the conditional logit allowing for simultaneous exports to multiple destinations and includes a full set of country-specific dummies. The sample is reduced because destinations that are chosen only once are excluded. The estimated coefficients have the expected sign and are statistically significant.

[To be completed] Explain here differences between CLM with /without fixed effects. We proceed as follows. We introduce simultaneously several extended distance variables with different distance bands. Following this approach, Figure 4 reports the odds ratios of a single regression with and without country fixed-effects and with different distance bands ranging from 0 to 4,000 kilometers, in 500 kilometers steps. The results show important differences in the odd-ratios.

## 6. Robustness analysis

[TO BE COMPLETED]

## 7. Conclusions

How do firms choose new export destinations? While there are many factors that are important for this decision, an empirical regularity strikes out: Firms tend to choose new export markets that are geographically close and culturally related to their prior export destinations. We quantify the effect of this path-dependence geographical pattern using Mexican customs data. We control for destination specific effects and account for possible multiple new export destinations.

Our baseline results (CLM with country fixed effects) show that the probability to export to an additional country increases by more than 40 percent for each prior export destination with positive connectedness, either measured using network analysis or extended gravity measures.

[TO BE COMPLETED]

## References


Akhmetova, Z. and Mitaritonna, C. (2012) "A Model of Firm Experimentation Under Demand Uncertainty with an Application to Multi-Destination Exporters", mimeo

Albornoz, F., Calvo-Pardo, H., Corcos, G., and Ornelas, E. (2012). "Sequential exporting", Journal of International Economics, 86, 17-31

Chaney, T. (2011). "The network structure of international trade", NBER Working Paper 16753.

Combes, P.-P., Lafourcade, M., and Mayer, T. (2005). The trade-creating effects of business and social networks: evidence from France. Journal of International Economics, 66(1):1-29.

Eaton J., Eslava, M., Krizan, C.J., Kugler M. and Tybout J. (2010), "A Search and Learning Model of Export Dynamics", Working Paper, mimeo

Fabling, R., Grimes, A. and Sanderson, L. (2009). “Whatever Next? Export Market Choices of New Zealand Firms”, Reserve Bank of New Zealand Discussion Paper Series DP2009/19.

Freund C. and Pierola, M.D. (2010), "Export Entrepreneurs: Evidence from Peru", Policy Research Working Paper Series 5407, World Bank.

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

Jackson, M. O. and Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? American Economic Review, 97(3), 890-915.

Lawless, M. (2009), "Firm export dynamics and the geography of trade", Journal of International Economics, 77 (2), 245–254

Lawless, M. (2011). Marginal Distance: Does Export Experience Reduce Firm Trade Costs?. Central Bank of Ireland Technical Working Paper 02/RT/11.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (eds.), Frontiers in Econometrics, New York: Academic Press, 105–142.

Morales, E., Sheu, G., and Zahler, A. (2011). Gravity and extended gravity: Estimating a structural model of export entry. mimeo, unpublished working paper.

Newman, M.E. (2006). Modularity and community structure in networks. PNAS, 103(23), 8577-8582

Nguyen, D. X. (2012). "Demand uncertainty: Exporting delays and exporting failures", Journal of International Economics, 86, 336-344 .

Rauch, J. and Watson, J. (2003). Starting small in an unfamiliar environment. International Journal of Industrial Organization, 21(7):1021-1042.

Rauch, J. E. (2001). Business and social networks in international trade. Journal of Economic Literature, 39(4):1177-1203.

Requena Silvente, F. and Castillo Giménez, J. (2007) "Information Spillovers and the Choice of Export Destination: A Multinomial Logit Analysis of Spanish Young SMEs", Small Business Economics, 73, 256-273.

Sheard, N. (2011). "Multimarket entry in exporting", mimeo.

Train, K. (1986). "Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand", Cambridge, MA: MIT Press.

