Global Ownership Patterns

Asier Mariscal *

November 16, 2009

Abstract

I assemble a new dataset of global corporate acquisitions and sales at the firm level to analyze multinational expansion patterns. Four facts emerge:
1) the number of firms making a given number of acquisitions has a Pareto shape; 2) firms expand into industries and countries with a similar Pareto frequency; 3) entrants into unpopular industries or countries are larger, as measured by global sales; and 4) at the firm level, industrial diversification and internationalization are simultaneous. I first present a statistical model of balls falling into bins to understand the mechanics of multiproduct firms. I find that it quantitatively predicts some, but not all, of the observed entry regularities: multinational firms are approximately like a random collection of identical size goods. However, the model fails to explain firm entry by popularity of the destination. In light of this finding, I propose a multi-industry-country model where firms produce in their core competence activities. Heterogeneous producers branch out into new industries if they overcome the fixed costs of entry, and also choose whether to produce abroad. At the firm level, ability pushes companies to enter into new industries and countries, but distance in the physical and product dimension generates resistance to expansion. This theory generates an industry-country gravity equation for aggregate bilateral sales, closely matching the data.

*I am very grateful to Thomas Chaney, Ralph Ossa and especially to Samuel Kortum for their advice and encouragement. I also thank Robert E. Lucas Jr. and Nancy Stokey for useful comments and guidance through my PhD. Seminar participants at Capital Theory and International Economics working groups of The University of Chicago provided great feedback. Steve Crawford was of great help while constructing the dataset. Financial support from Fundación Rafael del Pino, La Caixa and the Bank of Spain is gratefully acknowledged. The usual disclaimer applies. Please send comments to asier@uchicago.edu.

1 Introduction

The activities of multinational firms are the most prominent feature of globalization. Table 1 summarizes data on aggregate exports and foreign direct investment, FDI, by affiliates of multinational firms. Sales of foreign affiliates represent nearly double the value of world exports[1]. Interestingly, affiliates concentrate on selling to local markets rather than on exporting. In fact, only 14% of affiliate sales were exports. Together, these facts point to FDI as being quantitatively more relevant than exports as means to reach foreign markets.

	Percentage
World sales of foreign affiliates/world GDP	58
World exports of goods and non-factor svces./world GDP	23
World exports of foreign affiliates/world exports	35
World exports of foreign affiliates/total sales of affiliates	14

Table 1: World trade and sales by foreign affiliates, UNCTAD

This paper presents a new dataset and novel facts on the activities of multinational firms, provides a simple statistical model to understand the patterns in the data and presents an economic model that replicates the observed facts. In the empirical part, I construct a new firm level dataset on international acquisitions and global sales of the acquiror to analyze the expansion patterns of multinationals from and into multiple country-industry pairs[2]. I then uncover the determinants of multinational companies expansion across industries and countries. In particular, I ask the following questions. How do multinational companies expand? Are industry and country dimensions similar? What firm, industry, and country characteristics influence expansion?

[1]Foreign direct investment, FDI, is recorded when either a merger or an acquisition occurs or when greenfield investment is performed. The former occurs when a foreign citizen either acquires more than 10% of the capital of a company, while the latter when a new plant is opened.

[2]The dataset combines two different sources. The fist source, SDC Platinum, provides data on international mergers and acquisitions, M&A, while the second, Worldscope, has accounting data for firms that are listed in a stock market. For more details see section 2

The first contribution is to provide evidence on multinational firms' FDI *expansion* patterns. Unlike the previous literature, I look at firm population distributions. Four main new facts emerge.

The first finding is that few firms do many acquisitions and most just a few; more precisely, the number of firms that make exactly n acquisitions is proportional to 1/n. This is known as a Power Law[3]. To my knowledge, it is the first time that such a relation has been documented in the context of FDI originating from multiple sources[4]. This is an interesting fact because it is well known that firm size follows a pareto density, see e.g. Axtel (2001), Gabaix (2008) or Sutton (1997), and so the data suggests a connection between the two.

The second fact looks at the number of acquisitions in two independent dimensions: industry and country. Interestingly, the same Pareto distribution emerges when one looks at the number of *different* industries or countries. This is surprising because the microeconomic determinants of entry into an industry are expected to be very different from entry into countries. Still firms acquire affiliates in both dimensions with a similar intensity. The result suggests there is a common driving force pushing for expansion in both the country and the industry dimension.

Third, to analyze the determinants of entry into new industries and countries, I again open up both dimensions independently. For each dimension, I look at the relation between sales and the number of entrants in a given destination, which is a country or an industry depending on the dimension we are looking at, *conditional* on the multinationals' home industry or country respectively. I find that entrants into less popular destinations are characterized by higher average global sales. To fix ideas, take the industry dimension: conditional on the main business line of a firm, firms that enter

[3] I obtain the Pareto exponent using the method described in Gabaix and Ibragimov (2007).

[4] See Irrarazabal, Moxnes, and Opromolla (2008) for related evidence on Norwegian FDI.

industries where few other firms from their home industry do are large. Firms that overcome the larger implicit hurdle required to enter a difficult new industry also have larger global sales on average. The same basic pattern holds both for industries and countries.

Fourth, I find a positive relation between entry into multiple countries and industries, suggesting that firms expand similarly in both dimensions exploiting profitable opportunities in available markets.

The second contribution is to use a statistical model of balls and bins to ask the question of how are multinational firms *like* a random collection of identical size goods. The physical experiment is as follows. Take a given number of balls and throw them independently into K bins. Each ball lands in a category k, according to some probability p_k. Balls are the equivalent of acquisitions and bins are either countries or industries, depending on the experiment. The idea is that larger firms are represented by more balls, but each ball has the same size for all firms and the structure of the bins is constant. Using this experiment, I can separately study the role of the distribution of the number of acquisitions and how these are allocated across industries and countries. Hence, I extend Armenter and Koren (2008) by obtaining results for a wider set of implications which are firm population distributions as opposed to just aggregate means. The model is successful in quantitatively matching some, but not all, of the regularities uncovered in the data. Three lessons are drawn. First, the model opens an avenue for future research in the theory of firm growth. Second, the firm size distribution is key in the success of the model. Third, the model fails in relating firms sales and entry patterns by the popularity of the destination, revealing that further firm heterogeneity is needed to explain those entry patterns.

The third contribution is to develop an economic model that captures firm selection into markets where the statistical model fails. It is a unified theory of industry and country firm entry choice based on Chaney (2008). In the model, firms concentrate their activities around their *core competency*, which is defined by the country of origin and industrial expertise. It uses the following structure. Firms are born in a country-sector pair and are heterogenous in their cost. Each firm receives a random number of goods to produce in each and every sector and it's production costs are larger in more different sectors and in more distant countries. For each good, there are bilateral industry-country fixed costs of production that induce selection into markets so that only the most productive firms enter in more remote industries and countries. Hence, the key new ingredients of the model are a random number of products at the firm level and asymmetries in the industry-country entry costs. Explicitly modeling core competencies, Eckel and Neary (2010) and Mayer, Melitz, and Ottaviano (2009) deal with multiproduct exporters but consider asymmetries only in the country and industry dimension respectively[5]. Bernard, Redding, and Schott (2009a) model the dynamics of product switching at the firm level but do not attempt to explain production of technically related goods within firms.

The FDI literature has typically not dealt with firms that produce in several different industries[6]. Navaretti and Venables (2004) contributed to the evidence of multinational companies' behavior and modeled mergers and acquisitions, M&A, and greenfield investment, the two components of FDI, as having different fixed costs. Using aggregate bilateral measures of FDI, Ramondo (2006) uses a model of head-to-head competition and decreasing returns to scale at the firm level to deliver implications about sales and

[5]Both papers have different goals. The former is concerned with multiproduct firms in Cournot competition while the latter analyzes how destination country economic conditions affect the product mix of exporters.

[6]Although not directly related to my paper, relevant to the field are the theoretical foundations of the property rights approach to multinational ownership, Antràs (2003) and Antràs and Helpman (2004), as well as Nunn (2007), which provides evidence on the effects of input contractibility on international trade. Also relevant, are Carr, Markusen, and Maskus (2001) who simultaneously model vertical and horizontal FDI and provide simulated results for the nonlinearities the model delivers.

the number of affiliates for each bilateral country pair and performs counterfactual experiments of bilateral tax reductions. Focusing on M&A only, Head and Ries (2008) propose a proximity-ability trade-off to explain headquarter ownership in a multicountry model and use aggregate bilateral M&A to estimate a structural gravity equation that resembles the one in the trade literature. The so called proximity-concentration is the basis of Helpman, Melitz, and Yeaple (2004), who propose a model of heterogenous producers that choose to access a market through exports or FDI. Nocke and Yeaple (2007) analyze the same margin but also consider the effect of combining firm abilities through mergers. I do not focus on the proximity-concentration trade-off since I want to emphasize the asymmetries in industry and country entry patterns of multinational firms, and furthermore half my data sample is services where such a trade-off doesn't apply. Nocke and Yeaple (2008) focus on the assignment of firm resources that occurs through mergers. While I produce a rich set of firm patterns that I go on to model, intrafirm reallocation of resources is something I do not observe.

The related empirical contributions are more limited in number. Feinberg and Keane (2005) used BEA data on US multinationals to show evidence on the small fraction that affiliates ship back to parents. The result is in line with Hortaçsu and Syverson (2007) who find the same for US plants. They both put an upper bound on the importance of intrafirm trade for vertical links. Alfaro and Charlton (2010) is the only reference that has comparable data to mine but uses it in a very different way. They use a large ownership cross-section of companies to regress measures of multinational activity on bilateral industry-country factor intensity variables in a comparative advantage regression. They conclude that SIC proximity matters for ownership. Eaton, Kortum, and Kramarz (2009a) use export shipment data and sales for French firms to estimate an extended Chaney (2008)-Melitz (2003) model.

The remainder of the paper is organized as follows. Section 2 of the paper includes an explanation of the construction of the dataset and the empirical regularities I uncover. In section 3 I introduce the statistical model and in section 4 the economic model. Section 5 concludes.

2 Evidence

2.1 Firm level data sources

I combine two datasets with complementary information. The first data source is SDC Platinum, which is the most comprehensive data source on international mergers and acquisitions, M&A. It has been recording deals since 1985 to date. Specifically, industry, country of origin, deal value and several other data items[7] are collected for target and acquiror as well as their Ultimate Parents, UP.

The second source, Worldscope, tracks listed companies in 59 countries since 1980. It is the source for financial and balance sheet data. Both datasets are matched using common identifiers: sedol and datastream code[8]. The result is a micro dataset at the firm level on global M&A deals from 1985-2007 with global sales data for the acquiror UP[9]. Table 2 presents an overview of the data.

Variable	Observations
Number of deals	109,381
Number of deals with sales data for acquiror	59,753
Number of companies	32,605
Number of companies with sales data	9,293

Table 2: Data overview

There are nearly 110,000 completed cross-border deals in the dataset of which around

[7]See table 5 for all data items in SDC Platinum.
[8]Breinlich (2008) uses a similar procedure to match SDC Platinum and Compustat Global.
[9]The data item in Worldscope is wc1001.

60,000 have acquiror UP sales data. The number of companies refers to ultimate parents, which will be the basic firm unit in the empirical analysis. Figures are constructed using the most data available, so if sales data are not necessary for a given graph, all deals are used[10]. Table 9 shows that nearly 90% of deals are acquisitions the remaining being mergers so I will frequently refer to deals as acquisitions. As shown in table 8, the acquiror UP's in the dataset come mainly from developed economies. Because a large fraction of inward and outward FDI takes the form of M&A in developed countries[11], the dataset is likely to be representative of international investment to and from rich countries. In fact, over the 1985-2001 period and for 29 OECD countries, Head and Ries (2008) show that the ratio of SDC M&A data to UNCTAD FDI data is 0.82 for inflows and 0.7 for outflows[12]. The map in Figure 9 shows the distribution of acquisitions originated in a given country, confirming that OECD countries are well represented.

2.2 Empirical regularities

We start by looking at a Pareto distribution for the number of acquisitions in Figure 1. Note how the number of firms with given number of acquisitions decays linearly on a log-log scale. One is tempted to think that the firm size distribution is lurking behind such a regularity: that larger firms acquire more and in proportion to their size. Surprisingly, this regularity can be found again in Figure 2, which reports the number of firms that enter into any given number of different industries[13] or countries. Note

[10] While I include all completed deals, in fact, 95% of them are classified as FDI using the standard definition, ie, the percentage of capital acquired is larger than 10%. Results are robust to dropping observations that do not conform to the FDI definition. In terms of actual corporate control, 79% of deals with percent acquired 50% or more. I exclude LBOs, spin-offs, recapitalizations, self-tender offers, exchange offers, repurchases and privatizations. Results are robust to dropping merger deals.

[11] See UNCTAD (2000).

[12] See table 2 in their paper. In another perspective on the data comparability, the correlation of FDI and the M&A data from SDC is 0.9, for both outflows and inflows.

[13] The figure reports the number of different 4-digit SIC industries in which a firm has acquired. Nearly identical results are obtained when I construct this figure at 2-digit SIC.

the remarkable similarity between the country and industry figures. While entry into industries is theoretically driven by very different primitives from entry into countries, there seems to be an underlying force pushing with a very similar intensity in both directions. The Pareto shape of the figures is also noteworthy, showing that few firms go to many destinations and most go to a few. It is known that the upper tail of the sales distribution is Pareto and that firms that engage in international businesses are much larger than the average. It follows that a firm effect, whether we call it quality, efficiency or ability, is a candidate responsible for the similarity of the industry and country figures and the inherited Pareto shape.

So the next step is to figure out the role of ability. I use global sales as a proxy for efficiency because I do not observe sales by destination or industry. Following Eaton, Kortum, and Kramarz (2009a), Figure 3 shows the relation between selling in k-or-more destinations and average global sales. Again, both industry and country dimensions have very similar Pareto shapes. A possible interpretation of this figure is that more able firms enter more markets. However, in this figure there is a mechanical relation since firms that are present in more destinations will be on average larger. To address this potential reverse causality, I look at the data from the perspective of the implicit popularity of destinations.

Figure 4 shows the relation between average global sales and entry patterns by destination popularity. It is constructed in the following way. From each source, be it an industry or a country, I rank all destinations based on the number of entering firms. Then, I compute average sales of firms that enter each ranked destination. Each dot then represents firms from all sources into their k-th most popular destination and hence the southeasternmost dot is rank number one and it has the largest number of firms entering; the closest dot to the west is rank number two and so on. To provide an example, let's look first at table 6 which shows the most popular country destinations for firms from USA, Japan and Germany. Now, in the southernmost dot, there are US firms going to UK, and German and Japanese firms going to US, since, from their countries of origin those are the most popular destination countries. This is done for all ranks.[14].

The figure shows a negative relation between popularity of destinations and sales of entrants, ie, entrants into unpopular destinations are on average larger than entrants into more popular destinations. Again, note the similarity between industry and country figures and it's Pareto shape. This figure again points to a common underlying determinant of entry into countries and industries. Note that there is no in-built relation between market popularity and the scale of global sales. A company entering an unpopular destination, will automatically increase its sales by some amount but, being unpopular, if anything that amount would be smaller and not larger than if the destination was more popular. So global sales in the figure are measuring a firm size effect that is independent from destination market size.

Figure 5 shows that the relationship just described disappears when I do not condition on industry or country source. I interpret this result to mean that the force that impedes entry into a destination is source dependent, ie, there is an implicit bilateral distance effect. What industries are more frequently entered by a firm depends on it's "home" industry, and what countries are more frequently entered by a firm depends on it's "home" country. For example, a firm in the "Plastic materials" industry, SIC 2821, will be more likely to enter "Chemicals and Chemical Preparations", SIC 2899, than "Industrial Machinery and Equipment", SIC 5084. Analogously, a french firm is more likely to go to Belgium than to China, whereas for a Taiwanese firm the opposite would be true.

The similarity between the industry and country evidence suggest that firms take

[14]See table 7 for a list of industries ranked by popularity regardless of industry of origin.

industries and countries as sources of profits. However, the previous figures were one-dimensional in that they looked at either country or industry entry in isolation, abstracting from the other dimension. So it could be that some firms specialize in industry expansion and others in country expansion. They could also expand unrelatedly or on both dimensions simultaneously. If any of the two former expansion strategies were observed, it would cast doubt on the idea that industries and countries are similarly used as means to enhance profits and that a core ability drives expansion.

