Heterogeneity and Uncertainty in the Dynamics of Firm Selection into Foreign Markets

Mehmet F. Ulu*
Central Bank of the Republic of Turkey

July 12, 2013

Abstract

Firm-level data indicates a positive relationship between a firm's revenues from a market and the number of markets penetrated by that firm, and previous presence in that market. After studying the role of different types of firm and market-specific shocks in firms' selection decisions, I quantify an entry-cost-reducing effect of previous presence in a market, and increasing returns to being in more markets. I find that being in an additional market increases the demand in other markets between 1% and 3% across different sectors. Additionally, a variance decomposition between firm and market-specific heterogeneity and idiosyncratic uncertainty in firms' selection problem indicates that 1) firm-specific heterogeneity explains more of the total residual variation in revenues from foreign markets as opposed to idiosyncratic variation in technology intensive industries than less technology intensive ones and 2) the relative importance of idiosyncratic components diminishes as the level of per capita income of a destination market increases.

1 Introduction

Various aspects of firm turnover in export markets and the variation in countries that firms sell their products have been studied, and we already know that successful firm-country matches persist over time, and indicate presence of some sunk costs of operating in a market. However, whether there are external returns to being in more markets at the firm-level has not been completely studied. The basic setup with firms that are heterogeneous only along with their

*Email: fatih.ulu@tcmb.gov.tr. I am grateful to Sam Kortum, Thomas Chaney and Chad Syverson for their encouragement and advice. I'd like to thank Costas Arkolakis, Kerem Cosar, Ralph Ossa and seminar participants at the University of Chicago. All errors are mine. Correspondence: Fener Kalamis cad. Atlihan sok. 30/A Fenerbahce Kadikoy Istanbul.

productivities and common entry costs could not explain the variation in country groups that firms penetrate into. Hence, researchers need to add more dimensions of heterogeneity such as demand and entry cost shocks into the study of firm selection into foreign markets.

In this paper, I study firms' selection decisions in order to understand observed hysteresis of firm presence in a market and to test for external increasing returns to being in multiple markets. I study how the sunk and fixed costs of operating in a market affect the selection decisions of forward-looking firms in an environment where the sources of uncertainty are firm-time-specific productivity shocks and firm-time-country-specific demand and entry cost shocks. After studying the decision problem, I ask the question: How can we handle the unobserved firm and market-specific heterogeneity in these decision problems while having empirical tests for the mentioned returns? Using the data from Turkish manufacturing firms and accounting for the unobserved heterogeneities across firms and destination markets this paper will quantify the returns to previous presence in a market and being in more markets.

Previous studies have analyzed the returns to previous presence in export markets and also returns to being in more markets.[1] Some of these studies do not observe the decisions about specific destination countries.[2] Morales et al. (2011) observe the decisions about specific destination markets. They measure the impact of exporting to a bundle of countries in the previous period on the sunk costs of exporting to another country during the current period, and name this impact "extended gravity". While seeking the extended gravity forces, they project the sunk costs and fixed costs of exporting on four dummies -border, language, continent, and GDP group. Whenever an entry decision on a new market shares one of the mentioned attributes with a market that was served in the previous period, the related dummy variable gets the value one and reduces the sunk costs of entry along the shared market characteristics.

This paper expands on the empirical studies of Bernard and Jensen (2004) and Das et al. (2007) by incorporating firms' selection decisions regarding specific markets, whereas the question in the two studies is about only gaining the exporter status. The discrete choice study of the decisions about different markets requires the researcher to control for the unobserved heterogeneity across both firms and the destination markets. Having dummy variables for each market in the model causes loss of identification.[3] This study suggests a way to handle

[1]Sunk entry costs have been incorporated into the empirical analysis of export decisions by Bernard and Jensen (2004), Das et al. (2007), and Morales at al. (2011), among others.

[2]Both Bernard and Jensen (2004) and Das et al. (2007) analyzed firms' decisions about gaining exporter status. Bernard and Jensen (2004) examine the factors that increase the probability of entry into exporting by using data from U.S. manufacturing firms. While trying to identify the role of sunk costs and unobserved plant heterogeneity, they produce their main results by employing a linear probability model. They find that exporting today increases the probability of exporting tomorrow by 39%, and unobserved plant heterogeneity is important in the export decisions. Das et al. (2007) employ a dynamic structural framework in their empirical study of export participation. They examine the export profits of firms by fitting an ARMA(2,1) structure for export profits and estimate a type II Tobit model for the observed export revenues. The data does not allow them to model export profits from each market entered, so they perform the analysis using firm-level aggregate export information.

[3]This is a standard problem when too many dummies are introduced into a discrete choice estimation.

this unobserved heterogeneity problem. In Morales et al. (2011) the extended gravity affects only the sunk costs.[4] However, in this study, I find that being in more markets significantly impacts the demand side, which should be added to the model. In Morales et al. (2011) the remaining error term in the decision problem is assumed to be idiosyncratic at firm-country-year and no correlation across observations regarding a firm or firm-country relationship is allowed. Nevertheless, my setup accounts for such correlations and my findings indicate that disregarding that covariance may yield biased measurements. Controlling for the differences in firm-specific productivity and accounting for the demand and entry cost shocks across markets, I find that being in one more market increases the demand in other markets between 1% and 3% across different sectors. For instance, else equal for two firms, one is already in 10 foreign markets and the other in no foreign markets. Then, the former one is expected to have 10% to 30% more demand in a new country than the latter one. If I do not control for these unobserved firm and market-specific heterogeneities, the estimated returns to being in more markets are clearly biased upwards.[5]

To understand the content of the empirical study of the export decisions, I first study the decision problem in a theoretical environment.[6] I augment the basic model with heterogeneity in only firm-specific productivities (e.g. Melitz (2003)) with more dimensions of shocks to capture the deviations from a hierarchy of popular markets to enter.[7] When sunk costs exist for entry into the markets, some future returns to these costs -option values- emerge in the decision problems. Das et al. (2007) compute these values in their study by allowing the firms to have perfect foresight up to 30 years, and assume the same stochastic rules to govern the profits from the export markets throughout this time horizon. Instead, I show that those option values are functions of the current period's state variable -unobserved productivity, demand, and entry cost shocks that hit firms, and also unobserved firm-specific abilities and attributes. If the researcher can observe sufficiently many selection decisions and resulting revenues from these decisions, accounting for those option values becomes possible.

Implementation of some semi-parametric estimation techniques gives full control of the variance terms in the empirical study which allows a decomposition of the initially unexplained

[4]In my empirical analysis I will be studying the effect of the number of markets instead of the "extended gravity" impacts. However, my discussion is still relevant since firm-country export transactions are quite persistent over time.

[5]When I do not account for these heterogeneities estimates go up from 1% - 3% to 7% - 11%.

[6]Das, Roberts et al. (2007), Chaney (2011), Albornoz et al. (2011) and Arkolakis (2011) are some examples of these studies. Albornoz et al. (2011) studies a theoretical model in which a firm learns about its profitability as an exporter only after engaging in exporting. Chaney (2011) provides a network model of exporters and importers to explain the cross-sectional distribution of the number of foreign markets accessed by individual exporters and the dynamics of firm-level exports. Arkolakis (2011) handles firm selection into export markets and growth in a calibration study, assuming specific stochastic processes. However, I don't make specific distribution assumptions. In this sense my model is more general. My study complements all these studies by discussing the possible sources of uncertainties, how they affect decisions and how they can be mapped into an empirical framework.

[7]The sources of shocks in my setup are similar to the ones in Eaton, Kortum and Kramarz (2011).

variation in firm revenues between the unobserved heterogeneity terms and the idiosyncratic shocks. This decomposition reveals some interesting findings. Firstly, as the technological requirements of sectoral production increase, the share of idiosyncratic components weakens in total variation of the revenues, and instead the share of firm specific heterogeneity increases. In Apparel and Food industries idiosyncratic shocks explain most of the residual variation. However, in Machinery and Automotive industries, firm-specific heterogeneity explains much of the residual variation. Secondly, the relative importance of the idiosyncratic components diminishes as the level of per capita income of the destination markets increases.

In the following sections, I describe the data, and then I present the model. After discussing the identification of the model, I explain the estimation procedures, and present the results. After discussing the results, I conclude the analysis.

2 Data Description

This paper uses firm-level data from Turkish manufacturing industries. Turkey is not a unique country like China and India in terms of its labor resources, though it is a good representative of countries with large populations like Pakistan, Bangladesh, Egypt, Philippines, Vietnam, and Thailand. Since it is not a naturally resource-abundant country, Turkey had to achieve diversification in its industrial production. Turkish governments adopted export-oriented growth policies several decades ago, and the recent performance of the Turkish economy indicates a vibrant economic environment inside the country.

In this study, I use two datasets from the Turkish Statistical Institute (TurkStat). The first dataset is the Structural Industry Dataset, which collects several balance sheet and income statement items of Turkish firms from 2003 to 2008. In this dataset Turkstat collects data for 20+ employee firms continuously, whereas it selects the firms in the sample that employ less than 20 people every year. I include only 20+ employee firms in the analysis. My second dataset is the Turkish customs-level export transactions data from 2002 to 2008. I combine the two datasets using a key derived from firms' tax identity numbers.

Presenting some salient features of the data may be helpful to describe the environment in which the data were collected. Throughout the analyzed period exports to every geographic destination have been rising. Exports to different economic development groups (EU, OECD, OPEC) have simultaneously increased. In all manufacturing industries, exports have increased. Export revenues, share of export revenues in firm revenues, and average sales of specific goods to specific destination markets have distributions similar to lognormal distributions.

I analyze the selection decisions and the resulting revenues only about the most popular 27 markets which make up about 95% of all foreign sales for Turkish firms during the analyzed period. [8] A country may not be within the first most popular 27 destinations each year in the

[8] Employing data for all destination markets increases the time and effort cost without adding much insight

analyzed time period which is why I selected a few countries that did not show up among the most popular 27 for a couple of years.[9]

I include only the firms with NACE codes 15 (Food), 18 (Apparel), 28 (Metals), 29 (Machinery), and 34 (Automotive) in the analysis. These industries make up for, on average, 40% of both revenue and value added in total Turkish manufacturing industries (NACE 15-37) throughout the sample period. They are also diverse in their production technologies and relative intensities in using different production factors. Table 1 shows the panel structure of the firms in the sample. For instance, 3448 firms in the sample are consistently within the sample over the given time horizon, whereas 1356 firms enter the sample at the beginning of the third year and stay until the end of the fourth year.[10]

Table 1: Panel Structure of the Firms in the Sample

Freq.	Percent	Cum.	Pattern
3448	34.18	34.18	111111
1356	13.44	47.63	..1111
781	7.74	55.37	...111
666	6.6	61.97	.11111
586	5.81	67.78	1
434	4.3	72.08	11
418	4.14	76.23	11....
247	2.45	78.68	1.....
203	2.01	80.69	.1....
1948	19.31	100	(other)
10087	100		

Table 2 displays the number of destination markets for the firms in the sample from 2003 through 2008. Throughout the time period the mean of the number of foreign markets in which a firm sells its products is around two. Ninety percent of exporter firms export to at most seven markets and 99% of exporter firms send their products to, on average, at most 16 markets.

Percentile information about the export revenue distribution is given in Table 3. It's evident that the export revenues are going up at each percentile throughout the sample time period.[11]

Table 4 shows that around 60% of all firm-country matches continue from previous years for

into the analysis.

[9]Germany, England, Italy, France, Russia, Spain, USA, Romania, U.A.E., Netherlands, Iraq, Greece, Bulgaria, Belgium, Israel, Saudi Arabia, Ukraine, Iran, Poland, Algeria, Kazakhistan, China, Switzerland, Cyprus, Egypt, and Austria were consistently within the first 30 countries throughout the sample period. Azerbaijan was on the list four times. Denmark showed up in the first 30 three times while Syria showed up among the first thirty twice.

Sales in these first twenty seven countries make up 78% of all exports in 2007, 79% in 2006, 80% in 2005, 81% in 2004 and 80% in 2003.

[10]The split panel structure of the firms exhibited in Table 1 among the firms of different industries is available upon request.

[11]The split distribution exhibited in Table 3 among the firms of different industries is available upon request.

Table 2: Number of Foreign Destination Markets for the Firms in the Sample

year	N	Mean	S.D.	p90	p99
2003	4,909	2.10	3.56	7	16
2004	5,773	2.05	3.54	7	16
2005	6,825	2.05	3.54	7	16
2006	7,332	2.02	3.57	7	16
2007	7,540	2.02	3.61	7	17
2008	7,674	2.01	3.65	7	17

Table 3: Distribution of Exports for All Firms in the Sample (thousand USD)

year	p25	Median	p75	p90	p99
2003	41	247	1,247	4,740	40,200
2004	48	258	1,257	5,238	48,600
2005	49	257	1,173	4,728	43,100
2006	58	302	1,481	5,397	47,800
2007	68	371	1,793	6,664	55,400
2008	81	458	2,123	7,494	71,800

Turkish exporter manufacturing firms throughout the sample period. Aside from the persistence in the shocks that hit firm-country relationships, sunk costs to entry into the markets may also contribute to the persistence in the observed firm-country transaction matches.

Table 4: The Share of Continuing Firm-Country Pairs in All Firm-Country Pairs

Year	Share of the Continuing Pairs
2003	0.63
2004	0.58
2005	0.58
2006	0.64
2007	0.67
2008	0.68

Notes: This table shows the yearly ratio of firm-country matches which were present in the previous year, as well as the current year for Turkish manufacturing firms. Own calculations from the data that has been provided by TurkStat.

One purpose of this paper is to measure the benefit of being in more markets, and in Table 5, firm revenues increase in different markets by the number markets penetrated by the firms. This positive correlation between the firm revenues and the number of markets penetrated by the firms can be a result of the firm-specific productivites such that more productive firms enter into more markets and also sell more in those markets. This table gives only a hint about the possibility of some positive returns to being in more markets. As it is obvious, this table carries the market size effects on the reported numbers.[12]

Table 5: Average Sales for Firms with Different Number of Markets Penetrated (Thousand USD in 2003)

	Number of Markets Penetrated							
Markets	1	2	3	4	6	11	16	21
AUT	0.45	3.0	4.9	15	23	34	120	390
BGR	2.9	6.5	13	21	62	88	500	350
CHE	0.49	2.2	3.0	4.4	15	50	31	270
CHN	0.97	0.3	1.0	10	4.5	7.0	44	28
DEU	33	96	160	230	530	520	1300	4800
DNK	2.4	3.2	5.4	6.4	38	31	55	140
DZA	0.55	1.3	7.0	14	28	100	980	120
EGY	0.69	2.0	6.3	7.1	34	52	450	610
ESP	1.5	13	40	65	99	160	530	170
FRA	10	35	59	70	120	260	9300	1500
GBR	230	640	60	75	130	390	540	3300
GRC	2.8	9.5E	22	27	39	74	590	430
IRN	1.7	3.7	14	18	43	80	540	1200
ISR	4.8	9.2	35	45	58	150	930	1000
ITA	10	27	67	150	220	390	1900	810
NLD	9.4	11	13	34	71	140	400	640
POL	0.5	6.4	5.4	8.6	13	89	810	240
ROM	2.9	8.1	15	33	29	97	2100	550
RUS	9.8	15	22	15	43	200	1800	1200
SAU	3.3	7.7	20	8.3	12	39	410	110
USA	8.9	50	83	94	210	580	1600	1800

Notes: This table shows how the average revenues change in different destination markets for Turkish manufacturing firms by the number of markets penetrated by these firms. Each cell represents the average sales in the row-country by the firms that have been in a specific number of markets, including the row-country, indicated by the columns.