Figure 1. The network of Mexican firms export destinations, 2002

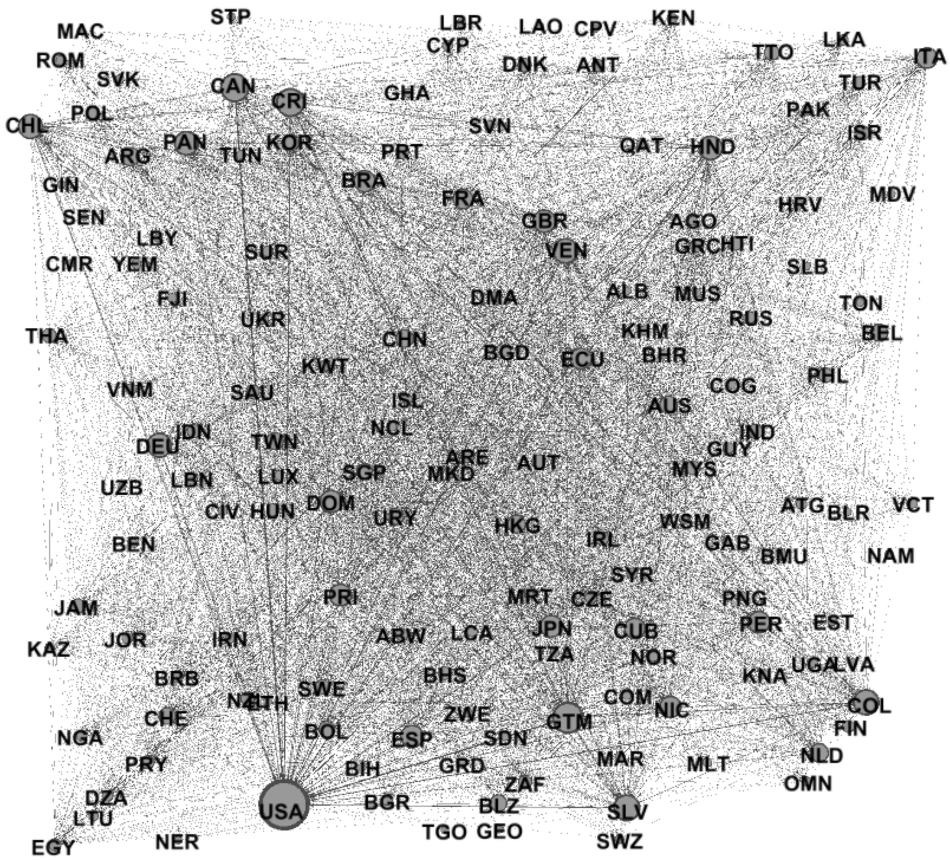


Figure 2. Communities in Mexican export destinations

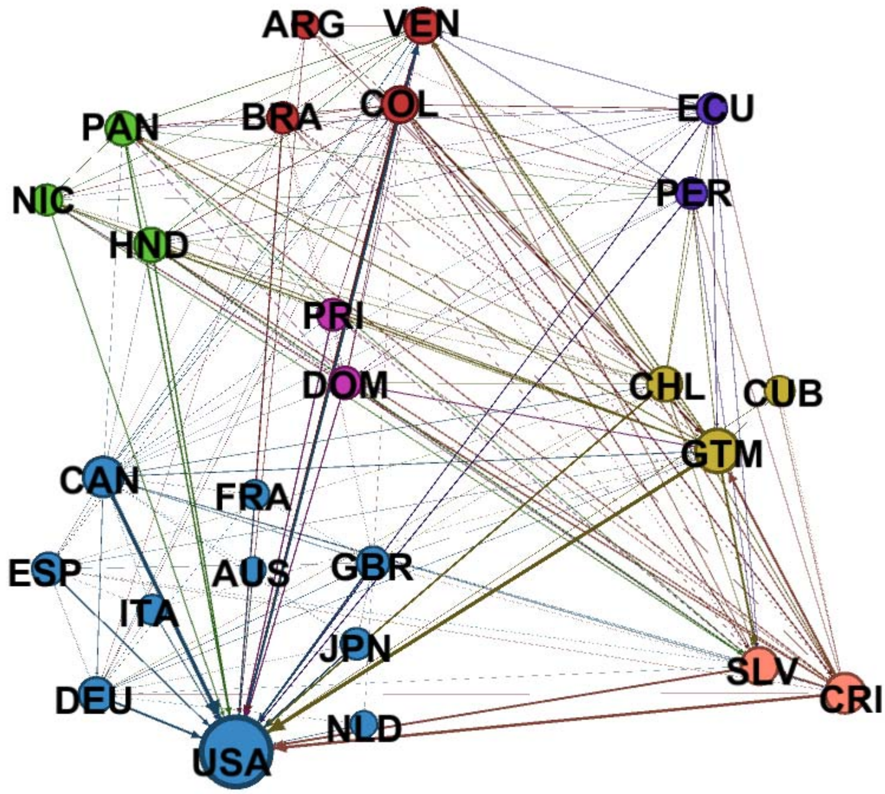


Figure 3. Histogram of destination-pair modularity (Year 2002; regular exporters)