Figure 6 shows the positive link between industry and country expansion for a given company. Companies that tend to acquire in multiple industries also do it in multiple countries. This supports the similarity of the country/industry entry patterns previously shown as well as reinforces the view of acquisitions as a means to replicate activity. In other words, firms don't care much whether they use their ability to enter into a new industry or a new country, they just try to extend their know-how into as many activities and locations as possible.

If this view is correct, just like physical geography imposes limits to expansion, we should find firms are clustering certain business lines. Firms should be bundling together the activities that they know how to perform. We already got a sense that such a pattern is important by looking at how figure 4 changes by not conditioning on the "home" industry. Another way to uncover such a pattern is to look at the rough data and ask whether the observed industry clustering that we observe could arise from chance. In categorical analysis, contingency tables are used to measure association and test whether populations fall into categories in a statistically independent way. For illustrative purposes, table 3 presents the distribution of deals across acquiror ultimate parent and target one-digit SIC industries. The first fact that stands out is the intensity of the diagonal, ie, horizontal deals are extremely prevalent at this level of aggregation[15].

[15] They still are the most prevalent one at 4-digit SIC as will be shown in subsection 6.5.

Second, given an acquiring industry, the fraction of deals into each target industry is far from uniform.

	Acquiror UP SIC								
Target SIC	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Total**
1:Mining, Construction	64.3	3.0	3.2	3.9	4.7	4.9	1.2	4.1	7.2
2:Manufacturing	6.1	65.3	7.0	4.4	10.7	12.2	2.5	6.8	17.1
3:Manufacturing	9.5	9.2	62.5	7.1	14.1	13.3	5.8	11.2	21.4
4:Transport, Communication	5.2	2.1	2.9	58.7	4.5	7.5	4.9	3.5	8.9
5:Trade	4.3	9.5	9.0	6.0	49.2	6.7	4.7	3.8	9.5
6:Finance, Insurance	4.3	2.8	3.5	3.8	5.3	38.2	4.2	5.0	13.2
7:Business, Hotels, Recreation svces	3.3	5.1	9.1	13.4	9.6	13.2	68.9	18.7	17.2
8:Health, Legal, Engineer. svces	3.1	3.1	2.8	2.7	1.8	4.1	7.9	47.0	5.5
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 3: Contingency table at one digit SIC

To formally measure the degree of association formally table 4 shows the Chi^2 statistics[16] of association at 2-digit SIC. I present results for all deals and for non-horizontal deals only. In both cases I obtain a high value for the statistic and consequently a negligible P-value, indicating a very small probability of the data having arisen from independent distribution of deals into categories[17]. This results are in line with Bernard, Redding, and Schott (2009a), who, using data for US exporters, find a strong tendency for product co-production, and with Alfaro and Charlton (2010) who find SIC proximity to be relevant in international ownership links.

	Observations	Chi^2	P-value
All deals	109381	1174384	0.000
Non-horizontal deals	68796	116845	0.000

Table 4: Contingency table: Association statistics at 2 digit SIC

The last piece of evidence that will be important for the modeling in section 4 is presented in Figure 7. The number of entrants rises with destination country size,

[16]The statistic is computed as $\sum_{i,j}\left(\frac{(Observed_{ij}-Expected_{ij})^2}{Expected_{ij}}\right)$ and it is the standard statistical test for association of categorical variables. The statistic has a Chi square distribution with degrees of freedom, dof=(Number of rows-1)(Number of columns-1)

[17]See section 6.4 for another perspective on this issue.

as measured by GDP, with a slope in log scale very close to one. This pattern is incompatible with Eaton and Kortum (2002) but in line with Chaney (2008) and Melitz (2003), and hence it pushes towards a model that includes fixed costs of acquisitions.

3 A balls and bins model of FDI

In this section, I use a purely statistical model to understand how the data compares to a simple mechanical model. The physical experiment consists of placing balls into mutually exclusive categories of different size. After a number of balls has been thrown into the set of bins, the joint distribution of the number of balls in each category is given by the multinomial probability function. The economic interpretation is that larger firms draw more balls, which, depending on the experiment, are placed into country or industry bins in proportion to their size. In expected terms, larger countries, as well as larger industries, attract more balls. The model will take as given the number of balls distribution, i.e. the number of acquisitions cross-section distribution, and it will make predictions about all the regularities introduced in section 2. It is an extension of Armenter and Koren (2008), who focus on understanding single destination and single product exporters as well as the number of zeros in bilateral trade.

The balls and bins model is a useful tool for two reasons. First, when the model misses a fact we can learn about the economics that are driving the results and the properties that the economic model cannot have. Second, when the model matches a fact, it informs us about what properties can deliver a successful model of firm random growth, thus opening a new research path. Let me elaborate on this. On the one hand, this model belongs to a tradition in economics that uses mechanical models to understand data regularities[18]. It takes the distribution of firm size as given and makes

[18] For example, see Ellison and Glaeser (1997) and Head and Ries (2005).

predictions about the allocation of goods into industry-country categories. On the other hand, there is a large literature exploring what mechanics of firm growth can deliver fat tailed distributions for firm sizes, see e.g. Simon (1955), Sutton (2001), Klette and Kortum (2004) and Gabaix (2008). The success of the balls and bins model in quantitatively explaining entry frequencies suggests that both statistical mechanics could be merged into a stylized model of firm growth through multiproduct and multinational operations. Since most of aggregate trade and FDI is explained by the extensive margin rather than the intensive margin[19], such a model could explain the firm size distribution, international and industrial branching, and aggregate bilateral sales.

Formally, there are $n \in \mathbb{N}$ balls and $K \in \mathbb{N}$ mutually exclusive categories indexed by $k \in 1, 2, ..., K$. Balls land in categories according to their size p_k, with $0 < p_k < 1$ and $\sum_k p_k = 1$ and each ball is placed into a bin independently from all other balls. The experiment can be represented by the multinomial probability distribution[20],

$$Pr(n_1, n_2, ..., n_K) = \frac{n!}{n_1!n_2!...n_K!} p_1^{n_1} p_2^{n_2} ... p_K^{n_K}, \tag{1}$$

which gives the probability of observing n_1 balls in the first category, n_2 in the second, etc... after $n = \sum_k n_k$ balls have been thrown. Depending on the experiment bins will represent industries or countries and the number of balls will be the number of acquisitions. I will use this model to obtain information on how acquisitions made by a firm are allocated to industries or countries, and, using the cross-sectional number of acquisitions distribution, make predictions about the population distributions.

[19]Here, the extensive margin is the number of firms and the intensive margin is sales per firm. Eaton, Kortum, and Kramarz (2009b) provide evidence that given destination country market size, a higher French market share in a destination typically reflects 88 percent more firms selling there and 12 percent more sales per firm. I find the same basic patter in my dataset.

[20]At the goods level, Chaney (2008) and Eaton and Kortum (2002) imply a multinomial probability distribution.

As mentioned, to calibrate the model two inputs are needed: the bin sizes and the number of balls distribution. For the former, remember bins are either countries or industries. I take each to be independent of the other and calibrate the category size k by calculating the proportion of acquisitions in k, $p_k = \frac{\text{Deals into category k}}{\text{Total number of deals}}$[21]. For the latter, I calibrate the number of balls distribution, $P(N = n)$, by dividing the number of firms in with a given number of acquisitions by the total number of firms.

Having calibrated the key elements of the model, we are ready to compute the model analogue to each figure in section 2. Let's start with Figure 2. In that figure, the number of firms is related to the number of categories they reach. Defining Y to be the number of categories reached and N the number of balls, we first need to compute the probability that exactly m categories are given that n balls were thrown, $P(Y = m|N = n)$. Then we can move to the population moment, $\sum_n P(Y = m|N = n)P(N = n)$, using the observed number of acquisitions distribution, $P(N = n) = \frac{\text{Number of firms doing n deals}}{\text{Total number of firms}}$. To obtain the number of firms with exactly m non-empty bins, I use $M(N = n) =$ Number of firms doing n deals to compute $M(Y = m) = \sum_n P(Y = m|N = n)M(N = n)$.

But what is $P(Y = k|N = n)$? It is an outcome of the model that we need to compute. Defining X to be the number of empty bins, we have that probability that exactly k bins are empty, conditional on n balls,

$$P(X = k|N = n) = \sum_{a \in a(k)} P_{a_1 \ldots a_k} = \sum_{a \in a(k)} \left(1 - \sum_{j=1}^{k} p_{a_j}\right)^n - \binom{k+1}{k} \sum_{a \in a(k+1)} \left(1 - \sum_{j=1}^{k+1}\right)^n + \binom{k+2}{k} \sum_{a \in a(k+2)} \left(1 - \sum_{j=1}^{k+2}\right)^n + \cdot + (-1)^{m-k} \binom{K}{k} \sum_{j=1}^{K} p_j^n, \quad (2)$$

[21]There are 193 countries and 924 industries with at least one deal.

where $\sum_{a\in a(k)}$ denotes summation over all subsets $a = (a_1, a_2, ..., a_k)$ of k integers from the $1, 2, ..., K$ categories. Note there are $\binom{m}{k}$ terms in the sum $\sum_a^{(k)}$[22]. From this expression, we can obtain $P(Y = m|N = n)$ since $Y = K - X$ but it is not so helpful in terms of implementing it. The reason is the very high number of subsets that are needed to compute even for small K[23]. Since the formula is not operative I simulate the model. For each n, I conduct 100,000 simulations for each of which I compute the number of non-empty bins. Then the estimated $P(Y = m|N = n)$ is just the fraction of simulations with exactly m non-empty bins. Figure 10 presents $P(Y = m|N = n)$ for selected m and n. Note how increasing the number of balls shifts the distribution to the right, but each time by less.

To obtain the simulated number of firms with exactly m non-empty bins, I compute $M(Y = m) = \sum_n P(Y = m|N = n)M(N = n)$. Figure 11 shows the results as they compare to the data. Simulated and observed data have a very similar slope. The size distribution dominates in expression $M(Y = m)$. However, in the upper tail of the country figure the number of firms decreases very fast. While in the beginning more balls means new categories are filled almost at the same rate, as categories are filled, it is less and less likely that a new category will become non-empty. This effect and the fact that few firms draw many balls generates the tail behavior. The number of categories is relevant too of course. Keeping constant the number of balls and say with a uniform bin size, the more categories the experiment has, the easier that balls fall into different bins. This is the reason the industry figure looks different from the country figure in the upper tail.

To construct Figure 12, since the statistical model is silent about sales, I need an assumption on the relation between the number of balls and the value per ball. I

[22] See appendix section 6.6 for the derivation of this result.

[23] Even for a small number of balls, the number of subsets for K categories soon approaches 2^K.

rescale the expected number of balls using average global sales per acquisition. The assumption is one of proportionality. It gives economic content to the question: are multiproduct/multinational firms a random collection of goods?. By assuming the number of acquisitions and total sales are proportional at the firm level and across firms, I am implicitly assuming no heterogeneity in ability. A large firm is just more goods, but each of them with the same size as those of a smaller firm. For this figure, I also need to compute the probability that at least k bins are non-empty when n balls have been thrown, $P(Y \geq k|N = n)$, and obtain its population number of firms counterpart, $M(Y \geq k) = \sum_n P(Y \geq k|N = n)M(N = n)$. Using the same method I have used for Figure 11, I know $P(Y \geq k|N = n)$. I also need the expected number of balls given that k-or-more categories have been filled,

$$E(N|Y \geq k) = \sum_n n \frac{P(N = n, Y \geq k)}{P(Y \geq k)} = \sum_n n \frac{P(Y \geq k|N = n)P(N = n)}{\sum_n P(Y \geq k|N = n)P(N = n)}, \quad (3)$$

which I have written as a function of observed probabilities. The simulation results are surprisingly close to the data both in the level of sales and in the slope. The level of sales is determined by the proportionality assumption. So for this figure, thinking of firms as a collection of goods of the same size approximates the data pretty well. The slope of the figure is the combination of two effects: the size distribution and $P(Y \geq k|N = n)$. When few balls are thrown it is easy to reach new categories, but once many are dropped it is increasingly hard to do so. This implies that when few balls are thrown the size distribution dominates and we obtain a linear shape. But to reach many different categories is very hard: the probability is small and there are few large firms. The combined effect creates the observed curvature for large k. Overall, it is remarkable how far the proportionality assumption and such a simple allocation mechanism can take us. If we were interested in modeling firm expansion, the results suggest that shuting down firm ability heterogeneity and focusing on firm expansion through discrete units can go a long way in quantitatively explaining many interesting features of entry.

Figure 13 shows the relation between sales and entry by popularity of the market. As we have seen in section 2, it is important to condition on source to rationalize the popularity entry patterns. In using the statistical model, rather than conditioning on source, generating data for all sources and then appropriately rearranging the output from all sources, I instead simulate the results using the bin sizes and number of acquisitions distribution I have used in the previous experiments. Conceptually, it's just as if the simulated data were coming from an "artificial" source. In this way, we can easily identify how the key "parameters" affect the results. To obtain the expected number of balls necessary to observe at least one in destination k, I compute,

$$\begin{aligned} E(N|n_k > 0) = \sum_n nP(N = n|n_k > 0) = \sum_n n\frac{P(N = n, n_k > 0)}{P(n_k > 0)} = \\ \sum_n n\frac{P(n_k > 0|N = n)P(N = n)}{\sum_n P(n_k > 0|N = n)P(N = n)} \end{aligned} \tag{4}$$

Expression 4 can be calculated since all components are known. In particular, $P(n_k > 0|N = n)$ can be obtained from the binomial distribution (n, p_k) as $1-(1-p_k)^n$ with the number of firms that enter market k given by $M(n_k > 0) = \sum_n P(n_k > 0|N = n)M(N = n)$. Again, I use the proportionality assumption to rescale number of acquisitions to global sales: I multiply equation 4 by the average global sales per acquisition.

This figure is the most counterfactual prediction of the model. Two key elements are off: the scale and the slope. Regarding the first, overestimation of sales is large in this figure for two reasons. First, because the probability of each bin is small, many balls have to be drawn so that at least one ball lands in the bin. Contrary to the number of non-empty destinations, where the first balls land in new destinations almost certainly, here many balls have to be drawn from the very beginning. Bin sizes explain why sales in the industry figure are higher than in the country figure. Second, the proportionality assumption shifts vertically the simulated data but doesn't alter its slope. Hence, the proportionality assumption makes the scale of global sales off. By assuming that all markets are the same, the model misses the connection between what happens in a particular market and the implications for the size in the rest of the markets. So to understand how the level of firms' global sales are related to their entry patterns by popularity, the proportionality assumption is not a good approximation and it appears that richer economic heterogeneity interacts in the data. In this dimension, the economic model will improve upon the statistical model by making sales in unpopular destinations smaller.

Regarding the slope, as opposed to Figure 12, here each dot is affected by the complete cross-section of firms. Slopes are constant, suggesting again that the firm size distribution dominates in equation 4. The simulated data has a steeper slope than the observed data inducing a larger discrepancy as we move towards less popular destinations: as p_k becomes smaller, more balls are needed for bin k to be non-empty. In other words, the model delivers too high a slope and hence misses on the size of large relative to small firms. By allowing ball sizes to be heterogenous and positively correlated with popularity of the market the slope could become flatter and make the scale smaller. The variation in sales by destination is a feature that the economic model will capture.

The last outcome of the experiment is Figure 14. The figure is constructed by separately calculating for industries and countries the expected number of non-empty categories given n balls have been thrown, $E(Y|N = n)$. Each dot represents firms with a given number of balls $N = n$ and the size of the dot is the density. The question is, for a given n, how many different industries and countries should the firm be expected to operate? The agreement between the data and the model is remarkable. By imposing that entry into industries and countries is not related, the model behaves quite symmetrically for both industry and country. Also, more fundamentally because the expected number of non-empty categories increases linearly when few balls relative to the K categories are drawn.

Figure 15 shows a counterfactual implication of the statistical model. The figure depicts the probability of observing exactly m non-empty bins, $P(Y = m) = \sum_n P(Y = m|N = n)P(N = n)$, against m. Note that, in the upper tail, the probability of observing a given number of non-empty categories falls sharply. There are two effects producing this result. First, $P(Y = m|N = n)$ falls sharply as n increases: after some balls have been dropped, more balls do not linearly imply more non-empty bins. In Figure 16, this is illustrated by the initial linearity in the expected non-empty categories and a gradual decay in the filling rate of new categories as we draw more balls. Second, few firms draw many balls, so decay in the firm size density cannot compensate for the decay in $P(Y = m|N = n)$.