The transition of the number of foreign markets in which firms sell their products is demonstrated in Table 6. A firm retains the number of its foreign markets with the highest probability, the next highest probability is that a firm adds one more foreign market or loses loses one in the following period. Large jumps between the number of destination markets within two consecutive years become less probable as the distance between the two states grows. The model to be presented in the following section will also show how sunk costs to entry and monotone shocks to demand and productivity may bring about such transition matrices.

[12]Seeking for some more preliminary evidence for the returns to being in more markets I specify a destination market, and the firms that operate in the specified market in two consecutive years. Then I group these firms with the change in the number of markets they cater in between these two years. When I compute the average revenue growth rates in this market for each firm group, I observe that there are some increases in average revenue growth rates as the number of markets catered increases. However, for several destination markets this evidence lacks sufficient number of observations and it is a bit noisy.

Table 6: Transition Matrix for the Number of Markets Penetrated (2002-2003)

	Number of Markets Penetrated at time t (n_t)											
n_{t-1}	0	1	2	3	4	5	6	7-9	10-13	14-20	21-50	50-max
0	0.79	0.16	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.44	0.37	0.11	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
2	0.22	0.27	0.24	0.14	0.06	0.03	0.02	0.02	0.01	0.00	0.00	0.00
3	0.14	0.16	0.20	0.19	0.14	0.07	0.05	0.04	0.01	0.00	0.00	0.00
4	0.08	0.09	0.14	0.16	0.17	0.13	0.08	0.12	0.03	0.01	0.00	0.00
5	0.05	0.05	0.10	0.12	0.14	0.16	0.12	0.18	0.05	0.01	0.00	0.00
6	0.06	0.03	0.04	0.10	0.09	0.14	0.13	0.27	0.11	0.02	0.00	0.00
7-9	0.05	0.02	0.02	0.03	0.04	0.09	0.10	0.35	0.23	0.06	0.01	0.00
10-13	0.03	0.01	0.01	0.02	0.02	0.02	0.04	0.19	0.38	0.25	0.03	0.00
14-20	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.04	0.15	0.50	0.25	0.00
21-50	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.11	0.81	0.04
50-max	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.92

Notes: Each cell shows the probability of going from n foreign markets at time t to m foreign markets at time $t+1$. These probabilities were calculated for the universe of Turkish manufacturing firms in 2002 and 2003. Own calculations from the data that has been provided by TurkStat.

3 Economic Environment

Let $x_{dt}(j,m)$ be the demand that the domestic firm j of industry m receives from country d, where $p_{dt}(j,m)$ is the price of firm j of sector m selling its good in country d at time t, and σ^d is the elasticity of demand in country d. $u_{djt} \in U = [0,\infty]$ is the demand shock that firm j receives in country d at time t, and allowed to be correlated over time.[13][14]

$$x_{dt}(j,m) = \frac{u_{djt}}{p_{dt}(j,m)^{\sigma^d}} \tag{1}$$

Firms

Goods in each country are Armington-differentiated and each good is produced by only one producer. Hence, we can label a firm in a country with the good it produces $j \in [0,1]$.

$z_{j,t} \in Z = [0,\infty]$ is firm j's productivity at time t. A firm with productivity z produces z units of output with one unit of labor. The productivity of the firm z receives Markov shocks over time. The third source of uncertainty in the model comes from the fixed costs of entering into a foreign market, $\psi_{djt} \in \Psi = [0,\infty]$.[15]

I write each equation about firm behavior below for a firm from a particular country and hence no index shows up for this country. Goods trade across countries is subject to iceberg

[13] I load all the shocks in the destination country that can hit the firm's demand in this country onto the demand shock term u_{djt}.

[14] Details about an economic environment and a utility function that can justify the functional form of consumer demand given in Eq.(1) are discussed in the Appendix.

[15] In the following sections, these shocks will be assumed to be a combination of firm-specific and idiosyncratic components across countries and time.

costs. Delivering one unit of a good with unit production cost c from the home country to the destination country, d, costs $\tau_d c$. For operating in country d, firm j has to pay a fixed cost ψ_{djt} every period regardless of whether it is a new entrant or an incumbent firm. If a firm is newly entering the foreign market, it has to pay a sunk cost, $\lambda_\psi \psi_{djt}$, where $\lambda_\psi > 0$ is a constant. As long as the firm stays in that market, it does not have to pay this amount again. But if it drops out of the market, it has to pay the sunk cost if it re-enters. Export participation is expected to be costly due to several factors (new packaging, adapting to destination markets' regulations, finding new customers in the destination markets, building distribution channels, etc.). Such entry costs in export markets have also been found in the literature.[16]

Let e^d_{jt} be the entry decision of firm j regarding country d at time t, and suppressing the firm index j, the cost of selling goods in country d at time t can be represented as $\mathcal{F}^d_t$.

$$\mathcal{F}^d_t(e^d_t, e^d_{t-1}) = \begin{cases} \psi_{dt}(1+\lambda_\psi) & \text{if } e^d_t = 1 \text{ and } e^d_{t-1} = 0 \\ \psi_{dt} & \text{if } e^d_t = 1 \text{ and } e^d_{t-1} = 1 \\ 0 & \text{if } e^d_t = 0 \end{cases} \tag{2}$$

At the time of entry decisions about N countries, firm j's state is $\mathbf{s}_t = \{z_{jt}, \mathbf{u}_{jt}, \boldsymbol{\psi}_{jt}\}$ where $\mathbf{u}_{jt} = \{u_{1jt}, ..u_{djt}, ..u_{Njt}\}$, $\boldsymbol{\psi}_{jt} = \{\psi_{1jt}, ..\psi_{djt}, ..\psi_{Njt}\}$ and $\mathbf{s} \in S = Z \times U^N \times \Psi^N$. The firm's state related to country d is $s_{dt} = \{z_{jt}, u_{djt}, \psi_{djt}\}$, $s_d \in S^d = Z \times U \times \Psi$. Before making a decision, the firm knows about its productivity, demand, and entry cost shocks for the current period. e^d_{jt-1} can also be added to the definition of s_{dt}. However, since it is endogenously determined I want to keep its notation separate. Figure 1 shows the timeline for the evolution of the state s_{dt} over time.

Figure 1: Timeline

$t-1$

e^d_{jt-1}

t

z_{jt}

u_{djt}

$$V(\mathbf{s}_t, \mathbf{e}_{t-1}) = \max_{p_{dt}, e_t^d} \sum_{t=T}^{\infty} \sum\nolimits_d \beta^t E(e_t^d(p_t^d - \tau_d c_t) x_{dt}(p_{dt}) - F_t^d(e_t^d, e_{t-1}^d)) \tag{3}$$

Then, firm's pricing problem is a static one and the optimal price conditional on entry -suppressing the firm indices- is

$$p_t^d = c_t \tau_d \frac{\sigma^d}{\sigma^d - 1} \tag{4}$$

Revenues and gross profits from country d for a firm in sector m conditional on entry are given in Eq.5 and 6.

$$R_{m,t}^d = u_{dt} \left[c_t \frac{\sigma^d}{\sigma^d - 1} \tau_d \right]^{1-\sigma^d} \tag{5}$$

$$\pi_{m,t}^d = \frac{R_{m,t}^d}{\sigma^d} \tag{6}$$

As seen in the above equations, firm-country-time specific demand shocks affect the gross profits from a country multiplicatively. The productivity shocks show up in unit cost of production c and demand elasticities control the effect of these shocks on the gross profit.

The dynamic part of the firm's problem is the entry decisisons. The corresponding Bellman equation for Eq.(3) is[17]

$$V(\mathbf{s}, \mathbf{e}) = \sum\nolimits_d V^d(s_{d,}, e^d) \tag{7}$$

Let $V^{d,1}(s_d, e^d)$ be the value from market d if the firm opts to sell in market d during the current period and $V^{d,0}(s_d, e^d)$ is the value if the firm opts not to sell in market d during the current period.[18]

$$V^d(s_{d,}, e^d) = \max\{V^{d,0}(s_{d,} e^d), V^{d,1}(s_{d,}, e^d)\} \tag{8}$$

$$V^{d,0}(s_{d,}, e^d) = 0 + \beta \int V^d(s_d', 0) \mathcal{P}_d(s_d' | s_{d,}) ds_d' \tag{9}$$

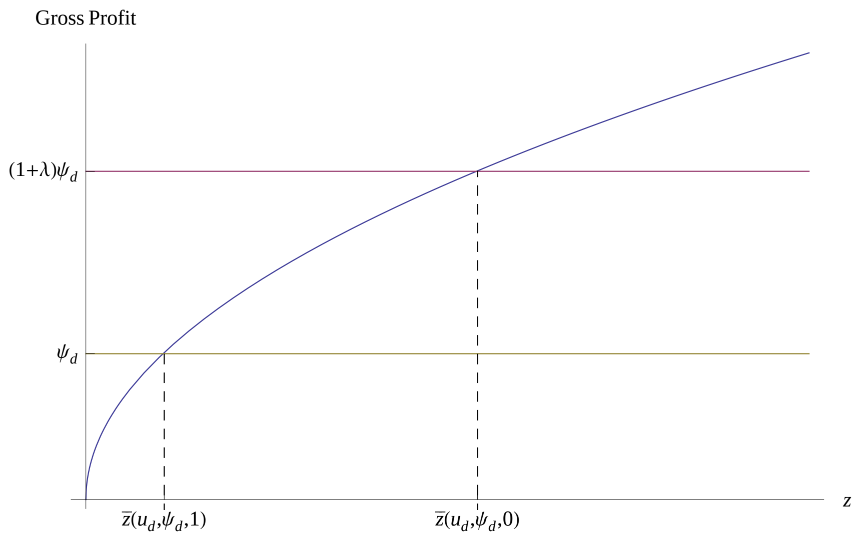
[17] Although, in the emprical section, I will be seeking for some returns to being in more markets, at this stage I do not allow any spillovers from decision about a market towards the decisions about other markets. In the following sections, I will discuss about possible extensions of the model that can allow to endogenize such spillovers, and still be compatible with the empirical analysis.

[18] At the time of decision e^d is a part of firm's state and it represents firm's status in market d in the previous period whereas $e^{d\prime}$ is firm's decision variable at the time of decision.

$$V^{d,1}(s_{d,}, e^d) = \pi^d(z, u_d) - (1 + (1 - e^d)\lambda_\psi)\psi_d) \qquad (10)$$
$$+\beta \int V^d(s'_{d,}, 1)\mathcal{P}_d(s'_d|s_{d,})ds'_d$$

Figure 2 shows how the cutoff productivity levels for given demand and entry cost shocks related to market d are determined. $\bar{z}_d(u_d, \psi_d, 1)$ is the cutoff productivity level for the firm if it is a continuing firm in that market and $\bar{z}_d(u_d, \psi_d, 0)$ is the cutoff productivity level for the firm if it is a new entrant into the same market. The gross profit function in the figure below represents $\pi^d(z, u_d) + \beta \int V^d(s'_{d,}, 1)\mathcal{P}_d(s'_d|s_{d,})ds'_d$.[19]

Figure 2: Entry Decisions



For showing the existence of a value function V and a policy function, I introduce conditional independence of shocks and monotone transitions for the shocks in Assumption 1-3, and Assumption 4 states an Inada-like condition regarding the discounted value of future at each state of nature. Existence of a Markov Perfect (Recursive) Equilibrium under the given assumptions, and a constant wage over time owing to a good that is costlessly traded across countries and produced with the same technology over time is discussed in the Appendix as well as some structural properties of the decision problem.

For any given $(u_{djt}, \psi_{djt}, e^d_{jt-1})$ tuple, $\bar{z}_{djt}(u_{djt}, \psi_{djt}, e^d_{jt-1})$ is the cutoff productivity level for selection into market d.

Let n be the number of export markets that a firm is in within a given year and $\mathcal{N}(n'|n)$ be the conditional distribution of this random variable over time.

[19] It is an increasing function in z. Yet it does not have to be concave as it is in Figure 2.

$$n_{jt}(z_{jt}, \mathbf{u}_{jt}, , \boldsymbol{\psi}_{jt}, \mathbf{e}_{jt-1}) = \sum_{d=1}^{N} e_{jt}^{d}(z_{jt}, u_{djt}, \psi_{djt}, e_{jt-1}^{d}) = \sum_{d=1}^{N} \mathbf{1}(z_{jt} \geq \bar{z}_{djt}(u_{djt}, \psi_{djt}, e_{jt-1}^{d}))$$

For any given firm j that was not in the destination market d at time $t-1$, the probability of entering into that market at time t is $P(e_{jt}^{d} = 1|e_{jt-1}^{d} = 0)$, and the probability of staying out of that market at time t is $P(e_{jt}^{d} = 0|e_{jt-1}^{d} = 0)$. Similar probabilities for an incumbent firm are denoted as $P(e_{jt}^{d} = 1|e_{jt-1}^{d} = 1), P(e_{jt}^{d} = 0|e_{jt-1}^{d} = 1)$, respectively. $P(e_{jt}^{d} = 0|e_{jt-1}^{d} = 0) = P(s_{dt} \in S^{d}$ such that $z_{jt} < \bar{z}_{djt}(u_{djt}, \psi_{djt}, e_{jt-1}^{d})|s_{dt-1} \in S^{d}$ such that $z_{jt-1} < \bar{z}_{djt-1}(u_{djt}, \psi_{djt}, e_{jt-1}^{d}))$.

The monotonicity assumptions that are introduced to each type of shock imply that if firm A has higher productivity today than firm B, firm A's odds are better than those of firm B for having better productivity in the following period. The other assumption of the model is that being an incumbent in a market reduces the sunk costs of selling in that market again. Owing to these assumptions, the probability that a firm's status in a market (incumbent or not) will remain the same is always higher than the probability that there will be a change in the status quo. When we consider a firm's status in several different markets and the evolution of the number of penetrated markets for that firm, maintaining the status quo (being in the same number of markets during two consecutive periods) will be the most probable outcome, and as the size of change (the change in the number of markets penetrated in between two consecutive periods) grows, the odds of that change diminishes.

Lemma 1. *Under the assumptions 2 and 3, we have* $P(e^{d\prime} = 1|e^{d} = 1) > P(e^{d\prime} = 1|e^{d} = 0)$ *for all* d.

For getting Lemma 1 the monotonicity assumptions (2-4) are sufficient. However, for getting Lemma 2 we have to strengthen the assumptions to Assumptions 2'-4' which are first order stochastic dominance assumptions.