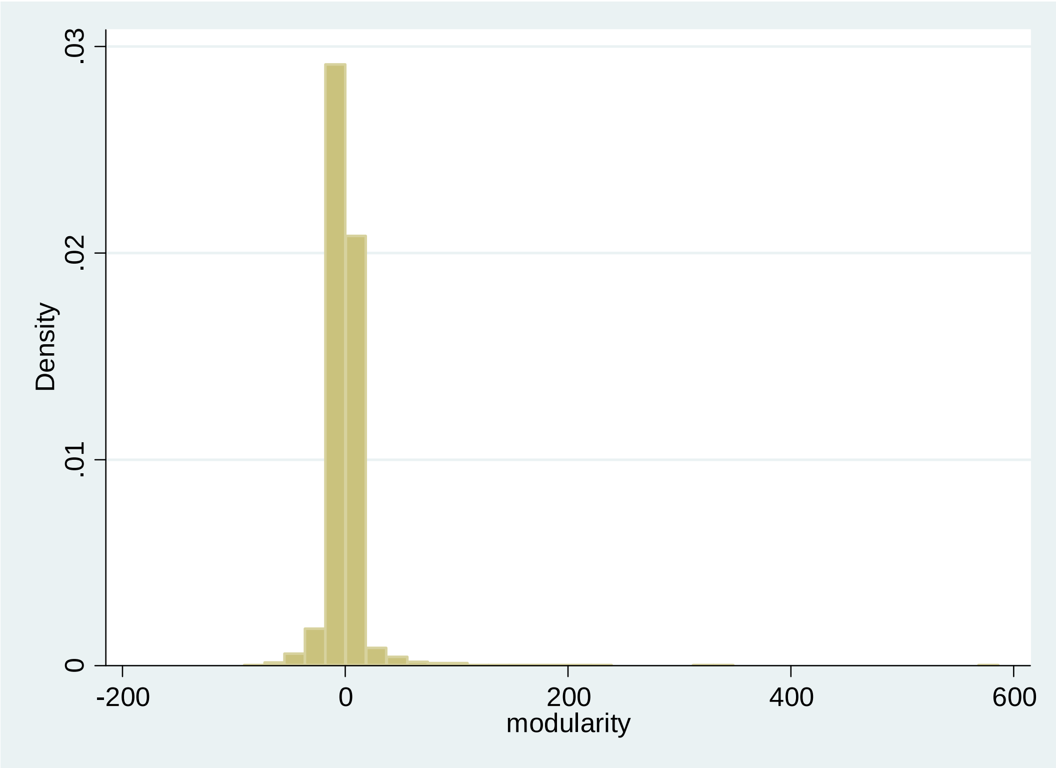
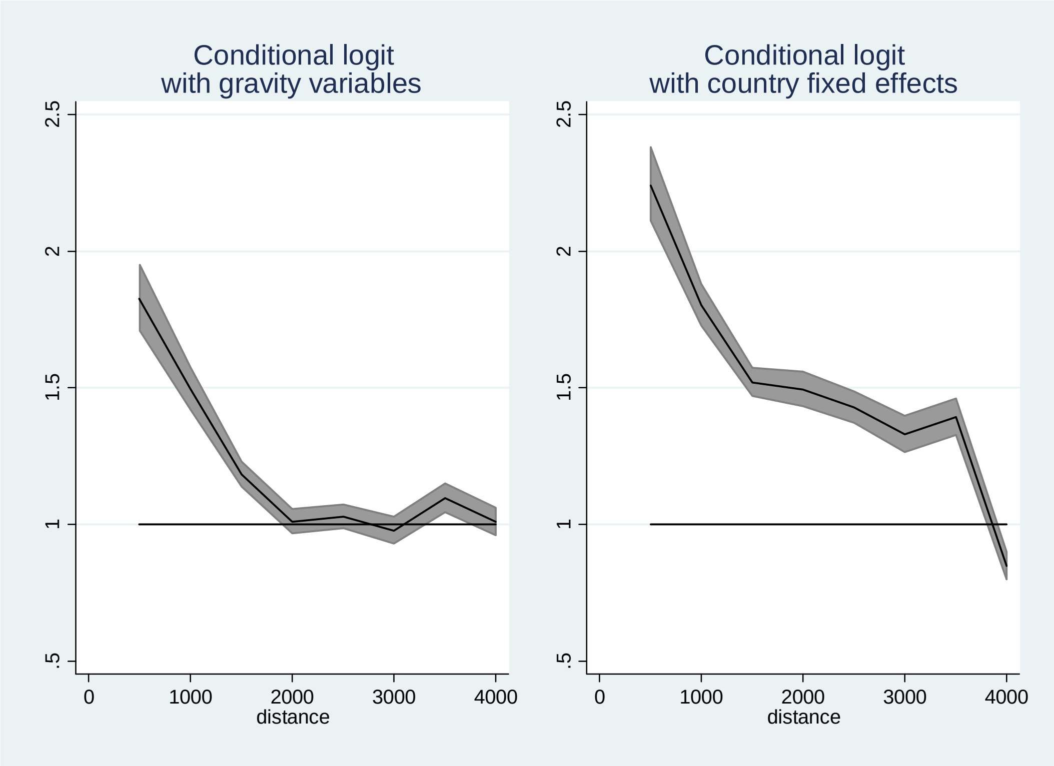


Figure 4. Impact of extended gravity based on physical distance on the probability to export to a destination.