The statistical model performs well in most dimensions except in capturing some selection patterns related to firm size. Since the model is silent about sales, I assume a proportional relation exists between global sales and the number of acquisitions. With this assumption, we learn that further heterogeneity must be relevant since the model misses behavior of larger firms compared to the small ones. In particular, the most notable counterfactual prediction of the model is in the relative sales of firms that enter the most versus the least popular destinations as well as the scale of sales. In the next section, I develop a model of firm behavior that captures this aspects of entry that are related to firm ability.

4 A multi industry-country entry model

While firms that engage in multi-industry production represent a large part of international sales, this feature has received relatively little attention. In my dataset, 68% of the firms entered just one industry and generated only 10% of total FDI value which implies that the bulk of the volume is performed by multi-industry firms[24]. The evidence section has also shown that firm entry depends on industry and country source, e.g. firms that start in chemicals tend to go to plastics more than they go to automobiles. Hence, the model takes seriously the idea that firms may supply products in rather different industries, e.g. General Electric may supply both consumer electronics and chemical products in some countries, or only one, or none of them in some other countries. Unlike the previous literature, the model will take into account the asymmetry of the product and geographic space and a random component in industry branching.

4.1 Production

There are N countries with each having I industries with differentiated varieties plus one homogenous sector. Hence, there are two broad types of goods, homogenous and differentiated. Producers are introduced in the next subsection. In the differentiated sector, producers are born in a country h and industry i and a realization of potential ideas for goods in all industries all of which have efficiency z. Motivated by the observation that firms enter multiple and remarkably different industries, I model firms as being able to potentially branch out into new industries different from their original or core industry.

[24]Similarly, in the export literature single product firms, at the 10-digit HS code, are 42% but they generate only 0.4% of export value.

Production uses only labor with constant returns to scale; the unit cost is $\frac{w_f \tau_{ijhf}}{z}$ where z is a firm effect, τ_{ijhf}[25] is an iceberg efficiency loss of producing goods in industry j, foreign country f given the firm was born in industry i country h and w_f is the wage rate in a foreign country f. The motivation for an iceberg efficiency loss of producing in a different sector is that carrying over ability to industries that are dissimilar from a firms' core industry is costly. Analogously, it captures the fact that monitoring assets in a distant and culturally different destination is hard[26]. The idea that, by focusing on their core abilities, firms can grow and sustain their competitive edge can be traced back to at least, Penrose (1959), where she argues that firms internal resources determines firm expansion.

As mentioned, each firm is born with a random discrete number of goods determined by a distribution $H(n|z)$ where $n \in 0, 1, 2...\infty$. The mean number of goods that a firm with efficiency z draws is denoted by $\lambda(z)$. A firm has a continuum of goods potentially available in each industry, but it will discover only a discrete number of them. This discrete number of goods is a random variable independent for each firm and is assumed to have a larger arrival rate for more efficient firms $\partial\lambda\partial z > 0$. Papers dealing with multiproduct firms, eg, Bernard, Redding, and Schott (2009a), Bernard, Redding, and Schott (2009b), Mayer, Melitz, and Ottaviano (2009) and Arkolakis and Muendler (2008), assume that firm level ability automatically determines a measure of goods to be produced. However, this is far from true in the data. As suggestive evidence, Figure 8 presents the estimated kernel density of global sales and the number

[25] I assume $\tau_{ijhf} > 1$ and $\tau_{iihh} = 1$. I also impose a triangle inequality to prevent export platform situations, i.e., $\tau_{ijhf} < \tau_{ijhk}\tau_{ijkf}$.

[26] In the most basic proximity-concentration hypothesis, variable trade costs are compared to fixed costs of opening plants to determine whether FDI or exports are used to reach a foreign market. Helpman, Melitz, and Yeaple (2004) extend the idea to include fixed costs of exporting so that larger firms perform FDI, middle ability firms export and the remaining only sell domestically. Implicitly they model FDI as having $\tau = 1$. I model FDI as being subject to variable monitoring costs in the spirit of Head and Ries (2008). Given the assumption on the fixed cost, all results would go through by setting $\tau = 1$.

of acquisitions[27]. For each level of sales, we see that there is substantial variation in the number of acquisitions. This points to the relevance of considering the nondeterministic nature of the number of industries in which a firm sells[28]. The data is saying that firms can be good in one or few goods, like e.g. in the case of a software developer, but not be able to extend that expertise to many other goods, like e.g. GE does. By taking seriously the observation that the number of products has a random component, I am able to study the macroeconomic implications of the mentioned uncertainty at the firm level.

In industry i in home country h the mass of potential producers that can produce their good with efficiency at least z given by,

$$G_{ih}(z) = T_h z^{-\theta} \text{ with } z > 0 \tag{5}$$

where $\theta > 0$ is an inverse measure of heterogeneity.

There is no heterogeneity in the production of the homogeneous good. It is produced under constant returns to scale using only labor, it is freely traded and it's price is normalized at 1.

4.2 Demand, competition and entry

Preferences are given by,

[27]The kernel density is estimated using an optimal bandwidth for estimating the probability density function. The Gaussian kernel can be viewed as a particular case of this method. The optimality is with respect to L_2, mean integrated squared error more precisely. However, for discrete variables it is merely a heuristic and one cannot guarantee optimality.

[28]In related evidence, using export shipment data from Brasil, Arkolakis and Muendler (2008) define the number of products as firm scope and average sale per product as scale and find that the correlation on logs is 0.34. Conditional on their firm ability proxy, this result leaves much unexplained heterogeneity in firms' number of products. Moreover the same positive but not strong correlation was found by Bernard, Redding, and Schott (2009b) for US data.

$$U = q_0^{\mu_0} \prod_{i=1}^{I} \left(\int_{\Omega_i} q_i(\omega)^{\frac{\sigma_i - 1}{\sigma_i}} d\omega \right)^{\frac{\mu_i \sigma_i}{\sigma_i - 1}} \tag{6}$$

where $\mu_0 + \sum_{i=1}^{I} \mu_i = 1$, σ_i is the elasticity of substitution between two varieties in industry i and Ω_i is the set of goods produced in equilibrium in industry i. Given the production assumptions, if the homogenous good is produced in all countries, then wages are equalized to 1 everywhere. I will focus only on this scenario[29].

A producer from (i, h) and efficiency z selling in destination (j, f) faces a CES demand, so he charges a price that is a constant markup over marginal cost, $p_{ijhf}(z) = m_j \frac{w_f \tau_{ijhf}}{z}$ where $m_j = \frac{\sigma_j}{\sigma_j - 1}$. Since the mass of producers is fixed there are profits in equilibrium and they need to be redistributed. I assume there is a mutual fund that divides global profits, π, in proportion to population sizes.

Sales revenue of a firm with productivity z born in (i, h) that sells in (j, f) are,

$$r_{ijhf}(z) = p_{ijhf}(z) q_{ijhf}(z) = \mu_j Y_f \left(\frac{p_{ijhf}(z)}{P_{jf}} \right)^{1-\sigma_j} \tag{7}$$

where P_{jf} is the ideal price index in destination (j, f) and $Y_f \equiv (\frac{\pi}{L} + 1) L_f$ is income in country f, which is composed of labor income and dividends from the profits that firms from all industries and countries accrue.

A firm from industry i country h must incur a fixed cost F_{ijhf} per good that wants to produce in industry j and country f. There is an asymmetric cost of entering industries that depends on the original expertise of the firm. More dissimilar industries will have a larger cost and hence firms will tend to be a collection of related industrial assets. This

[29]Since as I argue in the introduction, the dataset is likely to be representative of the OECD FDI flows, I do not consider this a drawback.

implies that even if all firms have ideas for goods in all other industries only the most productive firms will overcome the fixed cost and actually produce some of those goods. In this respect, it is similar to the innovation process that is the basis of the Frechet distribution in Eaton and Kortum (2002), where most ideas are not used because they are of too low quality to belong to the technology frontier. Here the technology frontier of the firm is determined by it's efficiency z, the monitoring costs τ and the cost of entering a given industry F. Net of the entry cost, a producer from (i, h) and efficiency z makes profits $\pi_{ijhf}(z) = \frac{\mu_j}{\sigma_j}(m_j \tau_{ijhf})^{1-\sigma_j} L_f(\frac{\pi}{L}+1) P_{jf}^{\sigma_j - 1} z^{\sigma_j - 1} - F_{ijhf}$ in destination (j, f). Hence, the cutoff efficiency level for producers from (i, h) to produce at (j, f) is,

$$\underline{z}_{ijhf} = P_{jf}^{-1} m_j \tau_{ijhf} \left(\frac{F_{ijhf} \sigma_j}{\mu_j Y_f} \right)^{\frac{1}{\sigma - 1}} \tag{8}$$

4.3 Equilibrium price index

Entry into a given destination needs to take into account all goods produced by entrants from all industries and countries that overcome the fixed cost of entry. Also, since there is a continuum of goods and efficiency is drawn independently for each good, we can use the LLN to write the price index as,

$$P_{jf}^{1-\sigma} = \sum_{h=1}^{N} \sum_{i=1}^{I} \sum_{n_z=0}^{\infty} \int_{\underline{z}_{ijhf}}^{\infty} n_z p(z)^{1-\sigma} H(n|z) dG_h(z) = \tag{9}$$

$$\sum_{h=1}^{N} \sum_{i=1}^{I} \int_{\underline{z}_{ijhf}}^{\infty} \lambda(z) p(z)^{1-\sigma} dG_h(z) \tag{10}$$

where $H(n|z)$ is the distribution of the number of goods conditional on efficiency and $\lambda(z)$ is the mean of the distribution. I assume $\lambda(z) = z^{\beta}$ and using equation 8 I obtain,

$$P_{jf}^{\beta-\theta} = c_2(j) \left(\left(\frac{\pi}{L} + 1 \right) L_f \right)^{\frac{\theta-\beta-\sigma_j+1}{\sigma_j-1}} \Psi_{jf} \tag{11}$$

where $\Psi_{jf} = \sum_h \sum_i T_h (m\tau_{ijhf})^{\beta-\theta} \left(\frac{F_{ijhf}\sigma}{\mu_j} \right)^{\frac{\sigma_j-1+\beta-\theta}{\sigma_j-1}}$ and $c_2(j) \equiv \frac{\theta}{1+\theta-\sigma_j-\beta}$. Convergence of the integral requires $\sigma_j - 1 + \beta - \theta < 0$. Ψ_{jf} generalizes the "multilateral resistance" term introduced in Anderson and Van Wincoop (2004) to include entry from all other industries. In this case, it is a measure of the remoteness of industry j country f that takes into account variable and fixed costs of entering the location as well as a firm size distribution effect.

Similarly, profits can be written as,

$$\pi = \sum_{h,f=1}^{N} \sum_{i,j=1}^{I} \sum_{n_z=0}^{\infty} \int_{\underline{z}_{ijhf}}^{\infty} n_z \pi_{ijhf}(z) H(n|z) dG_h(z) = \tag{12}$$

$$\sum_{h,f=1}^{N} \sum_{i,j=1}^{I} \int_{\underline{z}_{ijhf}}^{\infty} \lambda(z) \pi_{ijhf}(z) dG_h(z) \tag{13}$$

which using equation 11 and 8 become,

$$\pi = \frac{\sum_j \frac{\mu_j(\sigma_j-1)}{\sigma_j(\theta-\beta)}}{1 - \sum_j \frac{\mu_j(\sigma_j-1)}{\sigma_j(\theta-\beta)}} L = c_1 L \tag{14}$$

so income is $Y_f = \left(\frac{\pi}{L} + 1 \right) L_f = (1 + c_1) L_f$ and the price index,

$$P_{jf}^{\beta-\theta} = c_2(j)((1+c_1)L_f)^{\frac{\theta-\beta-\sigma_j+1}{\sigma_j-1}}\Psi_{jf} \tag{15}$$

Using 14 and 15, the cutoff for firms from (i,h) in (j,f) can be expressed as,

$$\underline{z}_{ijhf} = \left(c_2(j)\frac{(1+c_1)L_f}{\Psi_{jf}}\right)^{\frac{-1}{\theta-\beta}} \left(\frac{F_{ijhf}\sigma_j}{\mu_j}\right)^{\frac{1}{\sigma_j-1}} m_j\tau_{ijhf} \tag{16}$$

Finally, firms from (i,h) with ability $z \geq \underline{z}_{ijhf}$ have revenues in (j,f) given by,

$$r_{ijhf}(z) = \mu_j((1+c_1)L_f)^{\frac{\sigma_j-1}{\theta-\beta}} \left(\frac{m_j\tau_{ijhf}}{z}\right)^{1-\sigma_j} \Psi_{jf}^{\frac{\sigma_j-1}{\beta-\theta}} \tag{17}$$

4.4 Aggregates

In this section I introduce the aggregate implications of the model. Aggregate sales of affiliates from (i,h) selling in (j,f) are obtained by integrating over the mass of producers that overcome the cutoff ability level,

$$r_{ijhf} = \int_{\underline{z}_{ijhf}}^{\infty} \lambda(z) r_{ijhf}(z)\, dG_h(z) = (1+c_1)T_hL_f\mu_j(m_j\tau_{ijhf})^{\beta-\theta}\Psi_{jf}^{-1}\left(\frac{F_{ijhf}\sigma_j}{\mu_j}\right)^{1+\frac{\beta-\theta}{\sigma_j-1}} \tag{18}$$

Equation 21 is an industry-generalized gravity equation where bilateral sales depend one to one on both T_h and L_f, measures of size of countries[30], and both the variable and

[30] In fact, if T_i is proportional to L_i, as Alvarez and Lucas (2005) and others assume, then we have a symmetric effect of country sizes on bilateral sales. Given that total income in i is in equilibrium proportional to L_i we could then also substitute T_hL_f by Y_hY_f which makes the expression even more

the fixed cost of entry affect aggregate sales negatively but with different exponents. The next two comments elaborate on the latter observation.

The effect of τ on sales depends on θ, the firm heterogeneity parameter, and on β, the exponent on the arrival rate of goods but it doesn't depend on the elasticity of substitution. The fact that firms branch into multiple industries affects the result that sales depend only on θ that Chaney (2008) found. On the other hand, the effect of entry costs, F, on sales depends only on θ, β and σ. Again, firms branching into multiple industries also affects the equivalent result in Chaney (2008), since he found sales to depend only on θ and σ. The next proposition analyzes the effect on sales of changes in iceberg costs:

Proposition 1 *The elasticity of sales with respect to iceberg costs is* $\frac{d \ln r_{ijhf}}{d \ln \tau_{ijhf}} = \theta - \beta$.

Proof. First, I decompose the elasticity of sales with respect to entry costs into the corresponding intensive and extensive elasticities,

$$-\frac{d \ln r_{ijhf}}{d \ln \tau_{ijhf}} = -\frac{\tau_{ijhf}}{r_{ijhf}} \int_{\underline{z}_{ijhf}}^{\infty} \lambda(z) \frac{\partial r_{ijhf}(z)}{\partial \tau_{ijhf}} \, dG_h(z) + \frac{\tau_{ijhf}}{r_{ijhf}} \lambda(\underline{z}_{ijhf}) r_{ijhf}(\underline{z}_{ijhf}) G_h'(\underline{z}_{ijhf}) \frac{\partial \underline{z}_{ijhf}}{\partial \tau_{ijhf}} \tag{19}$$

First, assume that source (i, h) is small enough so that $\frac{\partial \Psi_{jf}}{\partial \tau_{ijhf}} \approx 0$. Then, using equation 17 I can write $\frac{\partial r_{ijhf}(z)}{\partial \tau_{ijhf}} = (1 - \sigma_j) \frac{r_{ijhf}(z)}{\tau_{ijhf}}$. The intensive margin elasticity becomes,

$$-\frac{\tau_{ijhf}}{r_{ijhf}} \int_{\underline{z}_{ijhf}}^{\infty} \lambda(z)(1 - \sigma_j) \frac{r_{ijhf}(z)}{\tau_{ijhf}} dG_h(z) = (\sigma_j - 1) \tag{20}$$

resemblant to a gravity equation.