Lemma 2. *Under the assumptions 2' and 3', we have* $P(e^{d\prime} = 0|e^{d} = 0) > P(e^{d\prime} = 0|e^{d} = 1)$ *for all* d.

Using the above two findings we can see that for a firm that is in n markets today, the probability of being in more or less markets tomorrow decreases as the change in the number of markets increases over time. This is also the behavior observed for Turkish firms in Table 6. I'll name this property the row-concavity (RCD) of the transition matrix, which is similar to the definition in Sandikci et al. (2008).

Definition. *A transition matrix* $P(n\prime|n)$ *is row concave with the maximum on the diagonal (RCD) if the maximum is on the diagonal and* (i) $P(n+i|n) \geq P(n+i+1|n)$ *for* $i = 1, ...N-n-1$

(ii) $P(n-i|n) \geq P(n-i-1|n)$ *for* $i = 1, ..n-1$

Proposition 1. *Let $\mathcal{N}(n'|n)$ be the transition matrix for the number of markets that a firm is in. If $\mathcal{Z}(z'|z), \mathcal{U}_d(u'|u)u, \mathcal{S}_d(\psi'|\psi)$ satisfy the assumptions A2'-A4', then $\mathcal{N}(n'|n)$ has the RCD property.*

We know that the transition matrix has positive probabilities in its diagonal components. Hence, we know that in the steady state, a unique distribution for the number of markets entered is attained.

In this section, we have analyzed how firms make up their selection decisions being exposed to different types of idiosyncratic shocks which are not easy to observe and include into any empirical study of these selection decisions. In the following section, we propose a way to account for these unobserved heterogeneities in empirical tests regarding these selection decisions.

4 Empirical Model

Revenue in a country conditional on entry

The revenue of a firm in a given country conditional on entry is given in Eq.(5). Taking the logarithm of the revenue equation we get

$$\ln R_{dt}(j,m) = (1-\sigma^d)\left[\ln c(j,t) + \ln \frac{\sigma^d}{\sigma^d - 1} + \ln \tau_d\right] + \ln u_{djt} \tag{11}$$

In Eq.(11), $c(j,t)$ is the firm-specific unit cost of production, $\ln u_{djt}$ is the firm-country-time specific demand shock. Profit is the fraction of the revenue from that market as given in Eq.(6), and $\ln \pi_t^d = \ln R_t^d - \ln \sigma^d$.

I disentangle this demand shock into firm-specific and idiosyncratic components. Let g_j be the firm-specific components of the demand shock and α_{dt}'s be their weights. Both factor groups and their weights in Eq.(12) are vector valued.[20]

$$u_{j,t}^d = (g_j)^{\alpha_{dt}} \exp(\eta_{djt}) \tag{12}$$

Entry Costs

I assume a functional form for firm-country-specific entry costs as given in Eq.(13). Z_{djt} is a vector of observable variables and ξ_{djt} is the unobservable shock component to entry cost. Here I'm tweaking the notation a bit by allowing Z_{djt} to be a vector and γ to be the corresponding vector of powers.

$$\psi_{djt}(Z_{djt}) = (Z_{djt})^{\gamma} \exp(\xi_{djt}) \tag{13}$$

$$\xi_{djt} = (g_j)^{\alpha_{\psi_{dt}}} \exp(\eta_{\psi_{djt}}) \tag{14}$$

[20]Factors in vector g_j are allowed to be correlated with each other. In the estimation procedure I will be able to handle these correlations and diagonalize the covariance matrices of the sets of factors.

The variables in Z_{djt} -Distance, Ext. Margin, Tariff- satisfy the exclusion restriction during the estimation of the choice decisions.[21]

Entry Decision

At the time of decision, a firm compares the expected discounted value of the two actions, entry into a market and no entry. Current period return of entry is the expected net profits today in the destination market after subtracting the fixed costs; current period return to no entry is zero. A firm sells in market d if and only if

$$\begin{aligned} V^{d,1}(s_d, e^d) - V^{d,0}(s_d, e^d) &\geq 0 \\ \pi^d(s_d) - (1 + (1 - e^d)\lambda_\psi)\psi_d + \beta E\Delta V^d(s_d, e^d) &\geq 0 \end{aligned}$$

where $E\Delta V(s_d, e^d) = E(V^d(s'_d, 1) - V^d(s'_d, 0)|s_d, e^d)$ (Option Value). A firm decides whether to enter into a market at the beginning of each period. At the time of decision, a firm knows that even if it does not enter the foreign markets today, it still has the option to enter at the beginning of the following period. At the beginning of each time period, a firm knows in which markets it sold its products during the previous period, and the firm has to decide about the markets in which it wants to be present during the period. As a result of its export participation decisions, between two time periods the firm evolves between the two states: being present in market d or not, $e^{d\prime} = 0$ or 1. Let Ω_{jt} be the information set of firm j at time t. With this information set, firm j calculates the expected returns from the current period and expected discounted value of future returns.[22]

The Content of the Option Values

The option value $E\Delta V(s_d, e^d)$ aggregates over all future states of nature given the current period's state. For analyzing the content of this aggragation we define the function $\Gamma(.)$. Given the current period's state $\Gamma(.)$ shows how the difference $V^d(s'_d, 1) - V^d(s'_d, 0)$ changes by z' after fixing u'_d and ψ'_d, and aggregates over $z\prime$. Then $E\Delta V(s_d, e^d)$ aggregates $\Gamma(.)$'s over u' and ψ'.

[21] In Helpman, Melitz, and Rubinstein (2008), similarly, fixed trade costs that are induced by trade barriers do not influence the variable (iceberg) trade costs and satisfy the exclusion restriction.

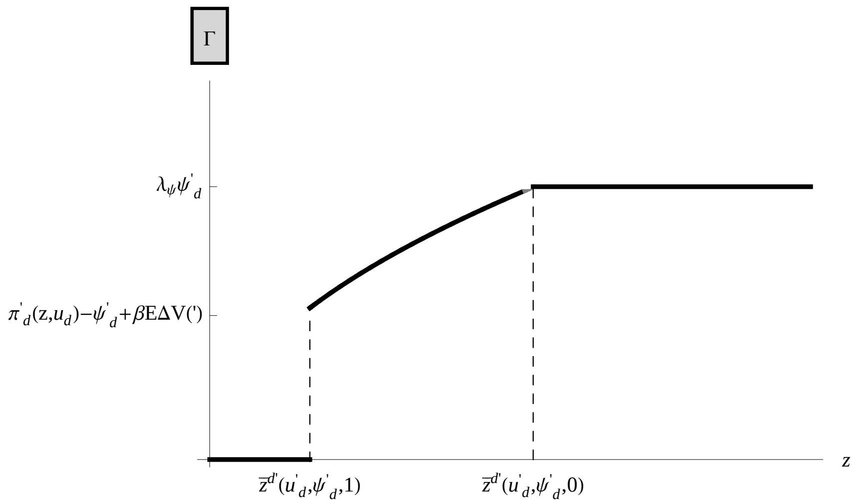
[22] Aside from the shock structure introduced so far, we may have suspicions about some trends in X and Z. Previous lemmas have assumed π^d_{jt} to have the same X - X^d_m, P^d_m in our setting - and Z. Let $\gamma_{X^d_m}, \gamma_{P^d_m}, \gamma_{Z^d}$ be the rates of growth for X^d_m, P^d_m, Z^d, respectively. In such an environment calculation of the expected option value, $E\Delta V(s_d, e^d)$, will change, but the directions of the relationships will not. For the estimations I de-trend the country-specific aggregates -market potential variable- with the country's GDP growth rate.

$$E\Delta V(s_d, e^d) = \int\int \Gamma(s_d, e^d, u'_d, \psi'_d) du' d\psi'$$

$$\Gamma(s_d, e^d, u'_d, \psi'_d) = \begin{cases} 0 & \text{for } z' \in (0, A = \bar{z}^{d\prime}(u'_d, \psi'_d, e^{d\prime} = 1|e^d)) \\ +E(\lambda_\psi \psi'_d | s_d, e^d) & \text{for } z' \in (B = \bar{z}^{d\prime}(u'_d, \psi'_d, e^{d\prime} = 0|e^d), \infty) \\ +E(\pi^d(z', u'_d) - \psi'_d + \beta E\Delta V(s'_d, e^{d\prime})|s_d, e^d) & \text{for } z' \in (A, B) \end{cases}$$

For some future states s', $\Gamma(s_d, e^d, u'_d, \psi'_d)$ is zero where productivity is low and $z' \in (0, \bar{z}^{d\prime}(u', \psi', e^{d\prime} = 1|e^d))$, for some s' with mediocre productivity it is equal to $E(\pi^d(z', u'_d) - \psi'_d + \beta E\Delta V(s'_d, e^{d\prime})|s_d, e^d)$ where $z' \in (\bar{z}^{d\prime}(u', \psi', e^{d\prime} = 1|e^d), \bar{z}^{d\prime}(u', \psi', e^{d\prime} = 0|e^d))$ and for some future states with high productivity levels it is equal to $E(\lambda_\psi \psi' | s_d, e^d)$ where $z' \in (\bar{z}^{d\prime}(u', \psi', e^{d\prime} = 0|e^d), \infty)$. We see that the option value is the expectation of discounted values of period returns and entry costs for different states of nature with different weights.

Figure 3: Content of the Option Values



In Figure 3, because of the sunk costs of entry, the future cutoff productivity levels for entry vary with today's incumbency status. Another point to notice is how option values are bounded by some expected discounted value of entry costs. $E\Delta V(s_d, e^d) < \beta E(\lambda_\psi \psi' | s_d, e^d)$.[23]

$$E\Delta V(s_d, e^d) = \mathcal{O}_\pi(s_d, e^d)\pi^d + \mathcal{O}_\psi(s_d, e^d)\psi_d \tag{15}$$

I define $\mathcal{O}(s_d, e^d) = (1 + \beta\mathcal{O}_\pi(s_d, e^d))/(1 + (1 - e^d)\lambda_\psi - \beta\mathcal{O}_\psi(s_d, e^d))$ and since it is a function firm's state, I assume $\mathcal{O}(s_d, e^d)$ to be a function of firm-specific factors as given in Eq.(16).

$$\mathcal{O}(s_d, e^d) = (1 + e^d)^{\alpha_d}(g_j)^{\alpha_{\mathcal{O}_d}} \exp(\eta_{\mathcal{O}_{djt}}) \tag{16}$$

[23] Derivation of Eq.(15) is given in the Appendix.

Using Eq.(16), Eq.(13), and Eq.(14) we can write an empirical equation for the entry decision

$$D^d_{jt}(s_d, e^d_{t-1}) = \mathbf{1}\left(\underbrace{E(\pi^d_{jt}(X_t) - (1 + (1 - e^d_{t-1})\lambda_\psi)\psi_{dt})}_{\text{Expected net current period profits}} \geq \underbrace{-\beta E(\Delta V(s_d, e^d_{t-1}))}_{\text{Expected discounted option value}}\right) \tag{17}$$
$$= \mathbf{1}\left(E(\mathcal{O}(s_d, e^d_{t-1})\pi^d_{jt}(X_t)) - E(\psi_{djt}) \geq 0\right)$$

The variable D^d_{jt} represents the entry decision into market d for firm j at year t. This decision depends on the firm's expectation about its net gains upon entry. Let's define G^d as the difference between the logarithm of today's expected revenue plus the option value of being in market d today and the logarithm of gross expected sunk entry costs.[24]

$$\mathbf{1}\{E(V^{d,1}(s_d, e^d) - V^{d,0}(s_d, e^d)) \geq 0\} \equiv \mathbf{1}\{E(G_d(s_d, e^d)) \geq 0\} \tag{18}$$

$$G_{dt}(s_{dt}, e^d_{t-1}) = \ln \pi_{dt} + \ln \mathcal{O}(s_{dt}, e^d_{t-1}) - \ln \psi_{dt}$$

D^d_{jt} is a function of observable variables that contribute to the profit in country d, X_{djt}, observable variables contribute to the entry costs, Z_{djt}. But, for notational ease I am suppressing the X_{djt} and Z_{djt} components of D^d_{jt} and the index j.

Extensive Margin

Another aspect of the entry decisions is the possible cost-reducing or demand-increasing effects of being in more countries.[25] For a production technology of constant returns to scale, the cost-reducing effect of size will not be on the costs of production. In Eq.(13) one item in the Z term will be the number of markets that the firm enters during a given period.

I exclude the effect of the outcome of the entry decision and the resulting outcome about a specific market while constructing the Extensive Margin variable for that specific market in order to prevent possible endogeneity issues. In other words, η_d does not take part in the calculation of the Extensive Margin variable used in the estimating equations about market d.

$$D^d = \mathbf{1}\{G_d = X_d\beta + f\alpha_d + \eta_d > 0\}$$
$$Ext.Margin = \sum\nolimits_{s \neq d} \mathbf{1}\{G_s = X_s\beta + f\alpha_s + \eta_s > 0\}$$

In the above equation system, the variable $Ext.\ Margin \in X^d$, and since it does not include information about the entry decision into market d, we abstract from the possible endogeneity

[24]Although the transformation in Eq.(18) causes loss of some information, it allows a discrete choice estimation of the decision process while keeping the signs of the estimated coeffcients correctly.

[25]If a firm's optimal decision for a given period requires it to enter n foreign markets, the Extensive Margin variable to be used in the estimating equation is Extensive Margin=$n - e^d_{jt}$.

problem due to including that variable into the choice equation.[26]

4.1 Factor Structure

Firms have different abilities, strengths and weaknesses. I treat a firm as a set of abilities -such as managerial ability, learning ability, etc.- that contribute to a firm's total productivity, as well as its demand and entry costs shocks. A more able firm may be more successful at reaching new customers and reducing its entry costs by handling the new conditions of different markets more effectively. Also, the monotonicity of the shocks that hit the firms signals the firm-specific abilities. A more able firm receives good shocks today and attracts better shocks in the future as long as it keeps its abilities. Though these abilities are not observed by economists, inference about them is possible. Defining factors as functions of these firm-specific abilities, I can impose some structure on the unobservables of the problem.