Note: The black solid line gives the estimated odds ratios from a conditional logit which includes as regressors the number of contiguous previous export destinations within 500 kilometers wide distance bands from 0 to 4,000 kilometers. The sample used is the same as in Table 1 (right habd side figure) and in Table 6 (left hand side figure). The gray area denotes the 95 percent-significance band using clustered standard errors at the firm level.

APPENDIX

Table A.1. Summary statistics

| Regression without fixed effects (177 countries) | | | | | |
|---|---|---|---|---|---|
| | N. Obser | Mean | S.D. | min | max |
| DEPENDENT VBL | 1324391 | 0,0125 | 0,11 | 0 | 1 |
| | | | | | |
| GRAVITY | | | | | |
| gdp | 1324391 | 249407 | 874415,8 | 107 | 1,44E+07 |
| distance | 1324391 | 10801 | 3807,0 | 1427 | 17757 |
| border | 1324391 | 0,01 | 0,1 | 0 | 1 |
| language | 1324391 | 0,11 | 0,3 | 0 | 1 |
| | | | | | |
| NETWORK | | | | | |
| all sectors | 1324391 | 3,05 | 73,8 | -922 | 1845 |
| specific-sector | 1324391 | 0,44 | 7,7 | -108 | 238 |
| rest sectors | 1324391 | 2,89 | 67,1 | -815 | 1629 |
| | | | | | |
| EXTENDED GRAVITY | | | | | |
| min distance | 1324391 | 7535,5 | 4288,1 | 94,27333 | 19781,39 |
| distance area | 1324391 | 0,08 | 0,4 | 0 | 11 |
| border | 1324391 | 0,05 | 0,3 | 0 | 7 |
| language | 1324391 | 0,56 | 1,3 | 0 | 30 |

| Regression with country-fixed effects (113 countries) | | | | | |
|---|---|---|---|---|---|
| | N. Obser | Mean | S.D. | min | max |
| DEPENDENT VBL | 700803 | 0,0234 | 0,15 | 0,0 | 1 |
| | | | | | |
| GRAVITY | | | | | |
| gdp | 700803 | 457562,5 | 1162277,0 | 465 | 14400000 |
| distance | 700803 | 9423 | 4180,9 | 1427 | 17757 |
| border | 700803 | 0 | 0 | 0 | 1 |
| language | 700803 | 0 | 0 | 0 | 1 |
| | | | | | |
| NETWORK | | | | | |
| all sectors | 700803 | 7,09 | 101,07 | -922 | 1845 |
| specific-sector | 700803 | 0,89 | 10,51 | -108 | 238 |
| rest sectors | 700803 | 6,62 | 91,95 | -815 | 1629 |
| | | | | | |
| EXTENDED GRAVITY | | | | | |
| min distance | 700803 | 6689 | 4403 | 94 | 19781 |
| distance area | 700803 | 0,12 | 0,45 | 0 | 10 |
| border | 700803 | 0,08 | 0,32 | 0 | 7 |
| language | 700803 | 0,59 | 1,35 | 0 | 35 |

Tabla A.2. Classification of HS2 categories into 16 broad product categories

| Group | HS2 products | Name of group of products |
|---|---|---|
| 1 | 01-05 | Animal & Animal Products |
| 2 | 06-15 | Vegetable Products |
| 3 | 16-24 | Foodstuffs |
| 4 | 25-27 | Mineral Products |
| 5 | 28-38 | Chemicals & Allied Industries |
| 6 | 39-40 | Plastics / Rubbers |
| 7 | 41-43 | Raw Hides, Skins, Leather, & Furs |
| 8 | 44-49 | Wood & Wood Products |
| 9 | 50-63 | Textiles |
| 10 | 64-67 | Footwear / Headgear |
| 11 | 68-71 | Stone / Glass |
| 12 | 72-83 | Metals |
| 13 | 84 | Machinery |
| 14 | 85 | Electrical |
| 15 | 86-89 | Transportation |
| 16 | 90-97 | Miscellaneous |