I now turn to the extensive margin. First, rewrite sales as $r_{ijhf}(z) = A_{ijhf} z^{\sigma_j - 1}$ to obtain total sales as,

$$r_{ijhf} = \int_{\underline{z}_{ijhf}}^{\infty} \lambda(z) A_{ijhf} z^{\sigma_j - 1} dG_h(z) = T_h \theta \int_{\underline{z}_{ijhf}}^{\infty} A_{ijhf} z^{\beta + \sigma_j - \theta - 2}\, dz$$

$$= \frac{\theta}{1 + \theta - \beta - \sigma_j} T_h A_{ijhf} \underline{z}_{ijhf}^{\beta + \sigma_j - 1 - \theta} = \frac{1}{1 + \theta - \beta - \sigma_j} T_h r(\underline{z}_{ijhf}) \lambda(\underline{z}_{ijhf}) G'_h(\underline{z}_{ijhf}) \quad (21)$$

Then, using equation 16 I rewrite $\frac{\partial \underline{z}_{ijhf}}{\partial \tau_{ijhf}} = \frac{\underline{z}_{ijhf}}{\tau_{ijhf}}$ and obtain elasticity of the extensive margin as,

$$\frac{\tau_{ijhf}}{r_{ijhf}} \lambda(\underline{z}_{ijhf}) r_{ijhf}(\underline{z}_{ijhf}) G'_h(\underline{z}_{ijhf}) \frac{\underline{z}_{ijhf}}{\tau_{ijhf}} = 1 + \theta - \beta - \sigma_j \quad (22)$$

Hence by adding the intensive and extensive margin elasticities I obtain the total elasticity of sales with respect to variable costs $\sigma_j - 1 + 1 + \theta - \beta - \sigma_j = \theta - \beta$. ■

The intuition for the result is the following. In the case of the intensive margin, the more substitutable goods are, σ high, the more sales will be affected by a change in τ. For the extensive margin, when σ is low, products are not very subtitutable and firms with low productivity are still able to capture a large market share and, hence, there is a large increase in sales. Overall, the effect of σ on both margins is opposite and the extensive margin dominates.

The next proposition deals with the effects on sales of changes in fixed cost entry barriers:

Proposition 2 *The elasticity of sales with respect to entry costs is* $\frac{d \ln r_{ijhf}}{d \ln F_{ijhf}} = \frac{\theta - \beta}{\sigma_j - 1} - 1$.

Proof. The elasticity of sales with respect to entry costs can be decomposed into the corresponding intensive and extensive elasticities,

$$-\frac{d\ln r_{ijhf}}{d\ln F_{ijhf}} = -\frac{F_{ijhf}}{r_{ijhf}}\int_{\underline{z}_{ijhf}}^{\infty}\lambda(z)\frac{\partial r_{ijhf}(z)}{\partial F_{ijhf}}\,dG_h(z) + \frac{F_{ijhf}}{r_{ijhf}}\lambda(\underline{z}_{ijhf})r_{ijhf}(\underline{z}_{ijhf})G_h'(\underline{z}_{ijhf})\frac{\partial \underline{z}_{ijhf}}{\partial F_{ijhf}} \tag{23}$$

Since $\frac{\partial r_{ijhf}(z)}{\partial F_{ijhf}} = 0$, the elasticity of the intensive margin with respect to the fixed cost is zero. Regarding the extensive margin, first note that $\frac{\partial \underline{z}_{ijhf}}{\partial F_{ijhf}} = \frac{1}{\sigma_j - 1}\frac{\underline{z}_{ijhf}}{F_{ijhf}}$. Then apply the same steps as for the intensive margin to obtain $\frac{1+\theta-\beta-\sigma_j}{\sigma_j-1} = \frac{\theta-\beta}{\sigma_j-1} - 1$. Overall, the elasticity of sales of affiliates from (h,i) selling in (j,f) with respect to F_{ijhf} is $\frac{\theta-\beta}{\sigma_j-1} - 1$.

■

In this case, note that the intensive margin doesn't depend on the fixed cost. A change in the fixed cost induces a lowering of the productivity threshold and more firms to enter and again, if σ is low, their sales are a large part of total sales. the extensive margin matters

I now turn to the study of other aggregates, like the number of affiliate plants from (i,h) selling in (j,f):

$$N_{ijhf} = \int_{\underline{z}_{ijhf}}^{\infty}\lambda(z)dG_h(z) = c_2(j)(1+c_1)T_hL_f(m_j\tau_{ijhf})^{\beta-\theta}\Psi_{jf}^{-1}\left(\frac{F_{ijhf}\sigma_j}{\mu_j}\right)^{\frac{\beta-\theta}{\sigma_j-1}} \tag{24}$$

This equation again looks like an industry-generalized gravity equation and all exponents are the same as in the equation 21 except for the one in the fixed cost. More precisely, the elasticity of sales with respect to the fixed cost is larger than the elasticity of the number of plants with respect to the fixed cost. The reason is that in the case of sales, the elasticity is the sum of the intensive and extensive margins, while the number of plants is only the extensive margin that matters. While equation 24 is directly observed in my data, equation 21 is not in my data sample.

Several models in international economics, including Eaton and Kortum (2002), Chaney (2008) and Ramondo (2006), deliver a multinomial model for aggregate fractions of bilateral sales. This model implies a multinomial structure that adds the missing industry dimension. Mathematically, sales from source (i, h) in destination (j, f) over total sales in (j, f),

$$\frac{r_{ijhf}}{r_{jf}} = \frac{T_h \tau_{ijhf}^{\beta-\theta} F_{ijhf}^{1+\frac{\beta-\theta}{\sigma_j-1}}}{\sum_h \sum_i T_h \tau_{ijhf}^{\beta-\theta} F_{ijhf}^{1+\frac{\beta-\theta}{\sigma_j-1}}} \tag{25}$$

Keeping barriers constant, a larger country will sell more in all sectors of a foreign destination. The same differences arise as comparing equation 21 and equation 24: both variable and fixed barriers impede multinational production but are affected by model parameters differently. While this fraction is not directly observed in my data, I do observe deal values for each acquisition. I also observe the fraction of plants from source (i, h) in destination (j, f) over the total plants in (j, f):

$$\frac{N_{ijhf}}{N_{jf}} = \frac{T_h \tau_{ijhf}^{\beta-\theta} F_{ijhf}^{\frac{\beta-\theta}{\sigma_j-1}}}{\sum_h \sum_i T_h \tau_{ijhf}^{\beta-\theta} F_{ijhf}^{\frac{\beta-\theta}{\sigma_j-1}}} \tag{26}$$

which has a nearly identical structure as equation 25 except for the exponent on the fixed cost. Since the number of acquisitions is in my sample, this equation can also be used for estimation purposes.

4.5 Empirical Analysis

[WORK IN PROGRESS]

5 Conclusion

Using firm level data, I provide new evidence on multinational production. The patterns that emerge have several features which to my best knowledge are new. First, there is a remarkable similarity to exporter behavior. Second, a robust similarity between industry and country entry emerges. Third, there is a role for firm heterogeneity, where a few firms enter several countries and industries while most enter just a few. Those that are larger as measured by global sales, are entering more distant countries and more different industries, but most firms enter just a few, more proximate destinations.

A statistical model of balls falling into bins provides insights into the economic content of the data. The balls and bins model is a useful tool for two reasons. First, when the model misses a fact we can learn about the economics that are driving the results and the properties that the economic model cannot have. Second, when the model matches a fact, it informs us about the properties that a successful random growth model of firm needs to have, thus opening a new research path. Let me elaborate on this. On the one hand, this model belongs to a tradition in economics that uses mechanical models to understand data regularities[31]. It takes the distribution of firm size as given and makes predictions about the allocation of goods into industry-country categories. On the other hand, there is a large literature exploring what mechanics of firm growth can deliver fat tailed distributions for firm sizes, see e.g. Simon (1955) and Klette and Kortum (2004)[32]. The success of the balls and bins model to explain entry frequencies suggests that both statistical mechanics could be merged into a stylized model of firm growth through multiproduct and multinational operations. Since most of aggregate trade and FDI is explained by the extensive margin rather than the intensive margin, such a model could explain the firm size distribution, international

[31]For example, see Ellison and Glaeser (1997) and Head and Ries (2005).

[32]For the case of cities, see Gabaix (1999).

and industrial branching and aggregate bilateral volumes.

The economic model has two novel features. First, it provides a unified framework to understand firm expansion into countries and into industries and allows asymmetric bilateral industry-country entry costs. In this it allows to consider the effects of the product space as well as geographic and cultural distances on firm ownership patterns. In particular, it crystalizes in a new way the core competencies idea, i.e., that firms cluster similar goods within their boundaries. Second, if one is to think about industry expansion, it is necessary to consider how intensively firms replicate their know-how. In that sense, I depart from the previous literature by breaking the mathematical connection between the firm ability and the number of products it produces. In particular, I allow a random number of goods per firm and study the macroeconomic implications. An industry-country gravity equation for aggregate bilateral sales is obtained and the evidence can be closely matched.

6 Appendix

6.1 Figures

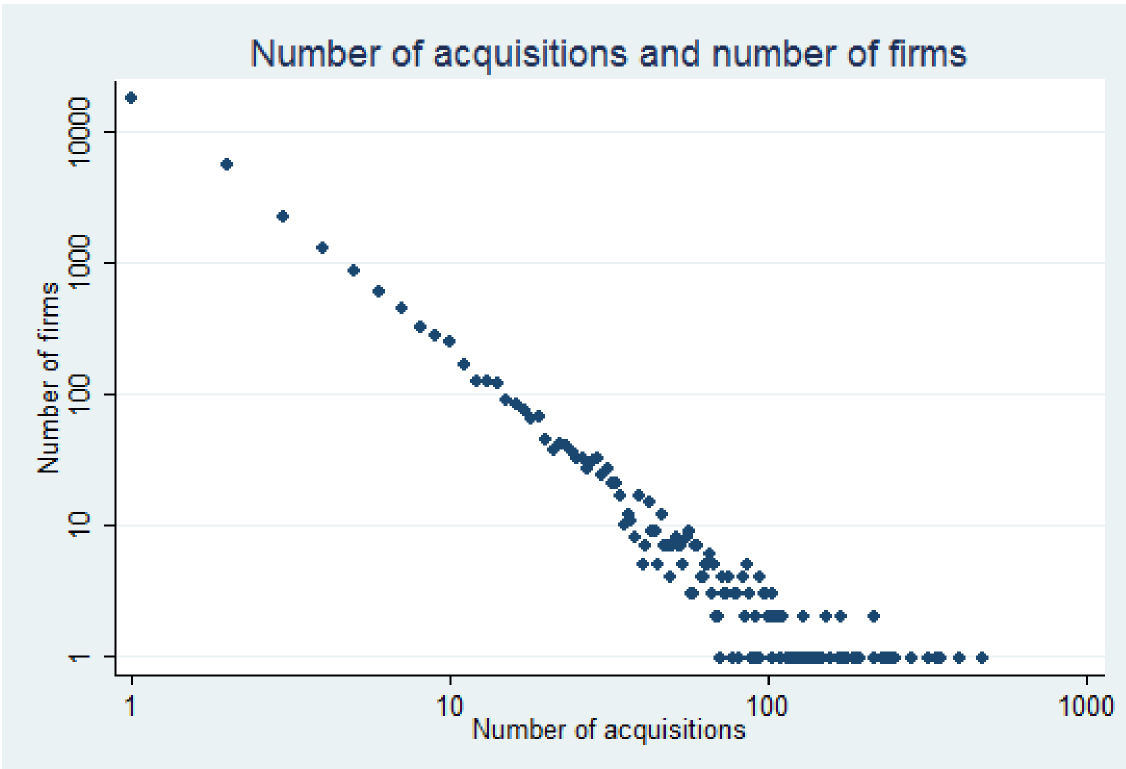

Figure 1: Number of companies with a given number of acquisitions. Note the Pareto shape, indicating great dispersion in the acquisition intensity in the cross-section of firms. The figure suggests that firm ability is transmitted to acquisition intensity.

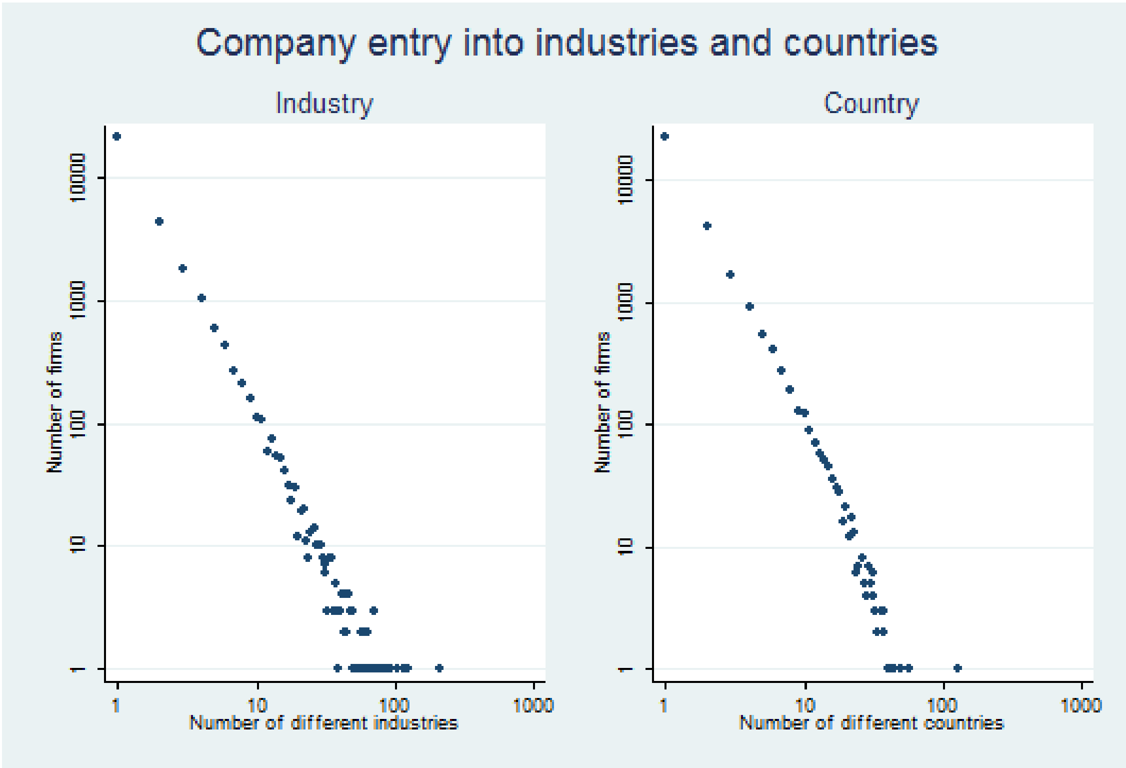

Figure 2: Number of companies that enter a given number of countries or industries, in log scale. The distribution is Pareto for both countries and industries. Overall, the figures suggest that the underlying firm size distribution affects industry and country entry similarly.

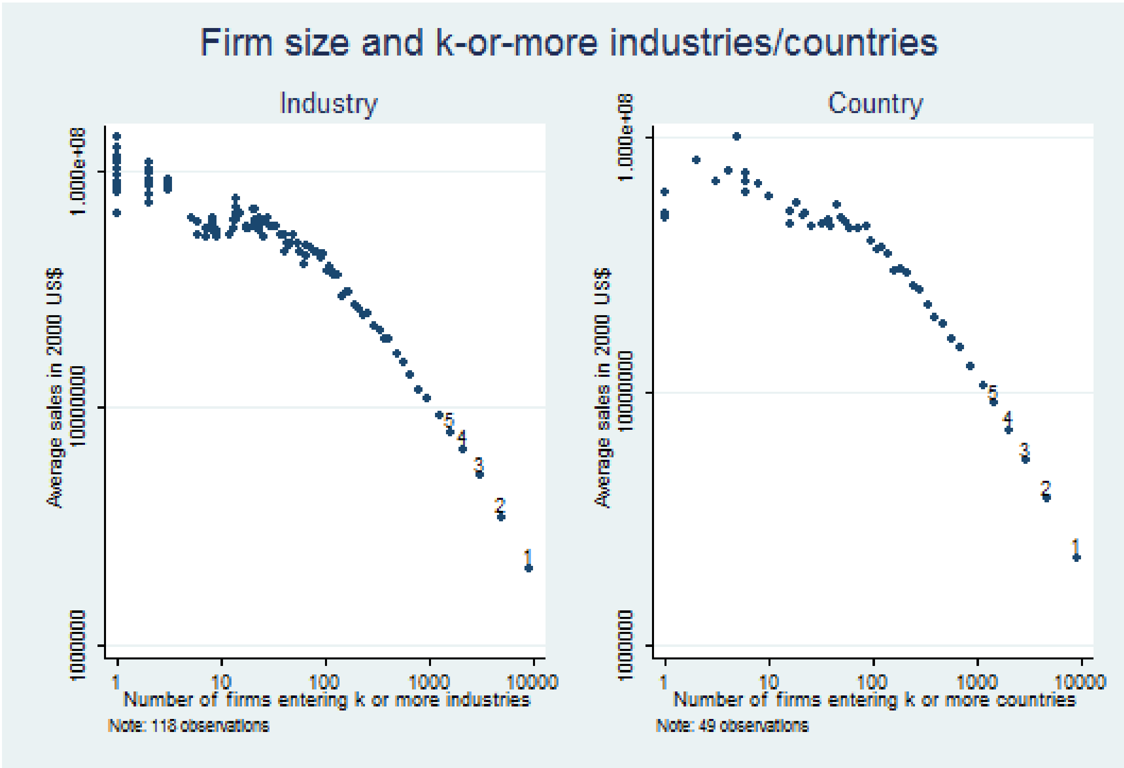

Figure 3: K-or-more country/industry destinations entered and average global sales of entrants, in log scale. The numbers over the dots indicate the number in "k-or-more markets" for k=1,2,...,5. For firms that enter k-or-more industries, I compute their average global sales. Country figure is analogous. Although the figure suggests that firms that enter many destinations have a larger firm effect, there is a mechanical relation between extra markets and larger global sales.