The structure that I have assumed for the demand shocks and the shocks to entry costs and productivity allows a factor analysis jointly in outcome (revenue conditional on entry) and choice equations. I also project the measured firm-specific productivities on these firm-specific factors as given in Eq.(19). The logarithms of the demand shock term in Eq.(12) and firm-country specific entry cost shock term in Eq.(14) and option value component of the choice equation are given below

$$z_{jt} = \alpha'_{zt} f_j + \eta_{z_{jt}} \tag{19}$$

$$\begin{aligned} \ln u^d_{j,t} &= \alpha_{dt} \ln g_j + \eta_{djt} \\ &= \alpha_{dt} f_j + \eta_{djt} \end{aligned} \tag{20}$$

$$\ln \xi_{djt} = \alpha'_{\xi_{dt}} f_j + \eta_{\xi_{djt}} \tag{21}$$

$$\ln \mathcal{O}(s_d, e^d_{t-1}) = \alpha'_{\mathcal{O}_{dt}} f_j + \eta_{\mathcal{O}djt} \tag{22}$$

The change between the two lines in Eq.(20) is only for notational purposes. Combining Eq.20-22, we get Eq.23

$$\ln u_{G_{djt}} = \alpha'_{G_{dt}} f_j + \eta_{G_{djt}} \tag{23}$$

where $\alpha_{G_{dt}} = (1 - \sigma^d)\alpha_{zt} + \alpha_{dt} + \alpha_{\xi_{dt}} + \alpha_{\mathcal{O}_{dt}}$. Assuming η_{djt}, $\eta_{\xi_{djt}}$ and $\eta_{\mathcal{O}_{djt}}$ are indepen-

[26] I discuss about some alternative extensions of the model that allows firms to endogenize the returns to being in more markets to some extent in the Appendix .

dently and normally distributed, the weighted sum of these random variables is also distributed normally. Hence, $\eta_{G_{djt}}$ is also normally distributed. In fact, independence of $\eta_{z_{jt}}$,η_{djt}, $\eta_{\xi_{djt}}$ and $\eta_{\mathcal{O}_{djt}}$ will be satisfied structurally as a result of our estimation procedure. All the covariance will be absorbed by the firm-specific factors and time-country specific factor loadings. These loadings allow each unobserved firm-specific factor to be weighted differently across different markets and time, and hence capture the unobserved firm-country-relationship-specific covariance. η's will be the remaining uniqueness terms, which are mutually independent of each other and independent of the factors. Factors are mean zero. Since we have a constant term in the estimating equation the non-zero mean of the factors will be captured by that constant term.

The factors absorb all the co-movement across outcome and choice equation unobservables, and the uniqueness terms represent the orthogonal parts of the unobservables. $cov(\eta) = Diag(\tau^2_{G_{dt}}, \tau^2_{dt}, \tau^2_{zt})$. The covariance across different components of the firm-specific factors will be absorbed by the factor loadings.[27]

For each firm I have three sets of equations: Choice (Eq.(17)), outcome (revenue conditional on entry Eq.(11)), and measurement (measured productivities Eq.(19)). To jointly estimate these equations using the factor structure I follow Bayesian MCMC estimation techniques.[28] Details of the estimation procedures are given in the Appendix.

4.2 Identification

The coefficients of the observed regressors, the unobserved firm-specific factors and their loadings and the distributions of the orthogonal shock (uniqueness) terms are identified.[29]

To understand the role of the unobserved firm-specific attributes and their interaction with different destination markets and the role of shocks orthogonal to the rest, I impose the functional form specifications given in Eq.(24).

Identification of the functional forms for factors and their loadings

The structure that has been put on the unobservable shocks can be justified under the following functional assumption and the following linearization.

[27] For showing $E(ff') = I$, let's assume $y = \alpha f + \eta$ has k dependent common factors where $\eta \sim N(0, T)$ and $f \sim N(0, \Sigma)$. By Cholesky decomposition we can find S where $\Sigma = SS'$. Then, $S^{-1}\Sigma(S')^{-1} = I$, and $\hat{f} = S^{-1}f$ has an independent common factor structure.

[28] Cunha, Heckman and Navarro (2005) and Cunha and Heckman (2007) have followed similar methodologies while explaining the life cycle earning with unobserved individual-specific heterogeneities. Carneiro et al. (2003) studied the returns to schooling.

[29] Identification of these components are discussed in the Appendix. An extensive discussion can be found in Cunha et al. (2005).

$$\begin{aligned} u_{djt} &= H_d(g_j) \cdot \exp(\eta_{djt}) \\ \ln u_{djt} &= \ln H_d(g_j) + \eta_{djt} \\ &= \mathcal{H}_d(g_j) + \eta_{djt} \end{aligned} \tag{24}$$

We can Taylor approximate around the means of the factors, μ_{g_j}.

$$\begin{aligned} \mathcal{H}_d(g_j) = \underbrace{\mathcal{H}_d(\mu_{g_j})}_{C} + \underbrace{(g_j - \mu_{g_j})}_{f_{j1}}\underbrace{\mathcal{H}'_d(\mu_{g_j})}_{\alpha_1} &+ \frac{1}{2!}\underbrace{(g_j - \mu_{g_j})^2}_{f_{j2}}\underbrace{\mathcal{H}''_d(\mu_{g_j}, \mu_{g_d})}_{\alpha_2} \\ &+ \frac{1}{3!}(g_j - \mu_{g_j})^3\mathcal{H}'''_d(\mu_{g_j}) + ... \\ &= \mathbf{C} + \mathbf{f}_j\alpha_d \end{aligned}$$

A similar strategy can be followed for z, ψ, and $\mathcal{O}$.

4.3 Variance Decomposition

If X and Y are random variables and on the same probability space, and variance of Y is finite, then

$$var(Y) = \underbrace{var(E(Y|X))}_{\text{Explained}} + \underbrace{E(var(Y|X))}_{\text{Unexplained}}$$

This is called the law of total variance. In this study I introduce the factor structure onto the unexplained components of the estimating equations, and quantify the sources of unexplained variance as firm specific heterogeneity and idiosyncrasy. If we define the unexplained regression residuals as $B = Y - X\beta$, I impose that

$$B = f\alpha + \eta$$

where $f \sim N(0, I), \eta \sim N(0, \Sigma_H)$, and $\Sigma_H = Diag(\tau_1^2 ... \tau_k^2)$.

Under this model, the variance-covariance is constrained by the form $var(B|\alpha, \Sigma_H) = \alpha\alpha' + \Sigma_H$, and for any element B_i in B $(i) var(B_i|f, \alpha, \Sigma_H) = \tau_i^2$, $(ii) var(B_i|\alpha, \Sigma_H) = \Sigma_{m=1}^k \alpha_{im}^2 + \tau_i^2$, $(iii) cov(B_i, B_j|f) = 0$, $(iv) cov(B_i, B_j|\alpha, \Sigma_H) = \Sigma_{m=1}^k \alpha_{im}\alpha_{jm}$.

5 Results

In the test of whether the production technologies in the selected industries were CRS, the p values were 0.0004, 0.6295, 0.0000,0.0000,0.003 for Food, Apparel, Metals, Machinery, and Automotive industries, respectively. Hence, we can say that production follows constant returns to scale technologies in our explored sectors. Except for the Apparel industry, these findings correspond to our prior assumption about CRS production technologies across industries.

The main parameter estimates of the selection problem are given in Tables 7 and 8 . The real exchange rate variable is defined to be the worth of Turkish Lira (TL) against other currencies.[30] In all sectors, increasing exchange rates obstruct entry into the foreign markets. In the Automotive industry this has a positive, but insignificant impact in the model with factors. From the theoretical section, we know that the revenues conditional on entry into a country are determined by the production costs and the elasticity of demand for a good in that market. Our results show that revenues increase with increasing exchange rates in all sectors. The reason behind this response of revenues may signal that some other forces are at work. If the demand from a specific country for the products of a specific sector is constant, then the decrease in the number of entrant firms means division of the constant demand among fewer firms, and hence these firms sell more in that market. The elasticity of revenue conditional on entry with respect to the exchange rate is the lowest at 0.07 in the Machinery industry, while the highest at 0.17 in the Apparel industry.

In Food and Metals sectors, lnc (logarithm of the constructed unit production cost) hinders entry into foreign markets. However, in Apparel and Automotive it promotes entry. Maybe in these sectors I'm capturing a quality effect that if the quality of the produced goods in these sectors are increasing over the period, and if higher quality goods are more costly to produce, then increasing unit costs of production may ease entry. On the revenue side, unit cost of production reduces revenue upon entry in all sectors. The response of revenue to the unit cost of goods produced is the highest in Automotive sector with an elasticity of -0.762, and the lowest in Food industry with an elasticity -0.038.

As expected, market potential tempts entry into a market in all sectors, and revenue upon entry increases with the potential of the market. It also seems that there has been a positive correlation between market potential and unobserved shocks that were not accounted for in the no-factors setup. With the introduction of firm-specific heterogeneity, the elasticity of revenue with respect to the market potentials goes down from 0.237 to 0.043, from 0.289 to 0.109, from 0.057 to 0.047, and 0.213 to 0.067 in Food, Apparel, Metals, and Automotive sectors, respectively. These positive and significant numbers indicate that as a destination market's size grows, firm revenues in that market grow, as well. These findings also indicate a ranking of competitiveness of Turkish firms in their destination markets. When the size of the market

[30]The values for the real exchange rate variable varies over time and countries.

doubles in one of the destination markets, revenues of an Apparel firm is expected to grow 10% in that market. However, the same increase in the market size causes only a 4% revenue increase in Food industry. These numbers suggest that Turkish firms in Apparel industry are relatively more competitive or have more market power in their destination markets than Turkish firms in Food industry.[31]

Lag presence in a market is a strong indicator re-entering the same market again during the current period in all industries. As shown in the previous sections, sunk costs of selling goods in a market (e.g, learning about distribution channels, packaging, and legal procedures) may be a reason.[32] When the firm-specific factors are introduced, in all sectors, the parameter estimates for the effects of lag presence in a market have slightly gone up. This situation indicates that at least one of the firm-specific factors was negatively correlated with this variable. If the recovered factors are mainly the unobserved firm-specific abilities, and if more able firms are more successful at handling the sunk costs so that they pay less sunk costs, then this increase in the parameter estimates is quite plausible. It's also clear from the parameter estimates that as the distance between the origin country and the destination market grows, entry becomes harder.

Measuring the synergetic effects of being in more markets -namely the Ext. Margin variable- is another interest of this paper We see that the positive demand effect of being in one more market is an increase in demand between 1% and 3% across different sectors. We can also talk about some entry-cost-reducing synergetic effect of being in more markets as we see that coefficients in the choice equations are higher than those for their outcome equation counterparts.[33] The endogeneity problem that was expected to occur in the case of no factors becomes apparent in these results, as well. Since the unobserved demand shocks and the impact of Ext. Margin are expected to be positively correlated, endogeneity is expected to bias the coefficient estimates for Ext. Margin upwards in the no-factors case. In the choice equation entry cost shocks are expected to be positively correlated with the Ext. Margin variable and since entry cost components enter into the choice equation with a negative sign in front of them, endogeneity is expected to cause a compound upward bias in the choice equation.[34] Tariff rates impede entry in Food, Textile, and Automotive industries, whereas the coefficient estimates are insignificant

[31] In Machinery industry the impact of market size on firm revenues in the destination markets are inconclusive. This result may indicate that Turkish machine producers have expertise in only certain kinds of machines, and changes in destination market sizes may be stemming from the increases in the demand for other types of machines that Turkish firms do not produce.

[32] Building customer lists in destination markets over time as it is studied in Drozd and Nosal (2012) may be another reason for this effect.

[33] Since we have employed a nonlinear model for estimating the selection behavior it is not straightforward to talk about how much a specific a variable increases the probability of selection. However, since we have normalized the latent returns to selection G in the estimation steps to be in between [-30,30], we can say that there are some entry-costs-reducing returns to Ext. Margin. Because of the nonlinear nature of the estimation we can not purge the effect on the demand from the estimated choice equation parameters. Hence we can't report some pure entry-cost-reducing impacts of Ext. Margin.

[34] The demand-increasing and the entry-cost-reducing impacts build up this compound bias.

in Metals and Machinery sectors. The reason behind this insignificance may be the imputation for missing tariff data, which is discussed in the data Appendix.

The factor loadings measure the effect of the interaction between firm-specific heterogeneity and the destination countries over time. If I put a restriction on one of the factors by fixing the value of that factor to 1 at the estimation step for each observation, then the factor loadings of that factor start capturing time-country-specific fixed effects. As a robustness check, I estimate the model by including these fixed effects, and in the Online Appendix, I present the results for the model, introducing the restriction for time-country-specific fixed effects. Inclusion of these fixed effects seemingly causes a multicollinearity problem among the regressors. We observe large changes in the estimated coefficients for some of the variables, especially the ones that vary only across time and country. For instance, the parameter estimate for tariff rates has jumped from -0.066 to 10.046 for the Automotive industry by the inclusion of country-time specific effects. Standard errors of the possibly affected coefficients (e.g., the ones for tariff, exchange rate) tended to be larger. Also, this inclusion may be causing an overfitting problem, as can be observed in Table 9. The fraction of explained entries increases drastically. However, it happens at the cost of biased estimates, and it will have a poor predictive power with some other data as it can exaggerate the effect of minor fluctuations in the data.

In Figure 4, I present the variances of the idiosyncratic terms for revenues conditional on entry for different sectors over destination countries and time. The model without firm-specific factors estimates idiosyncracies to be within the range of 30 to 80. Remembering that these idiosyncratic terms affect some equations in logarithms, we see that a big heterogeneity component in the model with no factors is hidden within the unexplained idiosyncratic component, and the size of the unobserved components is very large. When we add the firm-specific factors, the range of idiosyncratic terms goes down to the 0.2-1.5 interval.

For visualizing the difference between the two estimation strategies, Eq.25 may be helpful. When we do not control for the firm-specific unobserved heterogeneity, we are treating $\alpha f + \eta$ as the error term, whereas when we control for them, η is the error term that is expected to be orthogonal to the independent regressors. For comparing the size of the error terms in both models, let us take two firms with completely identical observed regressors. Then, the model without factors shows that in three standard deviations around the mean error shock, while one of these firms receives one unit error shock, the other can receive one billion fold of the other's shock. But after accounting for the unobserved heterogeneities, the same exercise gives the relative impact to be 1:20.

$$ln(R) = X\theta + \alpha f + \eta \tag{25}$$

In Figure 5, I present the estimated distribution of TFP's and factor score for the Food industry, which are jointly estimated with firms' entry decision equations and revenue condi-

Table 7: Estimates for the Choice Equation Parameters

No factor vs.Firm-Specific Factors

	Foods		Apparel		Metals		Machinery		Automotive	
	NF	F	NF	F	NF	F	NF	F	NF	F
Constant	−3.018*	−3.428*	−3.767*	−3.013*	−1.209*	−1.211*	−0.804*	−0.825*	−1.972*	−1.579*
	(0.001)	(0.273)	(0.001)	(0.202)	(0.001)	(0.161)	(0.001)	(0.129)	(0.001)	(0.187)
Exc. Rate	−0.061*	−0.087*	−0.019*	−0.107*	−0.025*	−0.022*	0.002*	−0.007	0.015*	0.015
	(0.000)	(0.014)	(0.000)	(0.022)	(0.000)	(0.008)	(0.000)	(0.006)	(0.000)	(0.009)
ln(c)	−0.036*	−0.057*	−0.003*	0.051*	−0.030*	−0.025*	0.003*	0.006	0.062*	0.095*
	(0.000)	(0.008)	(0.000)	(0.010)	(0.000)	(0.010)	(0.000)	(0.008)	(0.000)	(0.012)
Market Pot.	0.055*	0.031*	0.170*	0.156*	0.043*	0.046*	0.044*	0.027*	0.161*	0.138*
	(0.000)	(0.008)	(0.000)	(0.008)	(0.000)	(0.007)	(0.000)	(0.005)	(0.000)	(0.008)
Lag Presence	2.050*	2.114*	1.671*	1.703*	1.853*	1.940*	1.745*	1.775*	2.116*	2.156*
	(0.000)	(0.024)	(0.000)	(0.018)	(0.000)	(0.024)	(0.000)	(0.017)	(0.000)	(0.024)
Distance	−0.053*	−0.009	−0.116*	−0.154*	−0.251*	−0.263*	−0.254*	−0.224*	−0.253*	−0.219*
	(0.000)	(0.036)	(0.000)	(0.028)	(0.000)	(0.021)	(0.000)	(0.016)	(0.000)	(0.024)
Ext. Margin	0.140*	0.118*	0.186*	0.166*	0.126*	0.105*	0.118*	0.112*	0.115*	0.111*
	(0.000)	(0.004)	(0.000)	(0.004)	(0.000)	(0.003)	(0.000)	(0.002)	(0.000)	(0.003)
Tariff	−0.032*	−0.047*	−0.067*	−0.122*	0.034*	0.027*	0.009*	0.014	−0.011*	−0.066*
	(0.000)	(0.018)	(0.000)	(0.021)	(0.000)	(0.015)	(0.000)	(0.013)	(0.000)	(0.021)

Notes: This table shows the results for the joint estimation of Eq.(11), (17) and (19). Standard errors are in parentheses. All variables are in logs. * and ** means statistically significant at the 5% and 1% levels. Exc. Rate is the real exchange rate, lnc is a constructed measure of firm-specific unit cost of production. Lag presence in a market can be either 0 or 1. Distance is in kilometers. Ext. Margin stands for the number of markets penetrated during a given year. Refer to Appendix C for the details about the variable construction.