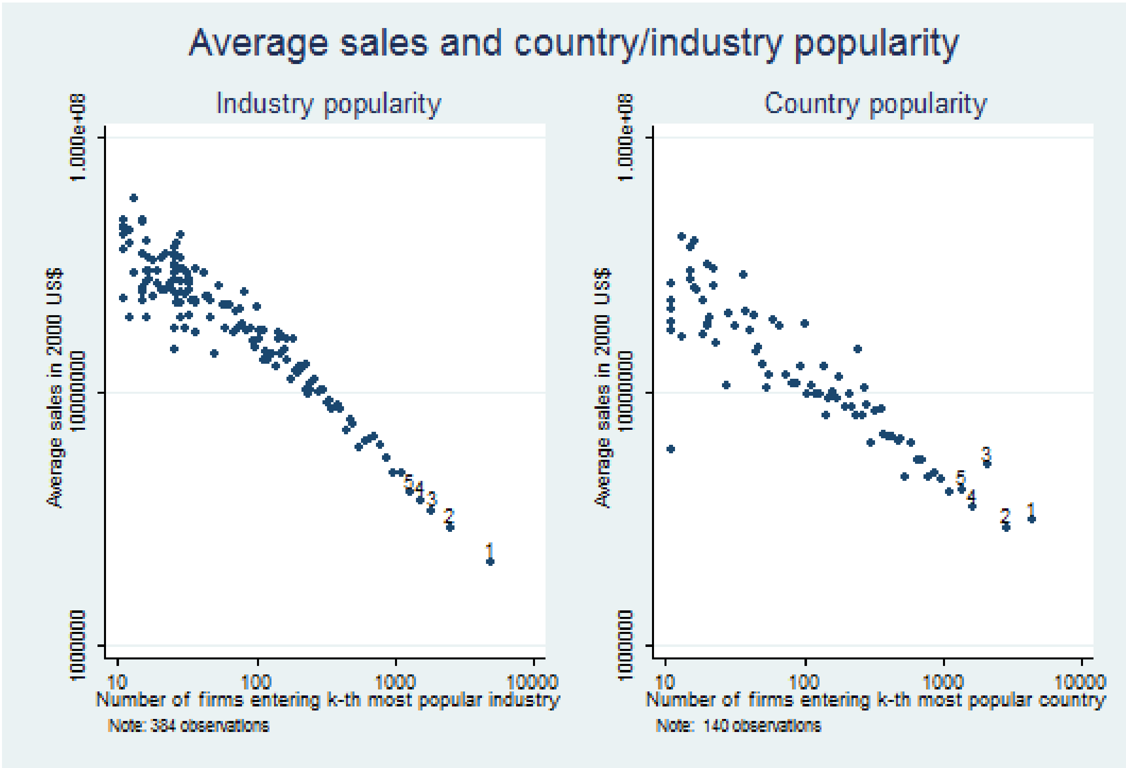

Figure 4: Country/industry popularity and average global sales of entrants, in log scale. The numbers over the dots indicate the number in "k-th" most popular market for k=1,2,...,5. I compute for each country the ranking of the most popular destinations. For each number in that ranking, for all source countries, I compute the average sales of firms that enter each ranked destination. Industry figure is analogous. Industry and country popularity is negatively associated with firm ability, ie, companies that enter unpopular countries/industries are larger as measured by sales. No mechanical relation is present in this figure.

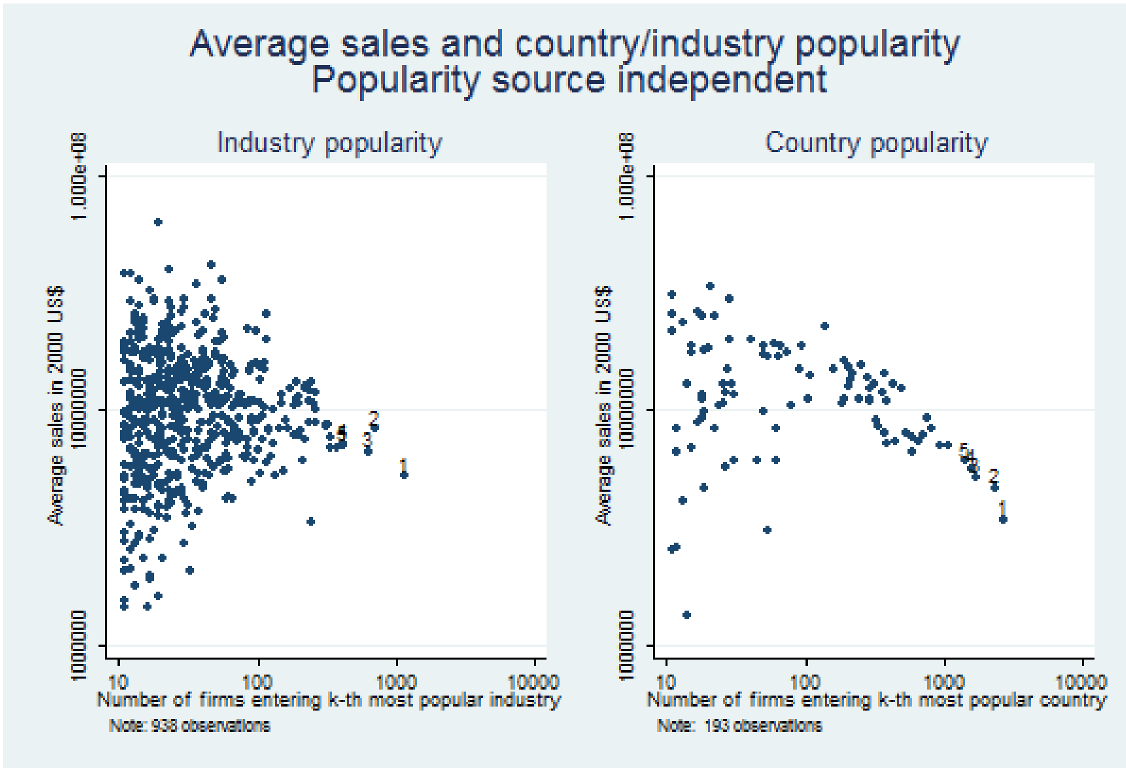

Figure 5: Country/industry popularity and average global sales of entrants, in log scale. The numbers over the dots indicate the number in "k-th" most popular market for k=1,2,...,5. In this figure, I do not construct a ranking for each source but rather I create a ranking independent of the originating firm information. The neat relationship found by conditioning on source disappears here suggesting that the implicit bilateral distances are necessary to rationalize the data.

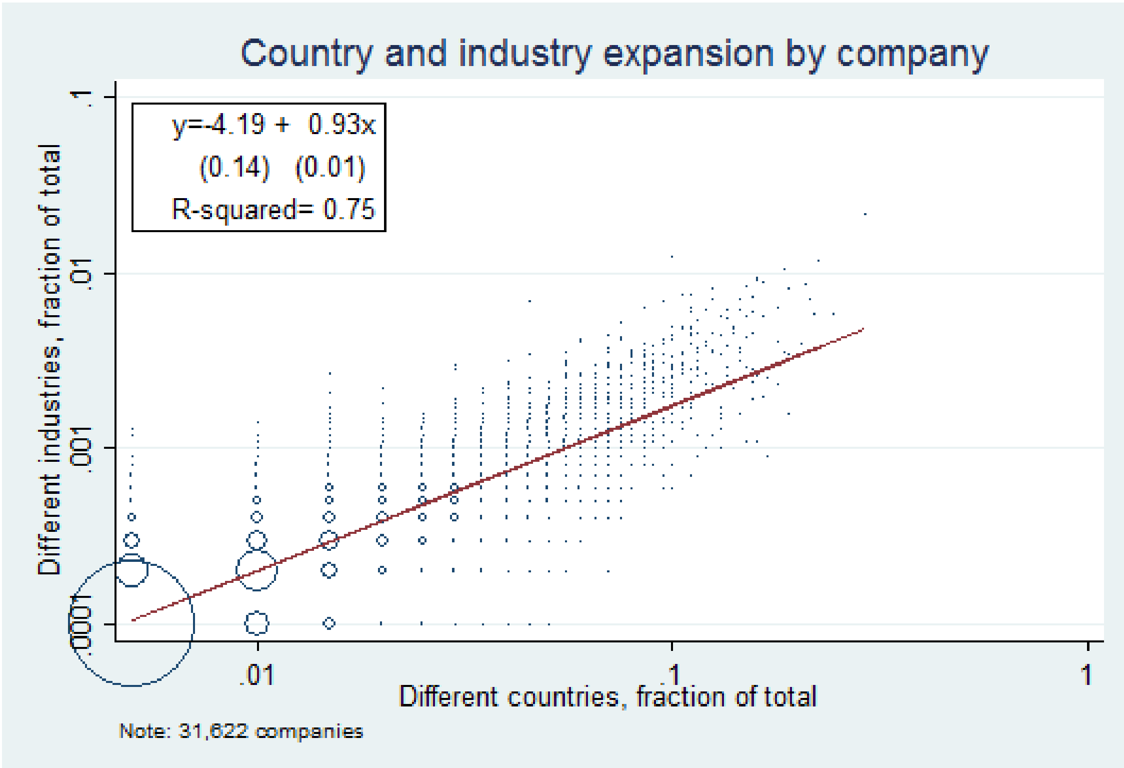

Figure 6: Number of industries and countries entered by a given company; dot sizes are proportional to number of firms. A strong positive correlation appears: firms that enter many countries also enter many industries.

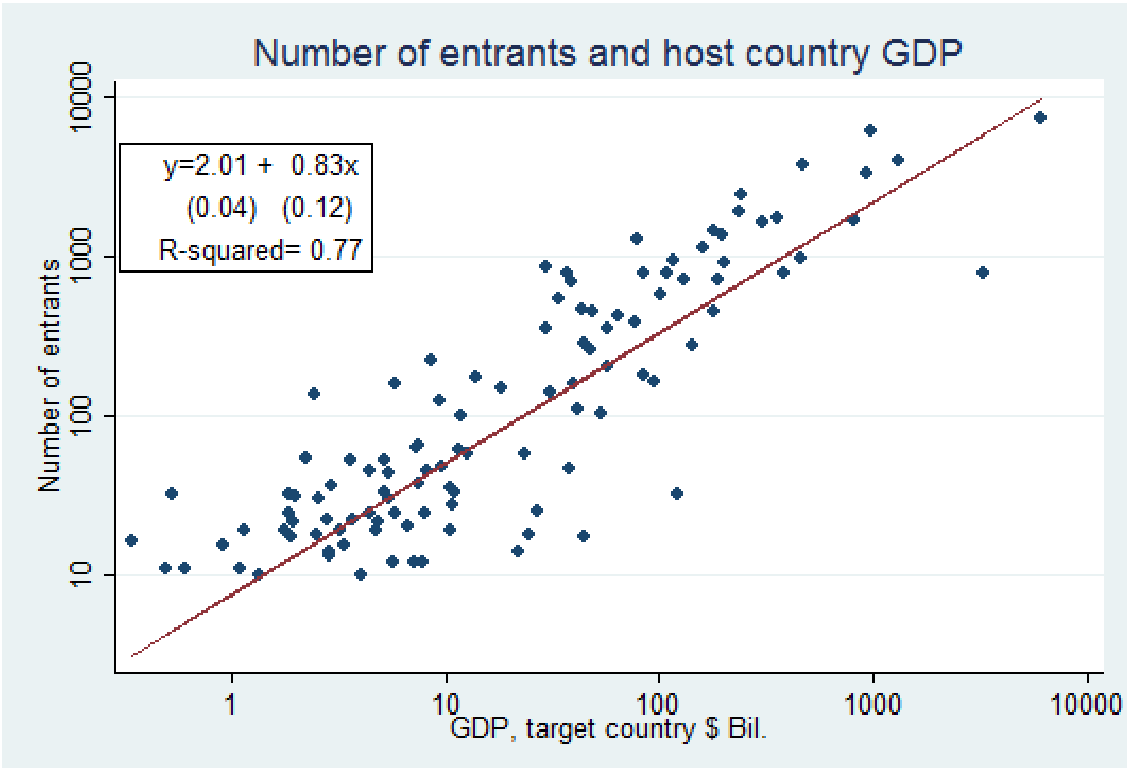

Figure 7: Number of entrants and destination country GDP. The number of entrants rises proportionately with destination GDP with a slope slightly smaller than 1.

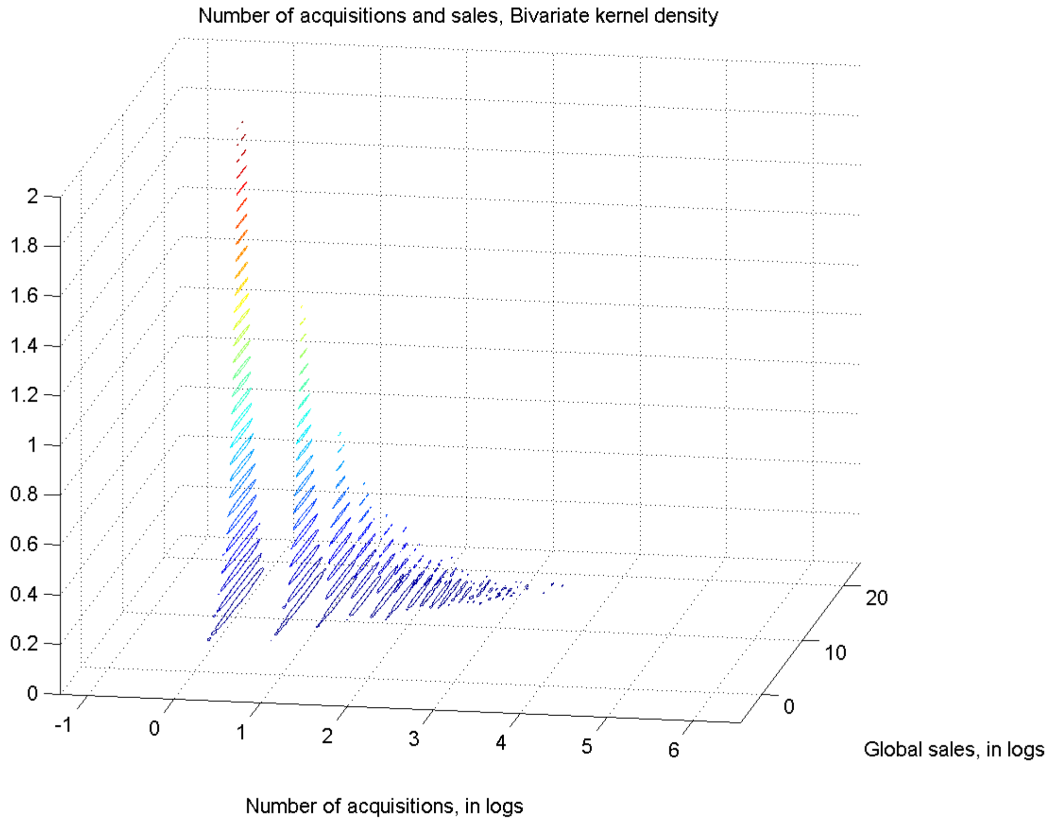

Figure 8: Bivariate kernel density of global sales and number of acquisitions. At each level of global sales we observe a distribution for the number of acquisitions suggesting that there is substantial heterogeneity in the connection between the two.

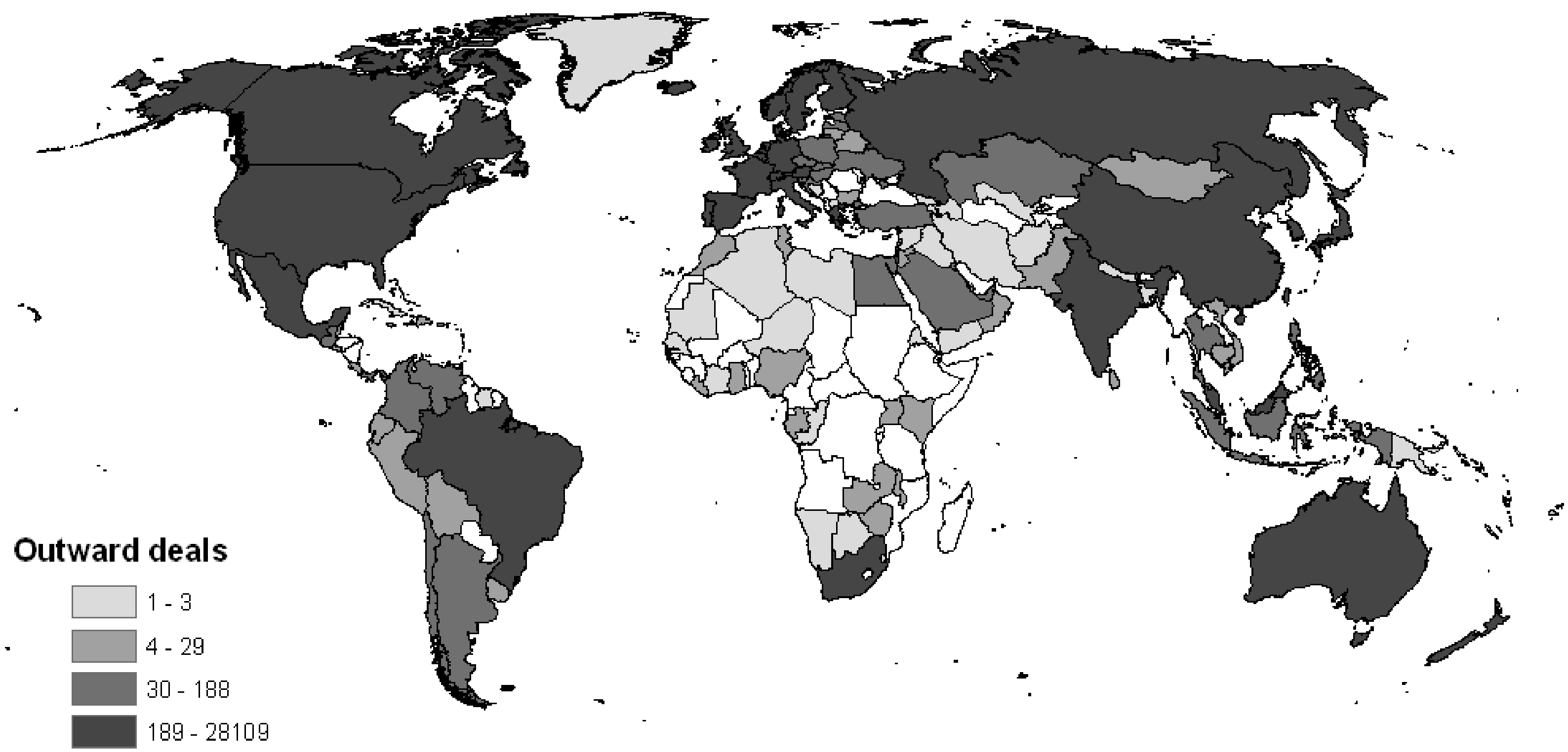

Figure 9: Outward deals originated in each country by quartile.

6.2 Simulation figures

6.2.1 Balls and bins

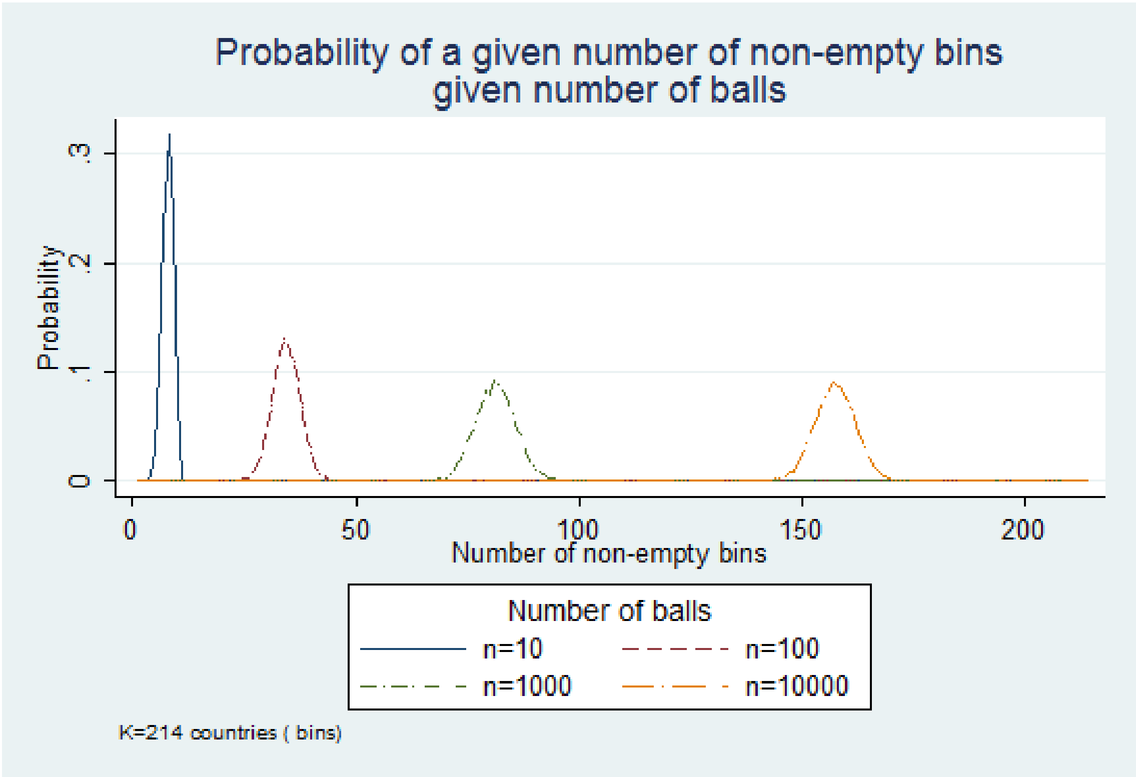

Figure 10: Simulated probability of given number of non-empty bins conditional on n balls.

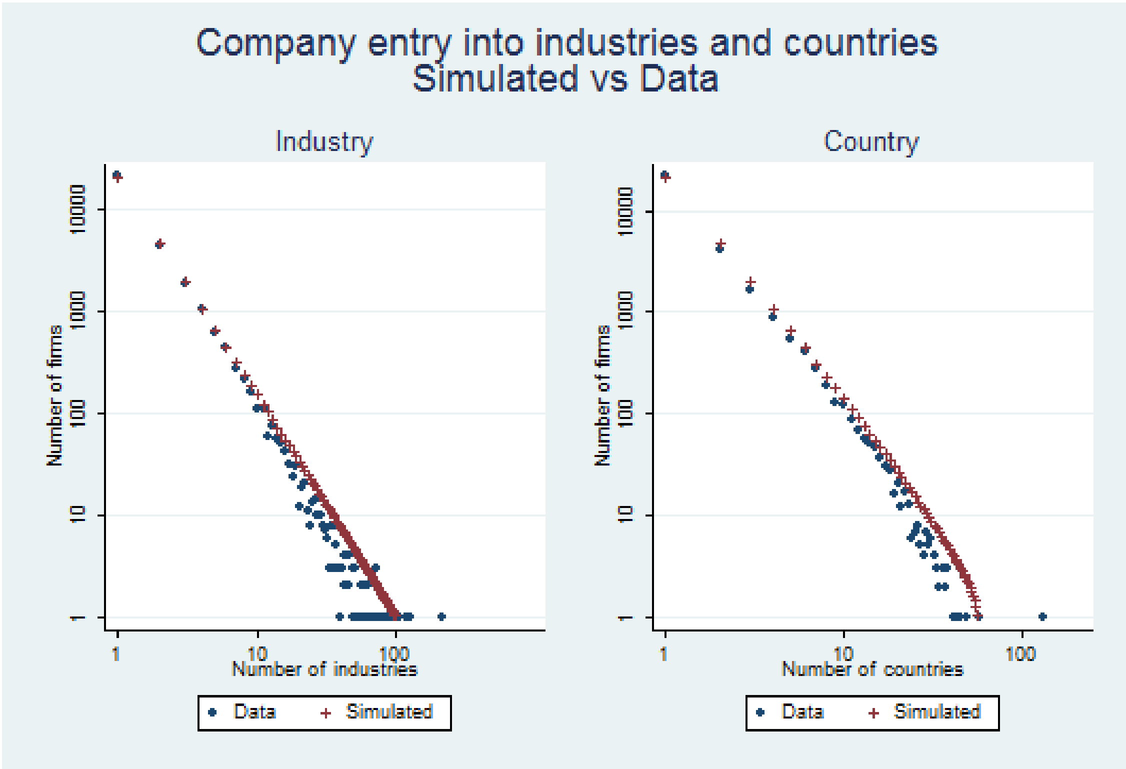

Figure 11: Model predictions and data on the number of firms that enter exactly k different industries or countries. Note the tight prediction of the model.

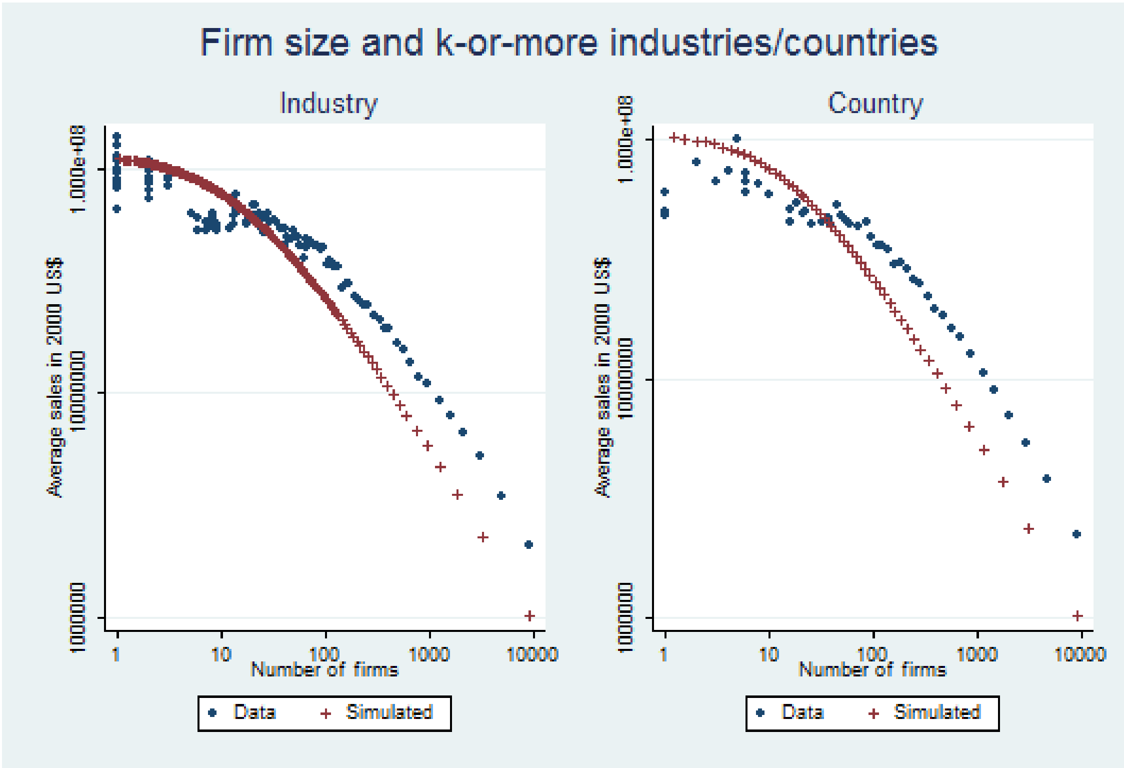

Figure 12: Model predictions and data on the number of firms that enter k-or-more different industries or countries. Note the scale is well predicted. This indicates that assuming firms to be a random collection of goods all of which are of the same size is remarkably close to fitting the data. Also note the similar shape which is induced by the number of acquisitions distribution and the decreasing probability of reaching many different bins. The reason for the curvature in the left tail is that non-empty categories are easily filled when a few balls are thrown but it becomes increasingly harder to fill new categories plus few firms throw many balls.

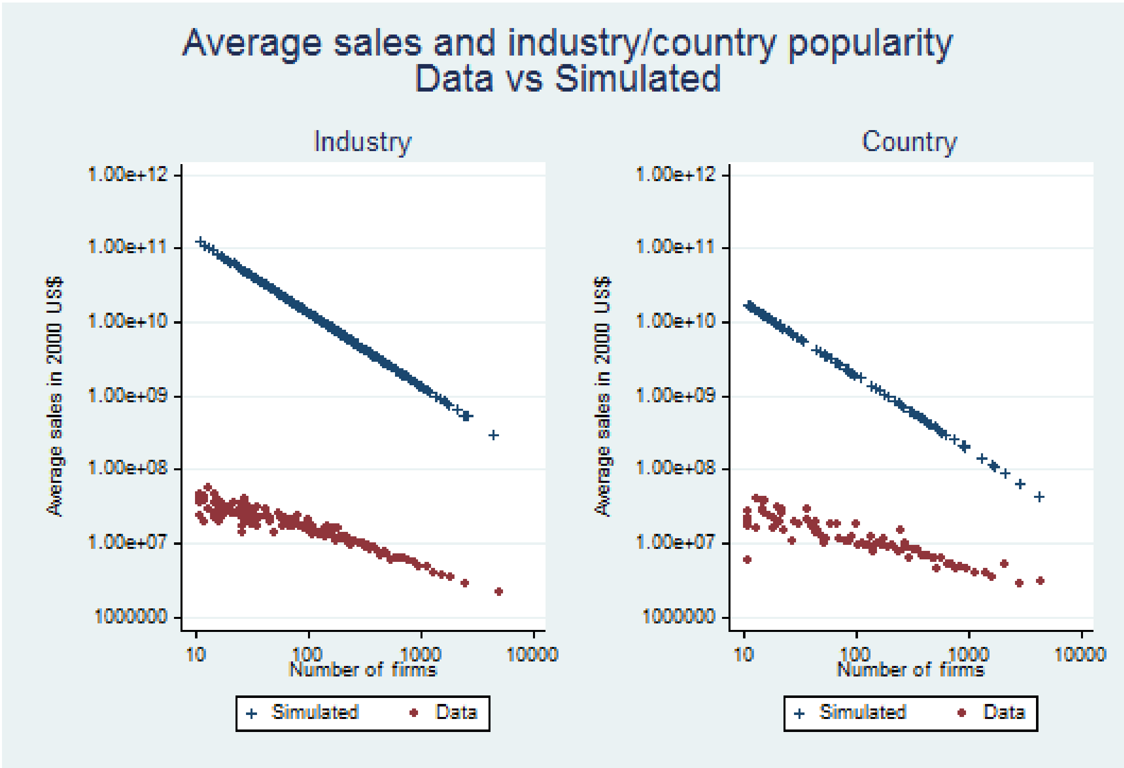

Figure 13: Model predictions and data on the number of firms that enter the k-th most popular industry or country. The large overprediction of sales happens for two reasons: the size of bins and the rescaling factor. The former explains why the country figure has a lower overprediction: bins are larger. The latter also induces shifts in the vertical axis but of the same magnitude on both figures. For this particular entry pattern, assuming that firms are a bundle of homogenous sized goods is not close to the data. Also note the linear shape inherited from the number of acquisitions distribution. The simulated data has a slope that is larger than the data, suggesting that the model misses the relative sizes of entrants into unpopular destinations with respect to those that enter easy destinations.