Table 8: Estimates for the Revenue Equation Parameters

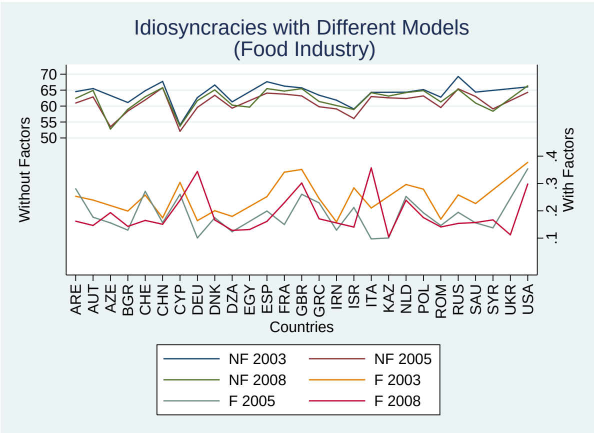
No factor vs.Firm-Specific Factors

	Foods		Apparel		Metals		Machinery		Automotive	
	NF	F	NF	F	NF	F	NF	F	NF	F
Constant	8.126*	9.894*	4.699*	7.629*	6.230*	6.467*	5.771*	8.078*	−2.569*	1.558*
	(0.007)	(0.083)	(0.009)	(0.104)	(0.008)	(0.105)	(0.005)	(0.077)	(0.007)	(0.129)
Exc. Rate	0.052*	0.096*	0.155*	0.174*	0.069*	0.090*	0.053*	0.070*	0.055*	0.102*
	(0.000)	(0.004)	(0.001)	(0.005)	(0.000)	(0.004)	(0.000)	(0.002)	(0.000)	(0.003)
ln(c)	0.033*	−0.038*	−0.195*	−0.148*	−0.334*	−0.343*	−0.301*	−0.205*	−0.934*	−0.762*
	(0.001)	(0.006)	(0.001)	(0.008)	(0.001)	(0.009)	(0.000)	(0.006)	(0.001)	(0.011)
Market Pot.	0.237*	0.043*	0.289*	0.109*	0.057*	0.047*	0.068*	−0.003	0.213*	0.067*
	(0.000)	(0.005)	(0.000)	(0.005)	(0.000)	(0.005)	(0.000)	(0.003)	(0.000)	(0.004)
Ext. Margin	0.115*	0.029*	0.079*	0.007*	0.070*	0.013*	0.101*	0.008*	0.092*	0.010*
	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)	(0.001)	(0.000)	(0.002)

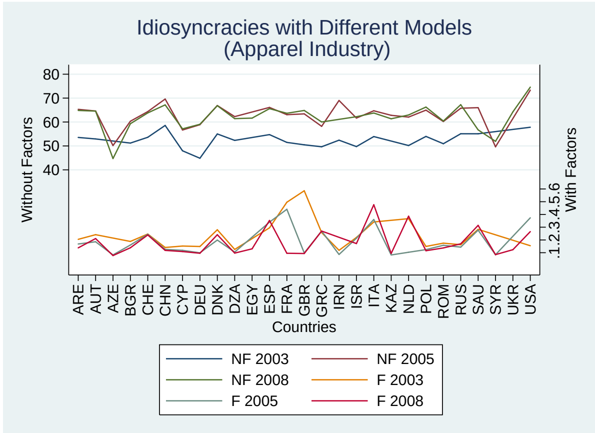
Notes: This table shows the results for the joint estimation of Eq.(11), (17) and (19). Standard errors are in parentheses. All variables are in logs. * and ** means statistically significant at the 5% and 1% levels. Exc. Rate is the real exchange rate, lnc is a constructed measure of firm-specific unit cost of production. Lag presence in a market can be either 0 or 1. Distance is in kilometers. Ext. Margin stands for the number of markets penetrated during a given year. Refer to Appendix C for the details about the variable construction.

Figure 4: Idiosyncracies for the Revenues

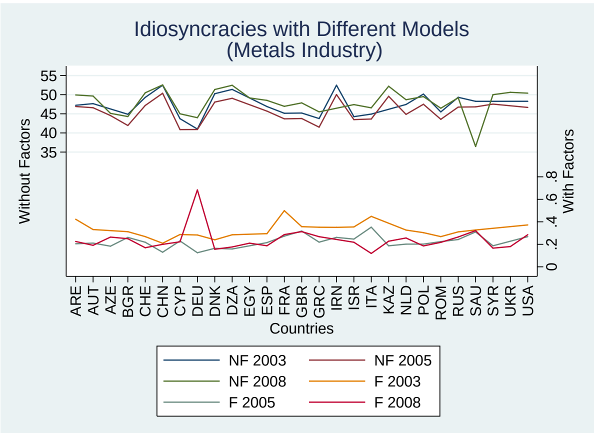
(a) Food Industry



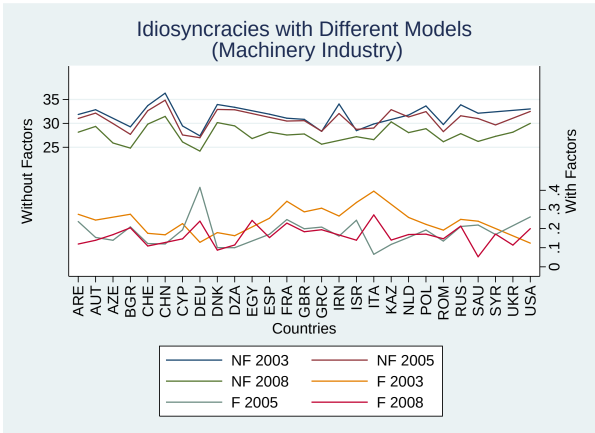
(b) Apparel Industry



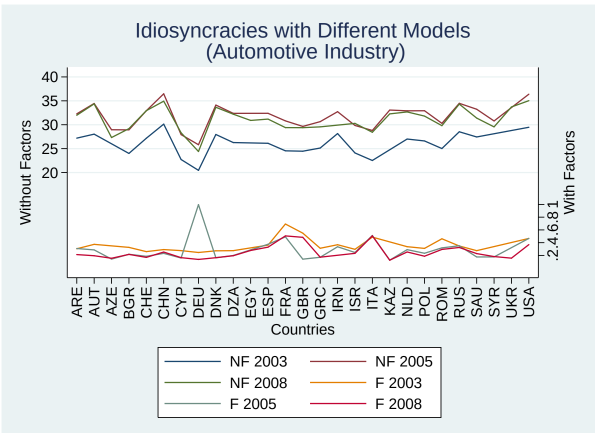
(c) Metals Industry



(d) Machinery Industry



(e) Automotive Industry



tional on entry equations.[35] A good fit of the actual TFP's and their factor score estimates is observed for 2003 and 2004 and the goodness of fit decays over time in the graphs. The main reason behind this is my estimation strategy. I have kept firm-specific factors constant over time to be consistent with the theoretical analysis in which I assume firm-specific abilities are constant over time. Hsieh and Klenow (2009) find the TFP ratio of a 90th percentile firm and a 10th percentile firm to be over 5:1 for China and India, while Syverson (2004) finds it to be 1.92 in the U.S manufacturing sector. In this study, I find the same ratio to be similar to the findings in Hsieh and Klenow (2009) for Turkish manufacturing firms.

Goodness of fit of revenues for the model with no firm-specific factors and with factors is shown in Figure 6.[36] With the inclusion of firm-specific factors, fit of the revenue estimates improves in all sectors. Table 9 shows the fraction of entries correctly captured by different models. The first column is for the model with no factors, the second column is for the models with only firm-specific factors, and the third one is for the model with both firm-specific factors and time-country fixed effects. Since the semiparametric estimation of the choice equation draws latent returns, $G's$, from truncated normal distributions, we do not observe very much differentiation in the correctly captured entry behaviors.[37][38]

Eaton, Kortum and Kramarz (2011) assume that the variation across firms in entry stems from the variation in productivities, which applies across all markets, and, hence universal, entry cost and sales shocks, which are idiosyncratic to firm market interactions. They search for the extent of variation explained by the universal rather than the idiosyncratic components. I also define similar sources for the shocks; however, I allow productivity shocks to be correlated with the demand and entry cost shocks. They assume entry cost and sales shocks to come from the same distributions for all firms across different markets. In contrast, I allow these shocks to be functions of some unobserved firm-specific attributes, which also contributes to the firm-level productivies. Then the remaining unexplained variation is from the idiosyncratic shocks at firm-market-time level and orthogonal to the unobserved attributes.

They find that 57% of the variation in entry into a market can be attributed to the core efficiency -productivity- of the firm, rather than to its draw of demand and entry cost shocks in the market. While only 4.8% of the variation in sales can be attributed purely to the efficiency, in total, 39% can be attributed to the efficiency and the other two shocks. The setup of this paper allows a decomposition of the variation in sales in between unobserved heterogeneity

[35] Similar figures for the other industries are presented in the Online Appendix.

[36] I get the best fit of the model with the data for all industries expect Automotive in a model that uses three independent factors. In Automotive industry the best fit comes with two firm-specific factors. However, the highest fraction of firm-specific heterogeneity is almost always explained by only one dimension of the heterogeneity.

[37] Introduction of the fixed effects increases the fraction of entries captured correctly in all sectors except the Metals industry. However, as mentioned before, this method is plagued with the multicollinearity problem, and most probably what is taking place is an overfitting issue.

[38] In the data, 95% of the observed choice decisions are no entry, and only the remaining 5% is entry.

and the idiosyncratic shocks. The results of this decomposition given in Figure 7 show that the share of idiosyncratic variance in total variation is not homogeneous across industries and countries. While, on average, the share of idiosyncratic components is higher in the Apparel and Food industries, it is, on average, lower in the Machinery and Automotive industries. Firm-specific heterogeneity explains more of the total residual variation in revenues in foreign markets as opposed to idiosyncratic variation in technology intensive industries than less technology intensive ones. One other finding is that the relative importance of the idiosyncratic components diminishes as the per capita income of the destination market increases.

The findings of the variance decomposition may also be helpful in understanding the returns to the policies that promote entry into the export markets. If policymakers' information sets about the firms are not larger than ours, we can have some policy implications for subsidizing the entry costs. In the less technology intensive industries like Food and Apparel and in lower per-capita income destinations, subsidizing the entry costs may be plausible because much of the unexplained sales variation will come from the idiosyncratic shocks, and the relative impact of the unobserved and hence uncontrolled heterogeneity will be low. However, in more technology intensive industries like Machinery and Automotive and in the richer destinations, the unobserved firm attributes make up for the biggest share of sales variation. Hence, in these industries and markets, instead of subsidizing the entry costs, subsidizing the firm-specific factors, skills and attributes, and skill formation may be a better option.

6 Conclusion

In this paper I provide a theoretical analysis of the dynamics of firms' foreign market entry decisions. I show the direction of firms' responses to different types of shocks and build a clear mapping between theoretical equations and an empirical analysis. I treat firms as sets of unobservable abilities so that I can deal with the endogeneity and misspecification problems by using nonparametric estimates of these unobservables. I exploit the covariance structure among choice (entry decision), outcome (revenue conditional on entry), and measurement (productivity estimates) equations for identifying these unobservables. My results have shown that being in an additional foreign market increases revenue conditional on entry into another market 1%-3%. It also reduces the entry costs for other markets so that eases the firm selection into new markets. Previous experience in a specific market has a very important impact on the choice decision about the same market today. These benefits to previous experience in a market and being in more markets can provide a rationale for the existing export subsidies by showing why a firm's status in a market can persist after initial entry into that market.

One other exciting outcome of the study is its ability to provide a variance decomposition of firms' revenues in different markets. The share of idiosyncratic variance in total variation is differs across both industries and countries. Firm-specific heterogeneity explains more of the total residual variation in revenues as opposed to idiosyncratic variation in technology intensive industries like Machinery and Automotive than less technology intensive industries like Apparel and Food. Another finding is that the relative importance of idiosyncratic components diminishes as the per capita income level of the destination market increases.

References

[1] Aguilar, Omar, and Mike West. 2000. Bayesian Dynamic Factor Models and Portfolio Allocation. Journal of Business & Economic Statistics 18, no. 3 (July): 338-357.

[2] Albornoz, Facundo, Héctor F. Calvo Pardo, Gregory Corcos, and Emanuel Ornelas. 2011. Sequential Exporting. mimeo.

[3] Arkolakis, Costas. 2011. A Unified Theory of Firm Selection and Growth.

[4] Arkolakis, Costas, and Marc-Andreas Muendler. 2010. The Extensive Margin of Exporting Products: A Firm-level Analysis. National Bureau of Economic Research Working Paper Series No. 16641 (December).

[5] Asplund, Marcus, and Volker Nocke. 2006. Firm Turnover in Imperfectly Competitive Markets. The Review of Economic Studies 73, no. 2 (April): 295-327.

[6] Bernard, Andrew B., Eaton, Jonathan, Jensen, J. Bradford and Samuel Kortum. 2003. Plants and Productivity in International Trade. American Economic Review 93, no.4: 1268-1290

[7] Bernard, Andrew B., and J. Bradford Jensen. 2004. Why Some Firms Export. Review of Economics and Statistics 86, no. 2 (May): 561-569.

[8] Campa, José Manuel. 2004. Exchange rates and trade: How important is hysteresis in trade? European Economic Review 48, no. 3 (June): 527-548.

[9] Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman. 2003. 2001 Lawrence R. Klein Lecture: Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. International Economic Review 44, no. 2 (May): 361-422.