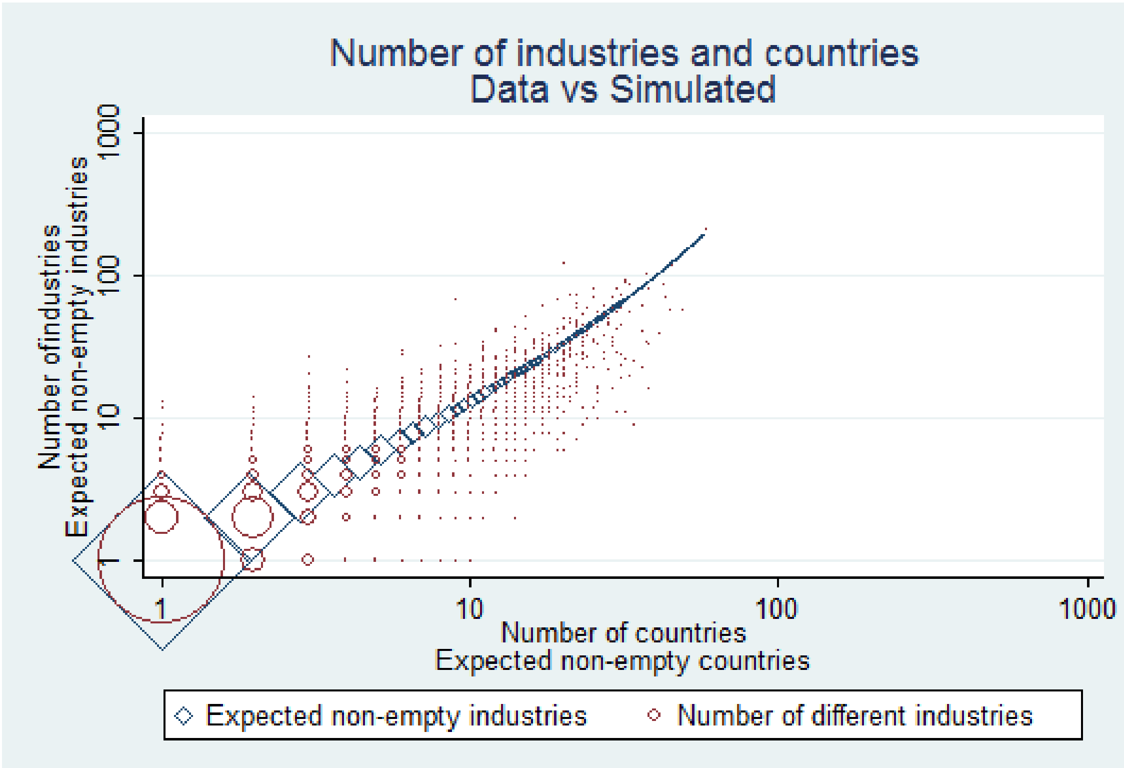

Figure 14: Number of different industries and number of different countries at the firm level. The model predictions are computed by doing two independent experiments. The simulation results and data on firm expansion are remarkably close. Firms that expand into new industries also expand into new countries.

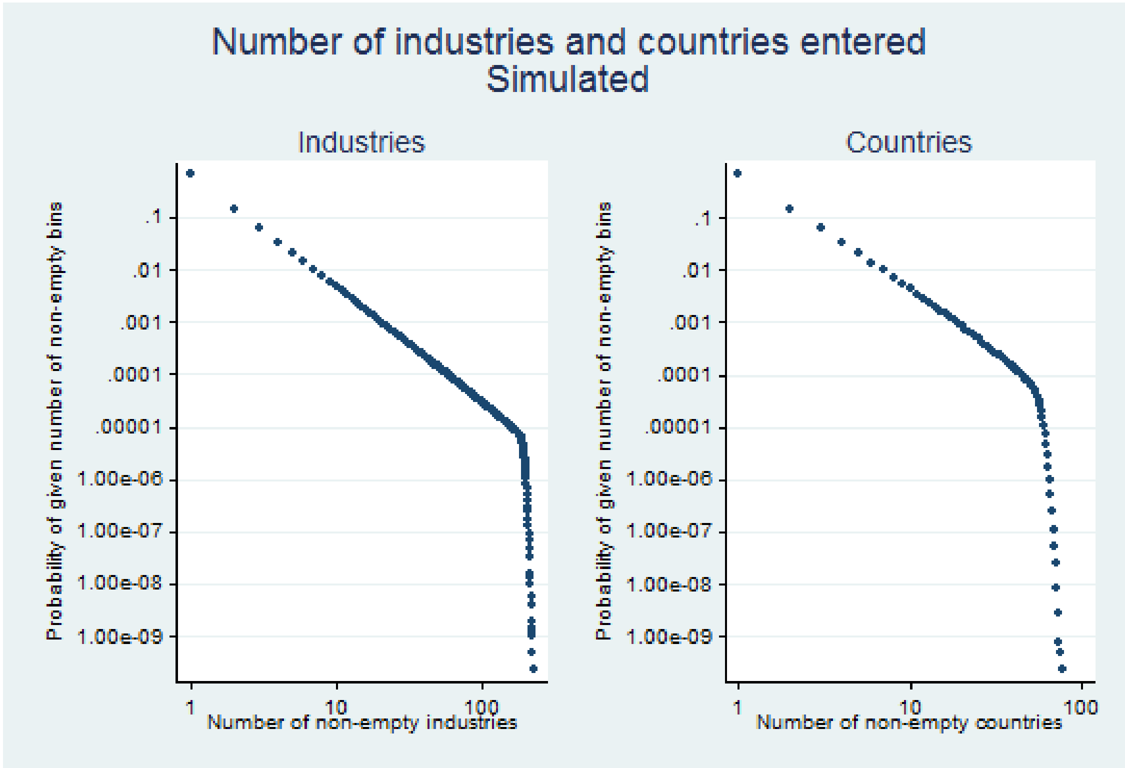

Figure 15: Probability of observing exactly k non-empty destinations. The figure is constructed using the observed deal number distribution and shows how on the upper tail the probability decays extremely fast. Two effects produce it; first, to reach new destinations becomes increasingly harder as the number of balls increases and second, few firms draw many balls.

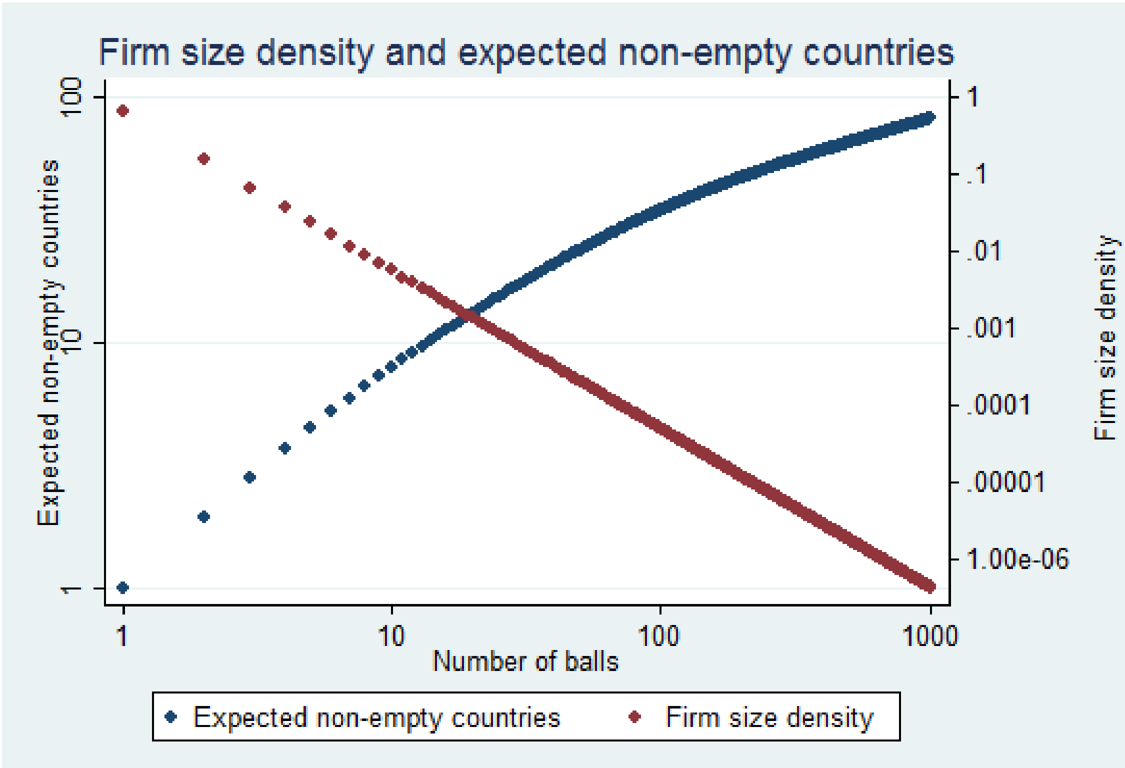

Figure 16: This figure illustrates the main failure of the balls and bins model. Because the firm size density falls linearly on logs with but the number of different destinations is less than linear on logs, the combined effect is a dramatic decrease in the predicted probability of high number of different destinations.

6.3 Tables

Deal data items
Deal id.
Deal value
Date of deal
Percentages acquired and owned before deal
Names
SEDOL, Datastream code
Country
Up to 5 4-digit SIC industries

Table 5: SDC Platinum: Acquiror, acquiror UP and target data items

Acquiror country	Target country	Rank number	Number of firms
United States	United Kingdom	1	1179
United States	Canada	2	1085
United States	Germany	3	643
United States	France	4	509
United States	Australia	5	400
United States	Netherlands	6	281
United States	Italy	7	270
Japan	United States	1	382
Japan	United Kingdom	2	105
Japan	Germany	3	89
Japan	France	4	74
Japan	Australia	5	67
Japan	China	6	60
Japan	Thailand	7	52
Germany	United States	1	153
Germany	United Kingdom	2	129
Germany	France	3	126
Germany	Switzerland	4	117
Germany	Netherlands	5	93
Germany	Austria	6	91
Germany	Spain	7	67

Table 6: Country bilateral popularity ranking

Industry name	**SIC**	**Entry ranking**
Prepackaged Software	7372	1
Business Services, n.e.c.	7389	2
Engineering Services	8711	3
Information Retrieval Services	7375	4
Plastics Products, n.e.c.	3089	5
Motor Vehicle Parts and Accessories	3714	6
Management Consulting Services	8742	7
Crude Petroleum and Natural Gas	1311	8
Semiconductors and Related Devices	3674	9
Plastics Materials, Synthetic Resins, and Nonvulcanizable Elastomers	2821	10
Electronic Components, n.e.c.	3679	11
Special Industry Machinery, n.e.c.	3559	12
Business Consulting Services, n.e.c.	8748	13
Chemicals and Chemical Preparations, n.e.c.	2899	14
Computer Integrated Systems Design	7373	15
Commercial Physical and Biological Research	8731	16
Industrial Machinery and Equipment	5084	17
Industrial Inorganic Chemicals, n.e.c.	2819	18
Pharmaceutical Preparations	2834	19

Table 7: Industry inflow ranking

Country	Number of firms
United States	15975
United Kingdom	8802
France	4202
Germany	3905
Japan	3288
Canada	3250
Netherlands	2507
Switzerland	1958
Australia	1737
Sweden	1411
Denmark	972
Ireland	945
Italy	938
Finland	908
Belgium	861
Spain	763
Norway	504
South Africa	501
Austria	440
Israel	254

Table 8: Number of acquiror's from each country with sales data.

Acquisition form	Number of deals	Percent
Acquisition of Certain Assets.	595	1
Acq. Majority Interest	15,442	14
Acq. Partial Interest	22,727	20
Acq. Remaining Interest	4,771	5
Acq. of Assets	56,063	49
Merger	15,494	11

Table 9: Type of deals.

6.4 Industrial co-production using Input-Output proximity measures

The extent of association in this analysis was measured implicitly, with no resort to any outside data other than the M&A data. Next, to provide evidence on what determines profitability of acquiring assets in industry j being a producer in industry i, I combine the dataset with the US Input-Output use table. In the same spirit that physical geography has been found to be relevant to explain international trade and finance flows, I construct a measure of industry proximity using IO data. Table 10 represents the IO use table for N industries where v_{ij} represents value of flows in dollars from industry i to industry j.

		Target						
	Industry	1	...	j	...	N	Total supply
Acquiror	1	v_{11}		v_{1j}		v_{1N}	v_{1s}
	...						
	i	v_{i1}		v_{ij}		v_{iN}	v_{is}
	...						
	N	v_{N1}		v_{Nj}		v_{NN}	v_{Ns}
	Total demand	v_{d1}		v_{dj}		v_{dN}	

Table 10: IO Use table. Values are flows.

I first construct input and output shares for each industry as follows. α_{ij} is j's input demand share in industry i, ie, $\alpha_{ij} \equiv \frac{v_{ij}}{v_{dj}}$. Analogously, δ_{ij} is $i-th$ supply share sold to industry j ie, $\delta_{ij} \equiv \frac{v_{ij}}{v_{is}}$. We are ready now to define an inverse distance measure that captures the degree of proximity or horizontality between any industry pair.

Definition 1 *Proximity measure between industry pair* (i, j)*:*

$$P_{ij}^{d} = \frac{\sum_{k \in \Omega_{ij}} (\alpha_{ik} + \alpha_{jk})}{2} \tag{27}$$

where Ω_{ij} *is the set of common inputs.*

Analogously, I construct a supply proximity measure, P_{ij}^s. Intuitively, I am interested in capturing the degree of overlap in the technology that an acquiror and target use, whether it is from the input use or the output supply perspective. The idea is that the greater the common input cost share of an industry pair, the lower is the potential efficiency losses due to monitoring distant activities. Analogously, the greater the common share of output, the easier it will be to carry over the knowledge between industries.

Figure 17 shows the M&A patterns compared to equal probability sampling of IO industries. Both input and output proximity measures show clustering of similar activities, in line with Alfaro and Charlton (2010) and the business relatedness literature[33]. M&A industry pairs are closer than IO industry pairs suggest. More importantly, the value of the proximity measure is large in both cases, with more than 80% of acquisitions falling closer than 50%, ie, an overwhelming majority of deals involve an acquiror and target pair that have at least an average of 50% input demand and output supply in common. An acquiror and a target in a deal share a quantitatively large fraction of their input structure and output destinations.

[33] See Fan and Lang (2000) and Teece et al. (1994).

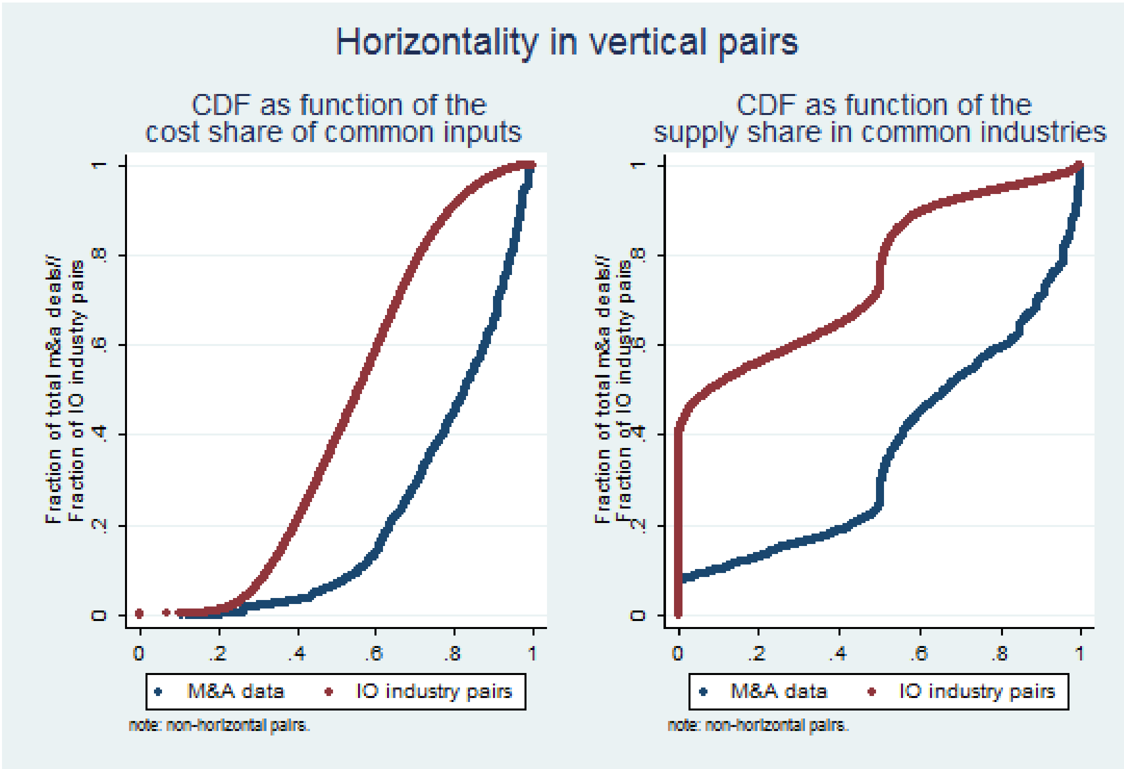

Figure 17: Input demand and output supply proximity measures, M&A vs IO industry pairs. Acquiror's and targets are much closer than IO data would suggest. Moreover, the proximity found is remarkably large: more than 80% of deals share at least 50% of input demand or output supply.

6.5 The role of inputs

To study the role of input sourcing in the dataset, I follow a standard procedure. The potential for vertical integration for an industry pair is measured, based on the Input-Output Use table, as an indicator using a cutoff for the bilateral input demand and/or supply share. More precisely, the measures I use are,

Definition 2 *VI*

VI=1 if $\max\{\alpha_{ij}, \alpha_{ji}\} > .05$*, VI=*0 *otherwise. Industries i and j are said to have* *vertical integration potential according to this measure if VI=1.*

Definition 3 *VI2*

VI2=1 if $\max\{\alpha_{ij}, \alpha_{ji}, \delta_{ij}, \delta_{ji}\} > .05$*, VI2*=0 *otherwise. Industries* i *and* j *are said to have vertical integration potential according to this measure if VI2=1.*

In words, if either input shares or either of both input and output shares are larger than 5% then the industry pair is said to have vertical integration potential for measures VI1 and VI2 respectively. Hortaçsu and Syverson (2007) use VI2 measure whereas other papers in the literature use VI.

A feature of the US IO use table that BEA provides is that it is constructed with more detail in manufacturing industries. Since out of 500 industries more than 350 are manufacturing and the rest are services, manufactures are approximately represented at 4 digit SIC. Accordingly, I check consistency of results for both all industries and only manufacturing. Also, another feature of the IO table is that there is intra-industry input sourcing for many industries and around 40% of them have vertical integration potential, ie, industries source from themselves and so purely horizontal deals appear as having VI potential. This will also imply the need to check robustness of vertical potential results along this dimension.

Table 11 compares the fraction of IO and M&A industry pairs with vertical integration potential for both measures when using the primary industry of Ultimate Parents and their respective targets. While the results show that there are more vertical relations in M&A data than IO industry pairs suggests, the majority of pairs are not vertically connected. Focusing on manufacturing makes the VI indexes drop. This is noteworthy because it reveals that the coarse IO classification available for services overestimates the relevance of vertical integration potential. Moreover, when excluding deals where both firms belong to the same industry in manufacturing, only 24% are connected. According to this table, around 20% of deals have purely vertical integration motivations. This is a lower bound on vertical integration potential because I use ultimate parent data and the primary industry. Next I present evidence on an upper bound.

	VI		**VI2**		**Obs**	
	IO	M&A	IO	M&A	IO	M&A
All data	2	39	3	44	214369	109381
All non-horizontal	2	24	3	29	213906	74020
Manufacturing	2	29	2	34	202705	53650
Manufact. nh	2	20	2	24	202350	42316

Table 11: Input relations and vertical integration potential at UP level, IO vs M&A. Displayed values are percent of industry pairs.

Table 12 presents evidence from acquirors, as opposed to UP's, and targets, hence I go deeper in the ownership structure of the acquiror to get as close as possible to the target. In trying to obtain an upper bound on vertical integration potential, I use for both acquiror and target involved in a deal, up to five 4-digit SIC codes and I keep the highest VI potential out of all possible acquiror-target industry pairs. In this case, the effect of horizontal deals is greater than in the previous table while the manufacturing-services distinction is less relevant. Removing the horizontal deals, makes the fraction of VI pairs fall to around 40%.

	VI		**VI2**		**Obs**	
	IO	M&A	IO	M&A	IO	M&A
All data	2	62	3	67	214369	104551
All, non-horizontal	2	36	3	42	213906	45491
Manufacturing	2	61	2	67	202705	54213
Manufact. nh	2	38	2	44	202350	27711

Table 12: Input relations and vertical integration potential at acquiror level, IO vs M&A. Displayed values are percent of industry pairs for IO data and percent of deals in FDI data. For each deal, I keep the maximum VI measure among all the possible industry links between target and acquiror.

Importantly, in unreported results, taking the average rather than the maximum of VI measures on all possible industry pairs within a deal, I obtain similar results to table 11. For this reason, I conclude that the role for pure vertical pairs is much closer to 24% of non-horizontal deals than to 44%, since on average most industry links in a deal are not related to input sourcing. Hence, while most of deals are horizontal, there are also vertical pairs, but more importantly, the largest part of deals are unclassified if we use this methodology.

In fact, a better description of the data is obtained when looking at IO proximity measures of non-horizontal deals. Two issues are worth mentioning in this respect. On the one hand, on average all inputs represent 50% of the sales of an industry, a $x\%$ input cost share cutoff implies a $\frac{x}{2}\%$ of sales. On the other hand, the proximity that we observe in IO proximity measures are quantitatively relevant, with more than 80% of the deals having 50% of their average cost expenditure and output supply in common. There is great IO proximity between non-horizontal deals, rationalizing not only unclassified or complex ownership links but also presenting an important reason for vertical pairs, ie, common know-how of the business lines.

Rather than attempting a theory that explains ownership *and* input trade in a multi-country, multi-industry setting, I focus my attention on the former. I do this for various reasons. First, table 1 shows that by focusing on FDI, I will be able to analyze a much larger part of international sales. Second, I have shown evidence on the relatively lower importance of vertical integration in ownership compared to ”horizontality” factors. Third, half of my sample is services, for which we know international sourcing is minuscule.

The model in section 4 will not be based on the proximity-concentration hypothesis[34] for two reasons. First, half my FDI data is on services where the export vs FDI trade-off does not exist. Second, while there is evidence that links both FDI and trade, this has not prevented the study of trade in isolation in numerous and relevant empirical and theoretical contributions. Accordingly, I focus on solely on FDI.

6.6 Derivation of equation 2

Let's slowly build our intuition[35]. For bin k, the probability of it being empty after n balls are assigned is $(1 - p_k)^n$. Similarly, the probability that bins j and k are both empty is $(1 - p_j - p_k)^n$. Hence, the probability that either j or k are non-empty is $(1 - p_k)^n + (1 - p_j)^n - (1 - p_j - p_k)^n$.

Obtaining the probability of a number m of empty bins say X is much harder. For the case $m = K - 1$, we have

$$P(X = K - 1) = \sum_{k=1}^{K} p_k^n \tag{28}$$

since we only need to consider cases in which all balls have to land in one bin. Now, define E_k as the event that bin k is empty and using the inclusion-exclusion principle,

[34]In its simplest form, the proximity-concentration hypothesis states that because FDI requires fixed costs of production and trade only variable costs, we should observe that FDI over trade is decreasing in the physical distance between supplier and host. See Brainard (1997) for an empirical study and Helpman, Melitz, and Yeaple (2004) for an elegant theoretical extension of this idea.

[35]The subsequent exposition follows Kotz and Johnson (1977)

$$P(X=0|n) = 1 - Pr(\bigcup_{k=1}^{K} E_k) = 1 - \sum_{k=1}^{K} Pr(E_k) + \sum_{k}^{K}\sum_{k<k'}^{K} Pr(E_k \bigcap E'_k) - ...$$

$$+(-1)^{K-1}\sum_{k}^{K} Pr(\bigcap_{k'\neq k}^{K} E'_k) = \quad (29)$$

$$1 - \sum_{k=1}^{K}(1-p_k)^n + \sum_{k}^{K}\sum_{k<k'}^{K}(1-p_k-p_{k'})^n - ...$$

$$+(-1)^{K-1}\sum_{k}^{K} p_k^n \quad (30)$$

The probability $P_{12...i}$ that the 1st, 2nd, ... i-th bins are empty and none of the other bins are empty is,

$$P_{12...i} = \left(1 - \sum_{k=1}^{i} p_k\right)^n xPr(\text{ no empty bins among (k+1)-st,..., K-th bins|n balls}) \quad (31)$$

The latter part of the expression relates to a situation with $(m-k)$ bins with probabilities $\frac{p_j}{1-\sum_{i=1}^{k} p_i}$ for $j = k+1, \cdots, K$. Changing the values of K and p_k, we can use 29 to write

$$P_{12...i} = \left(1 - \sum_{k=1}^{i} p_k\right)^n - \sum_{j'=k+1}^{K}\left(1 - p_{j'} - \sum_{j=1}^{K} p_k\right)^n + \sum_{j'\neq j}^{K}\sum_{j''>k}^{K}\left(1 - p_{j'} - p_{j''} - \sum_{j=1}^{k} p_j\right)^n - ... \quad (32)$$

By summing expressions similar to this for all $\binom{K}{k}$ possible subsets of size k chosen from $E_1, \cdot, E_K$,

$$P(X=k)=\sum_{a\in a(k)} P_{a_1\ldots a_k}=\sum_{a\in a(k)}\left(1-\sum_{j=1}^{k} p_{a_j}\right)^n-\binom{k+1}{k}\sum_{a\in a(k+1)}\left(1-\sum_{j=1}^{k+1}\right)^n$$
$$+\binom{k+2}{k}\sum_{a\in a(k+2)}\left(1-\sum_{j=1}^{k+2}\right)^n+\cdot+(-1)^{m-k}\binom{K}{k}\sum_{j=1}^{K} p_j^n \tag{33}$$

where $\sum_{a\in a(k)}$ denotes summation over all subsets $a=(a_1,a_2,...,a_k)$ of k integers from the $1,2,...,K$ categories. Note there are $\binom{m}{k}$ terms in the sum $\sum_a^{(k)}$. From this expression, we can obtain $P(Y=m|N=n)$ since $Y=K-X$.

References

Alfaro, L., and A. Charlton. 2010. "Intra-Industry Foreign Direct Investment." *American Economic Review, Forthcoming.*

Alvarez, F., and R.E. Lucas. 2005. "General Equilibrium Analysis Of The Eaton-Kortum Model Of International Trade." *mimeo, The University Of Chicago.*

Anderson, J. E., and E. Van Wincoop. 2004. "Trade costs." *Journal Of Economic Literature* 42, no.3:691–751.

Antràs, P. 2003. "Firms, Contracts, and Trade Structure." *The Quarterly Journal of Economics* 118 (4): 1375–1418.

Antràs, P., and E. Helpman. 2004. "Global Sourcing." *Journal of Political Economy* 112 (3): 552–580.

Arkolakis, Costas, and Marc-Andreas Muendler. 2008. "Extensive Margin of Exporting Goods: A Firm Level Analysis." *mimeo Yale University.*

Armenter, Roc, and Miklós Koren. 2008. "A Balls-and-Bins Model of Trade." *CeFiG Working Papers 3, Center for Firms in the Global Economy.*

Axtel, Robert. 2001. "Zipf Distribution of U.S. Firm Sizes." *Science* 293(5536):1818 – 1820.

Bernard, A. B., S. Redding, and P. Schott. 2009a. "Multi-Product Firms and Product Switching." *AER, forthcoming.*

———. 2009b. "Multi-Product Firms and Trade Liberalization." *NBER WP12782.*

Brainard, S. 1997. "An Empirical Assesment Of The Proximity-Concentrantion Trade-Off Between Multinational Sales And Trade." *American Economic Review*, vol. 87(4).

Breinlich, Holger. 2008. "Trade Liberalization and Industrial Restructuring through Mergers and Acquisitions." *Journal of International Economics*, vol. Vol. 76, No. 2.

Carr, D., J. Markusen, and K. Maskus. 2001. "Estimating The Knowledge-Capital Model Of The Multinational Entreprise." *American Economic Review*, vol. 91(3).

Chaney, T. 2008. "Distorted Gravity: the Intensive and Extensive Margins of International Trade." *American Economic Review,*, vol. Vol. 98, No. 4.

Eaton, J., and S. Kortum. 2002. "Technology, Geography And Trade." *Econometrica* 70(5):1741–1779.

Eaton, J., S. Kortum, and F. Kramarz. 2009a. "An anatomy of international trade: Evidence from French firms." *Unpublished.*

———. 2009b. "Dissecting Trade: Firms, Industries, and Export Destinations." *The American Economic Review* 94(2):150–154.

Eckel, C., and J.P. Neary. 2010. "Multi-Product Firms and Flexible Manufacturing in the Global Economy,." *Review of Economic Studies.*

Ellison, G., and E.L. Glaeser. 1997. ""Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach"." *Journal of Political Economy* 105(5):889–927.

Fan, J., and L. Lang. 2000. "The Measurement of Relatedness: An Application to Corporate Diversification." *Journal of Business*, vol. 73.

Feinberg, S., and M. Keane. 2005. Chapter The Intra-firm Trade of U.S. MNCs: Findings and Implications for Models and Policy of *The Impact of Foreign Direct Investment on Development: New Measurements, New Outcomes, New Policy Approaches.*, edited by Graham E. Blomstrom, M. and T. Moran. Institute for International Economics.

Gabaix, X. 1999. "Zipf's Law for cities: An explanation." *The Quarterly Journal of Economics*, vol. 114 (3).

———. 2008. "Power Laws in Economics and Finance." *Annual Review of Economics* 1:255–93.

Gabaix, X., and R. Ibragimov. 2007. "Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents." *NBER WP0342.*

Head, K., and J.C. Ries. 2005. "FDI as an Outcome of the Market for Corporate Control: Theory and Evidence." *Journal of the Japanese and International Economies* 19(2):215–232.

———. 2008. "FDI as an Outcome of the Market for Corporate Control: Theory and Evidence." *Journal of International Economics* 74(1):2–20.

Helpman, E., M. Melitz, and S. Yeaple. 2004. "Export versus FDI with Heterogeneous Firms." *American Economic Review*, vol. 94(1).

Hortaçsu, A., and C. Syverson. 2007. "Vertical Integration and Production: Some Plant-Level Evidence." *University of Chicago mimeo.*

Irrarazabal, A., A. Moxnes, and L.D. Opromolla. 2008. "The Margins of Multinational Production and the Role of Intra-firm Trade." *University of Oslo mimeo.*

Klette, T.J., and S. Kortum. 2004. "Innovating firms and aggregate innovation." *Journal of Political Economy* 112(5):986–1018.

Kotz, S., and N. Johnson. 1977. *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory.* Edited by Wiley Series in probability and mathematical statistics.

Mayer, Thierry, Marc J. Melitz, and Gianmarco I.P. Ottaviano. 2009. "Market size, Competition, and the Product Mix of Exporters." *Unpublished.*

Melitz, Marc. 2003. "The Impact Of Intraindustry Reallocations And Aggregate Industry Productivity." *Econometrica* 76:6:1695–1725.

Navaretti, G.B., and A. Venables. 2004. *Multinational Firms In The World Economy.* Edited by Princeton.

Nocke, V., and S. Yeaple. 2007. "Cross-border mergers and acquisitions vs. greenfield foreign direct investment: The role of firm heterogeneity." *Journal of International Economics* 72:336–365.

———. 2008. "An Assignment Theory Of FDI." *Review of Economic Studies* 75:529.

Nunn, N. 2007. "Relationship-Specificity, Incomplete Contracts and the Pattern of Trade." *Quarterly Journal of Economics* 122 No.2:569–600.

Penrose, Edith. 1959. "The Theory Of The Growth Of The Firm." *Oxford University Press*, p. 272.

Ramondo, N. 2006. "Size, Geography And Multinational Production." *mimeo UT.*

Simon, H.A. 1955. "On a class of skew distribution functions." *Biometrika*, vol. 42.

Sutton, J. 1997. "Gibrat's Legacy." *Journal of Economic Literature.* 35(1):40–59.

———. 2001. *Technology and Market Structure.* Edited by Cambridge. The MIT Press.

Teece, D. J., R. Rumelt, G. Dosi, and S. Winter. 1994. "Understanding corporate coherence: Theory and evidence." *Journal of Economic Behavior and Organization* 23 no.1:1–30.

UNCTAD. 2000. "World Investment Report."