[10] Chaney, Thomas. 2011. The Network Structure of International Trade. National Bureau of Economic Research Working Paper Series No. 16753 (January).

[11] Chib, Siddhartha, and Barton H. Hamilton. 2000. Bayesian analysis of cross-section and clustered data treatment models. Journal of Econometrics 97, no. 1 (July): 25-50.

[12] ———. 2002. Semiparametric Bayes analysis of longitudinal data treatment models. Journal of Econometrics 110, no. 1 (September): 67-89.

[13] Cunha, Flavio, and James J. Heckman. 2007. The Evolution of Inequality, Heterogeneity and Uncertainty in Labor Earnings in the U.S. Economy. National Bureau of Economic Research Working Paper Series No. 13526 (October).

[14] Cunha, Flavio, Heckman, James and Navarro Salvador. 2005. Separating uncertainty from heterogeneity in life cycle earnings. Oxford Economic Papers 57, no. 2 (April): 191.

[15] Das, Sanghamitra, Mark J. Roberts, and James R. Tybout. 2007. Market Entry Costs, Producer Heterogeneity, and Export Dynamics. Econometrica 75, no. 3: 837-873.

[16] De Loecker, Jan. 2007. Do exports generate higher productivity? Evidence from Slovenia. Journal of International Economics 73, no. 1 (September): 69-98.

[17] De Loecker, Jan. 2011. Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity. Econometrica 79, no.5 (September): 1407-1451.

[18] Drozd, Lukasz A., and Jaromir B. Nosal. 2012. Understanding International Prices: Customers as Capital. American Economic Review 102, no.1 (February): 364-395.

[19] Eaton, Jonathan, Marcela Eslava, Maurice Kugler, and James Tybout. 2007. Export Dynamics in Colombia: Firm-Level Evidence. National Bureau of Economic Research Working Paper Series No. 13531 (October).

[20] Eaton, Jonathan and Samuel Kortum. 2002. Technology, Geography, and Trade. Econometrica. 70, no. 5: 1741-1779.

[21] Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2011. An Anatomy of International Trade: Evidence from French Firms. Econometrica 79, no.5 (Septemeber): 1453-1498.

[22] Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? American Economic Review 98, no. 1 (March): 394-425.

[23] Geweke, John, and Guofu Zhou. 1996. Measuring the price of the Arbitrage Pricing Theory. The Review of Financial Studies 9, no. 2 (Summer): 557-587.

[24] Helpman, Elhanan, Marc Melitz, and Yona Rubinstein. 2008. Estimating Trade Flows: Trading Partners and Trading Volumes. Quarterly Journal of Economics 123, no. 2 (May): 441-487.

[25] Hsieh, Chang-Tai, Klenow, Peter, J. 2009. Misallocation and Manufacturing TFP in China and India. Quarterly Journal of Economics 124, no. 4 (November): 1403-1448

[26] Kee, Hiau Looi, and Kala Krishna. 2008. Firm-Level Heterogeneous Productivity and Demand Shocks: Evidence from Bangladesh. American Economic Review 98, no. 2 (May): 457-462.

[27] Kotlarski, Ignacy. 1967. On Characterizing the Gamma and the Normal Distribution. Pacific Journal of Mathematics 20, no. 1: 69-76.

[28] Ledermann, Walter. 1937. On the rank of the reduced correlational matrix in multiple-factor analysis. Psychometrika 2, no. 2 (June 27): 85-93. doi:10.1007/BF02288062.

[29] Levinsohn, James, and Amil Petrin. 2003. Estimating Production Functions Using Inputs to Control for Unobservables. The Review of Economic Studies 70, no. 2 (April): 317-341.

[30] Lopes, Hedibert Freitas. 2000. Bayesian analysis in latent factor and longitudinal models. United States – North Carolina: Duke University.

[31] Melitz, Marc J. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica 71, no. 6 (November): 1695-1725.

[32] Morales, Eduardo, Sheu, Gloria and Andres Zahler. 2010. Gravity and Extended Gravity: Estimating a Structural Model of Export Entry. (mimeo).

[33] Neyman, J., and Elizabeth L. Scott. 1948. Consistent Estimates Based on Partially Consistent Observations. Econometrica 16, no. 1 (January): 1-32.

[34] Pedersen, Niels Kjaergaard. 2009. Essays in International Trade. Northwestern University, June.

[35] Roberts, Mark J., and James R. Tybout. 1997. The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs. The American Economic Review 87, no. 4 (September): 545-564.

[36] Rossi, Peter, Allenby, Greg, and Robert Mcculloch. 2005. Bayesian Statistics and Marketing. Wiley Series in Probability and Statistics.

[37] Sandikci, Burhaneddin, Lisa M. Maillart, Andrew J. Schaefer, Oguzhan Alagoz, and Mark S. Roberts. 2008. Estimating the Patient's Price of Privacy in Liver Transplantation. Operations Research 56, no. 6: 1393-1410.

[38] Stokey, Nancy L., Robert E. Lucas Jr, and Edward C. Prescott. 1996. Recursive Methods in Economic Dynamics. Harvard University Press.

[39] Syverson, Chad. 2004. Product Substitutability and Productivity Dispersion. The Review of Economics and Statistics 86, no.2 (May): 534-550

[40] Van Biesebroeck, Johannes. 2005. Exporting raises productivity in sub-Saharan African manufacturing firms. Journal of International Economics 67, no. 2 (December): 373-391.

A Proofs and Derivations

A.1 Consumer Demand

Let N countries populate the world economy, and in each country, there exist $\mathbf{M}$ sectors in with a continuum of firms. In each country a continuum of consumers lives and each consumer maximizes his lifetime utility. Consumers gain utility from consuming a composite of goods. Then the utility function given in Eq.(26) can justify the functional forms for the demand used in the paper. In country d, consumer i's problem is:

$$\max U_{id} = x_{idt}(0)^{(1-\mu)} \sum_{t=0}^{\infty} \left(\sum_{m=1}^{\mathbf{M}} \left(\int_J x_{idt}(j,m)^{\frac{\sigma^d-1}{\sigma^d}} dj \right)^{\frac{\sigma^d}{\sigma^d-1}\frac{\varrho^d-1}{\varrho^d}} \right)^{\mu\frac{\varrho^d}{\varrho^d-1}} \tag{26}$$

$$s.t. \quad \sum_{m=1}^{\mathbf{M}} \left(\int_0^1 p_{dt}(j,m) x_{idt}(j,m) dj \right) \leq \tilde{W}_{idt} \text{ for all } t$$

where $\varrho^d > 1, \sigma^d > 1$ for all d. $\tilde{W}_{idt}$ is consumer i's disposable income at time t, and $W_{idt} = \mu \tilde{W}_{idt}$. $x_{idt}(j,m)$ is consumer i's demand for good j of sector m at time t, and $p_{dt}(j,m)$ is the price of the corresponding good. $x_{idt}(0)$ is a consumer's demand for the numeraire good in the economy. This numeraire good is produced in each country with the same technology and traded with no cost. All payments are in terms of this numeraire good in the economy. The consumer's first order condition gives rise to the following nested CES demand for goods. Since i is the representative consumer, I have suppressed the index i in the expressions.[39]

$$x_{dt}(j,m) = X_m^d(t) \left(\frac{P_m^d(t)}{p_{dt}(j,m)} \right)^{\sigma^d} \tag{27}$$

$$x_{dt}(j,m) = W_{dt} \frac{P_m^d(t)^{\sigma^d-1}}{p_{dt}(j,m)^{\sigma^d}} \tag{28}$$

where $P^d(t)$ is the price index for country d at time t.

$$X_m^d(t) = \left(\int_0^1 x_{dt}(j,m)^{\frac{\sigma^d-1}{\sigma^d}} dj \right)^{\frac{\sigma^d}{\sigma^d-1}} \text{ and } P_m^d(t) = \left(\int_0^1 p_{dt}(j,m)^{1-\sigma^d} d \right)^{\frac{1}{1-\sigma^d}}$$

$$X^d(t) = \left(\sum_{m=1}^{\mathbf{M}} X_m^d(t)^{\frac{\varrho^d-1}{\varrho^d}} \right)^{\frac{\varrho^d}{\varrho^d-1}} \text{ and } P^d(t) = \left(\sum_{m=1}^{\mathbf{M}} P_m^d(t)^{1-\varrho^d} \right)^{\frac{1}{1-\varrho^d}}$$

If E_m^d is total expenditure on the available goods of sector m in country d, and P_m^d is the ideal price index for those goods, then the multiplicative demand shock u_{djt} in the demand equation represents all the shocks that hit these entities.

A.2 About Existence of a Markov-Perfect (Recursive) Equilibrium

Assumption 1:

$$\mathcal{P}_d(z_j', u_{dj}', \psi_{dj}' | z_j, u_{dj}, \psi_{dj}) = \mathcal{Z}(z_j'|z_j) \mathcal{U}_d(u_{dj}'|u_{dj}) \mathcal{S}_d(\psi_{dj}'|\psi_{dj})$$

Assumption 1 introduces the conditional independence of firm-specific productivity shocks and firm-country-specific demand and entry cost shocks. In the empirical section the assumed

[39] Arkolakis and Muendler (2010) also use a similar demand structure.

factor structure will result in such a conditional independence across shocks.[40]

Assumption 2: $\mathcal{Z}(z'_j|z_j)$ is monotone such that $\mathcal{Z}(z'_j|z_j+\varepsilon) \geq \mathcal{Z}(z'_j|z_j)$ where $\varepsilon > 0$, and $z'_j > z_j + \varepsilon$. Similarly, $\mathcal{U}_d(u'_{dj}|u_{dj})$ is also monotone.

Assumption 3: $\mathcal{S}_d(\psi'_{dj}|\psi_{dj})$ is monotone such that $\mathcal{S}_d(\psi'_{dj}|\psi_{dj}-\varepsilon) \geq \mathcal{S}_d(\psi'_{dj}|\psi_{dj})$where $\varepsilon > 0$, and $\psi'_{dj} < \psi_{dj} - \varepsilon$.

From the assumed demand form and market structure it is evident that, conditional on entry, $\frac{\partial \pi_j^d}{\partial z_j} > 0$, and $\frac{\partial \pi_j^d}{\partial u_{dj}} > 0$. $\pi_j^d(z_j, u_{dj})$ is continuous in both z_j and u_{dj}.

Assumption 4: $\lim_{t\to\infty} \beta^t \int V^d(s_{dt}, e^d_{t-1}(s_{dt-1}))\mu(s_{d0}, ds_{dt}) = 0$ where $\mu(s_{d0}, ds_{dt})$ is the probability measure of the state being s_{dt} if s_{d0} is the initial state, $e \in \{0,1\}$, and all $(s_{d0}, e^d_0) = \hat{s}_{d0} \in \hat{S}^d$.

For the empirical analysis section Normal distributions will be assumed for the idiosyncratic components of the shocks. Normal distribution has moments of all orders, and for a Normal distributed variable-with mean μ and variance σ^2- the moment of order ρ exists and is finite for all ρ such that $\rho > -1$.

Every payment is made in terms of the numeraire good.[41] Under assumptions 1-4, the existence of V and an optimal policy function has been shown in Stokey et al. (1996). It is straightforward to show that for any $s \in S$, $V^{d,0}(s_d,1) \geq V^{d,0}(s_d,0)$ and $V^{d,1}(s_d,1) \geq V^{d,1}(s_d,0)$, $\int V^d(s'_d,1)\mathcal{P}_d(s'_d|s_d)ds'_d \geq \int V^d(s'_d,0)\mathcal{P}_d(s'_d|s_d)ds'_d$, and $V^d(s_d,1) \geq V^d(s_d,0)$. It is also straightforward to show that $V^d(s_d,e^d)$ is increasing in z and u_d and decreasing in ψ_d.

Proposition 2. *There exists a z_{jt}-based optimal control-limit policy for all $u_{djt} \in U$, $\psi_{djt} \in \Psi$, and $e^d_{jt-1} \in \{0,1\}$. Similarly, there exists a u_{djt}-based control-limit policy for all $z_{jt} \in Z$, $\psi_{djt} \in \Psi$, and $e^d_{jt-1} \in \{0,1\}$, and a ψ_{djt} based control-limit policy for all $z_{jt} \in Z$, $u_{djt} \in U$, and $e^d_{jt-1} \in \{0,1\}$.*

Using the established results in the literature we see that under Assumptions 1 and 2, $V^d(s_d,e^d)$ is increasing in z for all $\psi_d \in \Psi$, $u_d \in U$ and $e^d_{jt-1} \in \{0,1\}$. The proof in Stokey et al. (1996) requires boundedness of the period return function. In this setting, I assume the support of shocks is infinite. However, since we will also impose these shocks to be distributed Gaussian, and the shocks are multiplicative, we know that the expected discounted values and expected current period returns will be bounded. Similarly, $V^d(s_d,e^d)$ is increasing in u_d for all $z \in Z$, $\psi_d \in \Psi$ and $e^d \in \{0,1\}$.

Lemma 3. $E\Delta V^d(s_d,e^d) = E(V^d(s'_d,1) - V^d(s'_d,0)|s_d,e^d)$ *(Option Value) is nondecreasing in u_{dj} for given z_j, ψ_{dj} and e^d_{jt-1}, is nondecreasing in z_j for given u_{dj}, ψ_{dj} and e^d is nondecreasing in ψ_{dj} for given u_{dj}, z_j and e^d.*

Proof. (Lemma 3) In equilibrium $\mathcal{P}_d(s'_d, e^{d\prime}|s_d,e^d) = \mathcal{P}_d(s'_d|s_d,e^d) = \mathcal{P}_d(s'_d|s_d)$

[40] z_{jt} has all the firm-specific information and hence its transition is independent of the transition of firm-destination specific demand and entry cost shocks. Also, given u_{djt} and ψ_{djt}, entry costs and demand shocks follow independent Markov processes. This assumption is not necessary for the existence of the value and policy functions. However, it brings some expositional convenience and is in accordance with the following empirical analysis.

[41] If we allow a constant efficiency growth rate, γ_0, for the production of the numeraire good, then we have to impose a time discount factor $\beta > \gamma_0$ to keep the model identified.

$E\Delta V(s_d, e^d) = E(V^d(s'_d, e^{d\prime} = 1) - V^d(s'_d, e^{d\prime} = 0)|s_d, e^d)$ (Option Value)

$$\begin{aligned} V^d(s_d, e^d = 1) &= \max\{V^{d,0}(s_d, e^d = 1), V^{d,1}(s_d, e^d = 1)\} \\ V^d(s_d, e^d = 0) &= \max\{V^{d,0}(s_d, e^d = 0), V^{d,1}(s_d, e^d = 0)\} \\ &= \max\{V^{d,0}(s_d, e^d = 1), V^{d,1}(s_d, e^d = 1) - \lambda_\psi \psi_{dj}\} \end{aligned}$$

To prove this lemma we will follow a backward induction algorithm. First, we'll show nondecreasingness in u_d for $E\Delta V(s_d, e^d)$. Let's assume $E\Delta V(s'_d, e^{d\prime}) = \int V^d(s''_d, e^{d\prime\prime} = 1)\mathcal{P}_d(s''_d|s'_d, e^{d\prime})ds' - \int V^d(s''_d, e^{d\prime\prime} = 0)\mathcal{P}_d(s''_d|s'_d)ds''_d$ is nondecreasing in u_d.

If $s'_d \in S$ is such that firm opts to be out of the market even if it was in the market in the previous period, then $V^d(s'_d, e^{d\prime} = 1) = V^{d,0}(s'_d, e^{d\prime} = 1)$. Then we have $V^d(s'_d, e^{d\prime} = 1) = V^d(s'_d, e^{d\prime} = 0) = V^{d,0}(s'_d, e^{d\prime} = 1)$, and for these future states the value of being an incumbent in the previous period is 0.

If $s'_d \in S$ is such that firm opts to continue in the market then, $V^d(s'_d, e^{d\prime} = 1) = V^{d,1}(s'_d, e^{d\prime} = 1)$ then two cases are possible:

Case 1: If $V^d(s'_d, e^{d\prime} = 0) = V^{d,1}(s'_d, e^{d\prime} = 1) - \lambda_\psi \psi_d$, and we have $V^d(s'_d, e^{d\prime} = 1) - V^d(s'_d, e^{d\prime} = 0) = \lambda_\psi \psi'_d$

Case 2: If $V^d(s'_d, e^{d\prime} = 0) = V^{d,0}(s'_d, e^{d\prime} = 0)$.

then $V^d(s'_d, e^{d\prime} = 1) - V^d(s'_d, e^{d\prime} = 0) = V^{d,1}(s'_d, e^{d\prime} = 1) - V^{d,0}(s'_d, e^{d\prime} = 0) = (\pi^d(z', u'_d) - \psi'_d) + E\Delta V(s'_d, e^{d\prime}) < \lambda_\psi \psi'_d$. We already know that $(\pi^d(z', u'_d) - \psi'_d)$ is increasing in u'_d, and $E\Delta V(s'_d, e^{d\prime})$ is already assumed to be nondecreasing in u_d. Since this algorithm step is generic for any iteration step, we have that $E\Delta V(s_d, e^d)$ is nondecreasing in u_{dj}.

Similarly, we can show that $E\Delta V(s_d, e^d)$ is nondecreasing in z and nonincreasing in ψ_d. ■

Lemma 4. *$\bar{z}_{djt}(u_{djt}, \psi_{djt}, e^d_{jt-1})$ is decreasing in u_{djt} and increasing in ψ_{djt}. Similarly, $\bar{u}_{djt}(z_{jt}, \psi_{djt}, e^d_{jt-1})$ is decreasing in z_{jt} and increasing in ψ_{djt}, and $\bar{\psi}_{djt}(z_{jt}, u_{djt}, e^d_{jt-1})$ is increasing in z_{jt} and increasing in u_{djt}.*

Proof. (Lemma 4) Define $\Lambda(s_{djt}, e^d_{jt-1}) = V^{d,1}(s_{djt}, e^d_{jt-1}) - V^{d,0}(s_{djt}, e^d_{jt-1}) = \pi^d(z_{jt}, u_{djt}) - (1 + \mathbf{1}(e^d_{jt-1} = 0)\lambda_\psi)\psi_{djt} + \beta E\Delta V(s_{djt}, e^d_{jt-1})$.

In the previous lemma we've shown that $E\Delta V(s_{djt}, e^d_{jt-1})$ is nondecreasing in both z_{jt} and u_{djt}. Since the period return , $\pi^d(z_{jt}, u_{djt})$, is strictly increasing in z_{jt} and u_{djt}, $\Lambda(s_{djt}, e^d_{jt-1})$ is strictly increasing in both z_{jt} and u_{djt}. At $\bar{u}_{djt}(z_{jt}, \psi_{djt}, e^d_{jt-1})$, $\Lambda(\bar{u}_{djt}(z_{jt}, \psi_{djt}, e^d_{jt-1}), z_{jt}, \psi_{djt}, e^d_{jt-1}) = 0$. Hence $\bar{u}_{djt}(z_{jt}, \psi_{djt}, e^d_{jt-1})$ is strictly decreasing in z_{jt} and strictly increasing in ψ_{djt}. A similar strategy is valid for proving the remaining part of the lemma. ■

Proof. (Lemma 1) For this proof we will use the time subscpripts explicitly. Let us augment the state vector, s_{dt}, with the previous period incumbency status in market d, e^d_{t-1} and define $\tilde{s}_{dt} = (s_{dt}, e^d_{t-1})$. Also, let us define $\tilde{s}_{dt}(e^d_t = 0) \in \tilde{S}^{d,0}$ is the set of states where a firm's decision about country d at time t is no entry whereas $\tilde{s}_{dt}(e^d_t = 1) \in \tilde{S}^{d,1}$ is the set of states where a firm's decision about country d is entry, and $\tilde{S}^{d,0}, \tilde{S}^{d,1} \in \tilde{S}^d$.

$$
\begin{aligned}
&P(e_t^d = 1 | e_{t-1}^d = 1) \\
&\quad = \int \left(\int_0^\infty \int_0^\infty \int_{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 1)}^\infty \mathcal{P}_d(s_{dt} | \tilde{s}_{dt-1}(e_{t-1}^d = 1)) dz_t du_{dt} d\psi_{dt} \right) d\tilde{s}_{dt-1}(e_{t-1}^d = 1) \\
&P(e_t^d = 1 | e_{t-1}^d = 0) \\
&\quad = \int \left(\int_0^\infty \int_0^\infty \int_{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 0)}^\infty \mathcal{P}_d(s'_d | \tilde{s}_{dt-1}(e_{t-1}^d = 0)) dz_t du_{dt} d\psi_{dt} \right) d\tilde{s}_{dt-1}(e_{t-1}^d = 0)
\end{aligned}
$$

For each state vector $\tilde{s}_{dt}(e_t^d = 0)$ in $\tilde{S}^{d,0}$ containing $(z_t^0, u_{dt}^0, \psi_{dt}^0, e_{t-1}^{d0})$, there exists a $\bar{z}_{dt}(u_{dt}^0, \psi_{dt}^0, e_{t-1}^{d0})$. Let's take the state vector $(u_{dt}^0, \psi_{dt}^0, e_{t-1}^{d0}, \bar{z}_{dt}(u_{dt}^0, \psi_{dt}^0, e_{t-1}^{d0}))$ from $\tilde{S}^{d,1}$. It's obvious that $\bar{z}_{dt}(u_{dt}^0, \psi_{dt}^0, e_{t-1}^{d0})) \geqslant z_t^0$, and also we know that $\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 0) > \bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 1)$ for any (u_{dt}, ψ_{dt}) tuples.

First, I want to show that

$$
\int_{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 1)}^\infty \mathcal{Z}(z_t | \bar{z}_{dt-1}(u_{dt-1}^0, \psi_{dt-1}^0, e_{t-2}^{d0})) dz_t \;>\; \int_{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 0)}^\infty \mathcal{Z}(z_t | z_{t-1}^0) dz_t
$$

The supports of the two integrals overlap in the interval $(\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 0), \infty)$ and in this interval $\mathcal{Z}(z_t | \bar{z}_{dt-1}(u_{dt-1}^0, \psi_{dt-1}^0, e_{t-2}^{d0})) > \mathcal{Z}(z_t | z_{t-1}^0)$ for each value of z_t because of the monotonicity assumption. In the interval $(\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 1), \bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 0))$, $\mathcal{Z}(z_t | \bar{z}_{dt-1}(u_{dt-1}^0, \psi_{dt-1}^0, e_{t-2}^{d0}))$ has nonnegative values. Hence,
$\int_{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 1)}^\infty \mathcal{Z}(z_t | \bar{z}_{dt-1}(u_{dt-1}^0, \psi_{dt-1}^0, e_{t-2}^{d0})) dz_t \geq \int_{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 0)}^\infty \mathcal{Z}(z_t | z_{t-1}^0) dz_t$. Since this inequality is valid for each (u_{dt}, ψ_{dt}) tuple, we have $\int_0^\infty \int_0^\infty \int_{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 1)}^\infty \mathcal{Z}(z_t | \bar{z}_{dt-1}(u_{dt-1}^0, \psi_{dt-1}^0, e_{t-2}^{d0}))\mathcal{U}($
$\geq \int_0^\infty \int_0^\infty \int_{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 0)}^\infty \mathcal{Z}(z_t | z_{t-1}^0) \mathcal{U}(u_{dt} | u_{dt-1}^0) \mathcal{S}_d(\psi_{dt} | \psi_{dt-1}^0) dz_t du_t d\psi_t$.

For each $\tilde{s}_{dt}(e_t^d = 0) \in \tilde{S}^{d,0}$ we have a distinct point in $\tilde{S}^{d,1}$ that satisfies the mentioned inequality of the integrals. Then, we have $P(e_t^d = 1 | e_{t-1}^d = 1) \geq P(e_t^d = 1 | e_{t-1}^d = 0)$. ■

Assumption 2': $\mathcal{Z}(z'_j | z_j)$ follows first order stochastic dominance (FOSD) such that $\tilde{\mathcal{Z}}_{z_j + \varepsilon}(z'_j) \leq \tilde{\mathcal{Z}}_{z_j}(z'_j)$ for all z'_j and with strict inequality at some z'_j. $\tilde{\mathcal{Z}}(\cdot)$ is the cumulative distribution function, and $\varepsilon > 0$. Similarly, $\mathcal{U}_d(u'_{dj} | u_{dj})$ also follows FOSD.

Assumption 3': $\mathcal{S}_d(\psi'_{dj} | \psi_{dj})$ also follows FOSD rule such that $\tilde{\mathcal{S}}_{d, \psi_{dj} - \varepsilon}(\psi'_{dj}) \geq \tilde{\mathcal{S}}_{d, \psi_{dj}}(\psi'_{dj})$ for all ψ'_{dj} and with strict inequality at some ψ'_{dj}.$\tilde{\mathcal{S}}_d(\cdot)$ is the cumulative distribution function, and $\varepsilon > 0$.

Proof. (Lemma 2) Under the new set of assumption (Assumptions 2'-4')

$$
\begin{aligned}
&P(e_t^d = 0 | e_{t-1}^d = 0) \\
&\quad = \int \left(\int_0^\infty \int_0^\infty \int_0^{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 0)} \mathcal{P}_d(s_{dt} | \tilde{s}_{dt-1}(e_{t-1}^d = 0)) dz_t du_t d\psi_t \right) d\tilde{s}_{dt-1}(e_{t-1}^d = 0) \\
&P(e_t^d = 0 | e_{t-1}^d = 1) \\
&\quad = \int \left(\int_0^\infty \int_0^\infty \int_0^{\bar{z}_{dt}(u_{dt}, \psi_{dt}, e_{t-1}^d = 1)} \mathcal{P}_d(s_{dt} | \tilde{s}_{dt-1}(e_{t-1}^d = 1)) dz_t du_t d\psi_t \right) d\tilde{s}_{dt-1}(e_{t-1}^d = 1)
\end{aligned}
$$

For each state vector $\tilde{s}_{dt}(e_t^d = 1)$ in $S^{d,1}$ containing $(z_t^1, u_{dt}^1, \psi_{dt}^1, e_{t-1}^{d1})$, we can find an ε

such that $(u^1_{dt}, \psi^1_{dt}, e^{d1}_{t-1}, z^1_t - \varepsilon)$ is from $\tilde{S}^{d,0}$, and also we know that $\bar{z}_{dt}(u_{dt}, \psi_{dt}, e^d_{t-1} = 0) > \bar{z}_{dt}(u_{dt}, \psi_{dt}, e^d_{t-1} = 1)$ for any (u_{dt}, ψ_{dt}) tuples.

First, I want to show that $\int_0^{\bar{z}_{dt}(u_{dt},\psi_{dt},e^d_{t-1}=0)} \mathcal{Z}(z_t|z_{t-1}-\varepsilon)dz_t \geq \int_0^{\bar{z}_{dt}(u_{dt},\psi_{dt},e^d_{t-1}=1)} \mathcal{Z}(z_t|z^1_{t-1})dz_t$. The supports of the two integrals overlap in the interval $(0, \bar{z}_{dt}(u_{dt}, \psi_{dt}, e^d_{t-1} = 1))$ and in this interval $\tilde{\mathcal{Z}}_{z_{t-1}-\varepsilon}(\bar{z}_{dt}(u_{dt}, \psi_{dt}, e^d_{t-1} = 1)) \geq \mathcal{Z}_{z^1_{t-1}}(\bar{z}_{dt}(u_{dt}, \psi_{dt}, e^d_{t-1} = 1))$ becuse of the FOSD rules. Also, in the interval $(\bar{z}_{dt}(u_{dt}, \psi_{dt}, e^d_{t-1} = 1), \bar{z}_{dt}(u_{dt}, \psi_{dt}, e^d_{t-1} = 0))$, $\mathcal{Z}(z_t|z_{t-1}-\varepsilon)$ has nonnegative values. Hence, $\int_0^{\bar{z}_{dt}(u_{dt},\psi_{dt},e^d_{t-1}=0)} \mathcal{Z}(z_t|z_{t-1} - \varepsilon)dz_t \geq \int_0^{\bar{z}_{dt}(u_{dt},\psi_{dt},e^d_{t-1}=1)} \mathcal{Z}(z_t|z^1_{t-1})dz_t$.

Since this inequality is valid for each (u'_d, ψ'_d) tuple,we have $\int_0^\infty \int_0^\infty \int_0^{\bar{z}_{dt}(u_{dt},\psi_{dt},e^d_{t-1}=0)} \mathcal{Z}(z_t|z_{t-1}-\varepsilon)\mathcal{U}(u_{dt}|u^1_{dt-1})\mathcal{S}_d(\psi_{dt}|\psi^1_{dt-1})dz_t du_t d\psi_t$

$\geq \int_0^\infty \int_0^\infty \int_0^{\bar{z}_{dt}(u_{dt},\psi_{dt},e^d_{t-1}=1)} \mathcal{Z}(z_t|z^1_{t-1})\mathcal{U}(u_{dt}|u^1_{dt-1})\mathcal{S}_d(\psi_{dt}|\psi^1_{dt-1})dz_t du_t d\psi_t$.

For each $\tilde{s}(e^{d\prime} = 1) \in \tilde{S}^{d,1}$ we have a distinct point in $\tilde{S}^{d,0}$ that satisfies the mentioned inequality of the integrals. Then, we have $P(e^d_t = 0|e^d_{t-1} = 0) \geq P(e^d_t = 0|e^d_{t-1} = 1)$. ■

Derivation of Eq.(15) is as below

$$
\begin{aligned}
&E\Delta V(s_d, e^d) = \int\int\int_{\bar{z}^{d\prime}(u',\psi',e^{d\prime}=0|e^d)}^{\infty} \lambda_\psi \psi'_d \frac{\psi_d}{\psi_d} \mathcal{P}_d(s'_d|s_d)dz' du' d\psi' \\
&+ \beta \int\int\int_{\bar{z}^{d\prime\prime}(u'',\psi'',e^{d\prime\prime}=0|e^d)}^{\infty} \lambda_\psi \psi''_d \frac{\psi_d}{\psi_d} \mathcal{P}_d(s''_d|s_d)dz'' du'' d\psi'' + ... \\
&+ \int\int\int_{\bar{z}^{d\prime}(u',\psi',e^{d\prime}=1|e^d)}^{\bar{z}^{d\prime}(u',\psi',e^{d\prime}=0|e^d)} (\pi^d_t \frac{u'}{u} \exp((z'-z))^{(\sigma^d-1)} - \psi_d \frac{\psi'_d}{\psi_d}) \mathcal{P}_d(s'_d|s_d)dz' du' d\psi' \\
&+ \beta \int\int\int_{\bar{z}^{d\prime\prime}(u'',\psi'',e^{d\prime\prime}=1|e^d)}^{\bar{z}^{d\prime\prime}(u'',\psi'',e^{d\prime\prime}=0|e^d)} (\pi^d_t \frac{u''}{u} \exp((z''-z))^{(\sigma^d-1)} - \psi_d \frac{\psi''_d}{\psi_d}) \mathcal{P}_d(s''_d|s_d)dz'' du'' d\psi'' + ... \\
&= \mathcal{O}_\pi(s_d, e^d)\pi^d + \mathcal{O}_\psi(s_d, e^d)\psi_d
\end{aligned}
$$

Proof. (Proposition 1) We already know that for firm j

$$
\begin{aligned}
P(e^d_{jt+1} = 1|e^d_{jt} = 1) \geq P(e^d_{jt+1} = 0|e^d_{jt} = 1) \\
P(e^d_{jt+1} = 0|e^d_{jt} = 0) \geq P(e^d_{jt+1} = 1|e^d_{jt} = 0)
\end{aligned}
$$

The above set of inequalities tells us that a change in market status is punished probabilistically such that the probability of a change in status in a country is less than the probability of maintaining the status quo. For a firm that was in a given set of n countries in the previous period, we'll track the probabilities of moving to $n + i$ and $n + i + 1$ countries where $i \geqslant 0$. Let's compare the probabilities for the two possible future states assuming such that a firm moves to the same set of $n + i$ countries in the case of the transition to $n + i + 1$ countries. If the +1-th country is one of the initial n specific countries, this means that the transition probability related to the remaining $N - n - i$ countries, $\prod_{k\in\Omega_{N-n-i}} P(e^k_{jt+1}|e^k_{jt})$ where Ω_{N-n-i} is the set of $N-n-i$ countries, is lower for the case of transition towards $n+i+1$ countries. If the +1-th country is not one of the initial n specific countries, then it means that incumbency status has changed in another country, which is again less probable than maintaining the status quo in that country. For a firm that can be in any subset of countries at time t, going from state n to $n + i$ requires, on average, more change as i increases, and hence the probability of change nonincreases. Similar reasoning is valid for the $P(n - i|n) \geq P(n - i - 1|n)$ part of the proposition, as well. We know that for a given firm, the transition matrix for the number of markets penetrated is row-concave. $\mathcal{N}(n'|n)$ is a weighted sum of firm transition matrices. For any distribution of firms in the economy, $\mathcal{N}(n'|n)$ is a weighted sum of RCD matrices, and it is straightforward to show that a weighted sum of RCD matrices is also an RCD matrix. ■

A.3 Endogenizing the Returns to the Extensive Margin

In the basic setup, I have not allowed firms to incorporate any possible spillovers between markets into their decision processes or if such spillovers exist thay are external to the firms' decisions. In this section, I will discuss about the class of models that will not hurt the empirical findings of the paper, and allow somehow endogenizing the returns to being in more markets. Let us define $\bar{n}^{-d} = \sum_{k\in N-\{d\}} e^{k\prime}$ that we observe in the data.

$V^{d,0}(s_d, e^d, n^{-d}(\mathbf{s},\mathbf{e}))$ and $V^{d,1}(s_d, e^d, n^{-d}(\mathbf{s},\mathbf{e}))$ are the values for entering or not into the market d given the decisions about the other markets and hence $n^{-d}(\mathbf{s},\mathbf{e})$, respectively. $n^{-d}(\mathbf{s},\mathbf{e}) = \{\hat{n}^{-d}(\mathbf{s},\mathbf{e}), \tilde{n}^{-d}(\mathbf{s},\mathbf{e})\}$. $\hat{n}^{-d}(\mathbf{s},\mathbf{e})$ is the number of markets other than d that would have been penetrated in the case when the calculations for the other markets have assumed that the decision about market d would be entry, $e^{d\prime} = 1$. $\tilde{n}^{-d}(\mathbf{s},\mathbf{e})$ is the number of markets other than d that would have been penetrated in the case when the calculations for the other markets have assumed that the decision about market d would be no entry. $\hat{\mathbf{e}}'$ and $\tilde{\mathbf{e}}'$ are the corresponding decision vectors. δ_π is the percentage of more demand that the firm will be receiving in each country because of one extra market that it is operating in. Similarly, the constant δ_ψ represents the entry cost reducing impact of being in more markets.

$$\begin{aligned} V^{d,1}(s_d, e^d, \hat{n}^{-d}(\mathbf{s},\mathbf{e})) = \pi^d(z, u_d)(1+\delta_\pi \hat{n}^{-d}) \\ -(1+(1-e^d)\lambda_\psi)(1-\delta_\psi \hat{n}^{-d})\psi_d \\ +\beta \int V^d(s_d', 1, n^{-d\prime}(\mathbf{s}', \hat{\mathbf{e}}'))\mathcal{P}(\mathbf{s}'|\mathbf{s})d\mathbf{s}' \end{aligned}$$

$$V^{d,0}(s_d, e^d, \tilde{n}^{-d}(\mathbf{s},\mathbf{e})) = 0 + \beta \int V^d(s_d', 0, n^{-d}(\mathbf{s}', \tilde{\mathbf{e}}'))\mathcal{P}(\mathbf{s}'|\mathbf{s})d\mathbf{s}'$$

Let $V^d(s_{d,}, e^d, n^{-d}(\mathbf{s},\mathbf{e}))$ be the value from the optimal decision from market d given the decisions about all other markets.

$$\begin{aligned} V^d(s_{d,}, e^d, n^{-d}(\mathbf{s},\mathbf{e})) = \max_{e^{d\prime}\in\{0,1\}} \{ V^{d,0}(s_d, e^d, \tilde{n}^{-d}(\mathbf{s},\mathbf{e})) + \sum\nolimits_{k\in N-\{d\}} V^k(s_k, e^k, \tilde{n}^{-d}(\mathbf{s},\mathbf{e})), \\ V^{d,1}(s_d, e^d, \hat{n}^{-d}(\mathbf{s},\mathbf{e})) + \sum\nolimits_{k\in N-\{d\}} V^k(s_k, e^k, \hat{n}^{-d}(\mathbf{s},\mathbf{e}))\} \end{aligned}$$

Upon observing the state $(\mathbf{s},\mathbf{e})$, the firm optimally decides about which markets to enter, and its value is $V(\mathbf{s},\mathbf{e})$.

$$V(\mathbf{s},\mathbf{e}) = \max_{\{e^{1\prime}..e^{N\prime}\}} \sum\nolimits_d V^d(s_{d,}, e^d, n^{-d}(\mathbf{s},\mathbf{e}))$$

Solution of this system of equations gives us $V^{d,1}(\cdot)$, $V^{d,0}(\cdot)$, $\mathbf{e}' = \{e^{1\prime}..e^{N\prime}\}$, $V^d(\cdot)$ and $V(\cdot)$.

If the left-hand side of the below inequality is greater than zero then firm enters and vice versa.

$$\begin{aligned} & V^{d,1}(s_d, e^d, \hat{n}^{-d}(\mathbf{s}, \mathbf{e})) - V^{d,0}(s_d, e^d, \tilde{n}^{-d}(\mathbf{s}, \mathbf{e})) \\ + & \sum\nolimits_{k \in N-\{d\}} \{V^k(s_k, e^k, \hat{n}^{-d}(\mathbf{s}, \mathbf{e})) - V^k(s_k, e^k, \tilde{n}^{-d}(\mathbf{s}, \mathbf{e}))\} \geq 0 \end{aligned}$$

Net returns to be raised from country d

$$\begin{aligned} & V^{d,1}(s_d, e^d, \hat{n}^{-d}(\mathbf{s}, \mathbf{e})) - V^{d,0}(s_d, e^d, \tilde{n}^{-d}(\mathbf{s}, \mathbf{e})) \\ & = \pi^d(z, u_d)(1 + \delta_\pi \hat{n}^{-d}) - (1 + (1 - e^d)\lambda_\psi)(1 - \delta_\psi \tilde{n}^{-d})\psi_d \\ & + \beta E \Delta V^d(s_d, e^d, n^{-d}(\mathbf{s}, \mathbf{e})) \end{aligned}$$

For not complicating the model too much, I impose the below structure onto the firm's calculations about the synergetic returns from other markets: Different alternative specifications for these returns can comply with the empirical analysis of the paper.

$$\begin{aligned} & \sum\nolimits_{k \in N-\{d\}} \{V^k(s_k, e^k, \hat{n}^{-d}(\mathbf{s}, \mathbf{e})) - V^k(s_k, e^k, \tilde{n}^{-d}(\mathbf{s}, \mathbf{e}))\} \\ & \in \{\psi_d \delta_\psi n^{-d}, \pi^d(z, u_d)\delta_\pi n^{-d}, \pi^d(z, u_d)\delta_\pi n^{-d} + \psi_d \delta_\psi n^{-d}\} \end{aligned}$$

Under any of these specification the solution of the firm's decision problem will have components related to only the specific market about which a decision is being made, the number of markets penetrated within the set of other markets.

If we assume that $\hat{n}^{-d} \approx \bar{n}^{-d}$ and $\tilde{n} \approx \bar{n}^{-d}$, then the empirical findings of this paper can be supported under the models that are studied in this section.

B Estimation

B.1 Estimating the production function and productivity

Estimation of productivities is another ingredient for the purposes of this study. The literature has treated productivity as the main force behind selection into market decisions. I estimate firm-specific productivity and control for it along the empirical analysis. I define productivity as a combination of some firm-specific attributes and idiosyncratic shocks and trace the comovement across these attributes and demand shock and entry cost shocks.

So far, we've assumed only one production factor with one factor price. However, firms use capital, K, labor, L, and intermediate goods, M, for producing their outputs. I assume constant returns to scale Cobb-Douglas technologies for production of each good, so that we can have a unit cost of production, c, independent of the firm size. My estimation for sector-specific production functions also verifies this constant returns assumption.

$$Q_{j,t} = \exp(z_{j,t}) K_t^{\alpha_k} L_t^{\alpha_l} M_t^{1-\alpha_k-\alpha_l} \tag{29}$$

Given the production factor prices -wage, rental rate of capital, and price for intermediate inputs- optimal production factor bundle under the production function given in Eq.(29) for firms requires the cost of producing one unit of output as in Eq.(32)[42].

$$c = \frac{C_t}{Q_t} = \frac{w_t}{\alpha_l}\left[\frac{1}{\exp(z_t)(\frac{\alpha_k}{\alpha_l}\frac{w}{r})^{\alpha_k}(\frac{\alpha_m}{\alpha_l}\frac{w}{m})^{\alpha_m}}\right] \tag{32}$$

Q^m_{jt} is the physical output of firm j in industry m at time t as given in Eq.(29). Unfortunately, we cannot observe the physical output in the data. However, we can observe sales revenues R^m_{jt}. The customary solution to this problem is to deflate firm-level sales with an industry price index. This deflation may not eliminate all of the price effect.[43] Hence, for estimation of the sector-specific production technologies and productivity parameters we introduce a firm-specific demand shock, $\exp(\upsilon_{jt})$, to Eq(27).[44]

For estimating the production function parameters we assume that the demand for firm j in industry m, Q^m_{jt}, has the form $Q^m_{jt} = Q_{mt}\frac{(P_{st})^{\bar{\sigma}_m}}{(\bar{p}_{jt})^{\bar{\sigma}_m}}\exp(\upsilon_{jt})$, where Q_{mt} is the total demand for industry m's products, P_{mt} price aggregate of industry m's products, $\bar{p}_{jt}$ is the average price firm j is charging for its output, $\bar{\sigma}_m$ is the average elasticity of substitution for consumers within the demand for the domestic industry's products. $\exp(\upsilon_{jt})$ is a firm-specific idiosyncratic demand shock. Firm j's revenue is $R^m_{jt} = Q^m_{jt}\bar{p}_{jt}$. Using the demand function we have $R^m_{jt} = Q_{jt}^{\frac{\bar{\sigma}_m - 1}{\bar{\sigma}_m}} P_{st}^{\frac{\bar{\sigma}_m - 1}{\bar{\sigma}_m}} R_{st}^{\frac{1}{\bar{\sigma}_m}} (\exp(\upsilon_{jt}))^{\frac{1}{\bar{\sigma}_m}}$. Deflating the revenue with industry average prices, we get Eq.(33)[45]

$$\ln R^m_{jt} - \ln P_{mt} = \frac{\sigma_m - 1}{\sigma_m}\ln Q_{jt} + \frac{1}{\bar{\sigma}_m}\ln\frac{R_{mt}}{P_{mt}} + \frac{1}{\bar{\sigma}_m}\upsilon_{jt} \tag{33}$$

$$\tilde{r}^m_{jt} = \bar{\gamma}^m\alpha_k k_{jt} + \bar{\gamma}^m\alpha_l l_{jt} + \bar{\gamma}^m\alpha_m m_{jt} + \frac{1}{\bar{\sigma}_m}\ln\frac{R_{mt}}{P_{mt}} + \bar{\gamma}^m z_{j,t} + \frac{1}{\bar{\sigma}_m}\upsilon_{jt} \tag{34}$$

where $\tilde{r}^m_{jt} = \ln R^m_{jt} - \ln P_{mt}$ and $\bar{\gamma}^m = \frac{\bar{\sigma}_m - 1}{\bar{\sigma}_m}$. We have both productivity and demand shocks

[42] Taking the logarithm of both sides in Eq.(32) we have

$$\ln c = -\alpha_l\ln\alpha_l - \alpha_k\ln\alpha_k - \alpha_m\ln\alpha_m + \alpha_l\ln w_t + \alpha_k\ln r_t + \alpha_m\ln m_t - z_t \tag{30}$$

I obtain the data for industry-wide price indices and intermediate good price indices from the Statistical Institute (TurkStat) business inclination survey results and market interest rates proxy for the price of capital.

Plugging c into Eq.(11) we get

$$\begin{aligned}\ln R_{dt} = \bar{\kappa}^j_{dt} + (1-\sigma^d)\left[\alpha_l\ln w_t + \alpha_k\ln r_t + \alpha_m\ln m_t\right] \\ + \ln X^d_s(t) + \sigma^d\ln P^d_s(t) + (1-\sigma^d)\ln\tau_d - (1-\sigma^d)z_t + \ln u^d_{j,t}\end{aligned} \tag{31}$$

where $\bar{\kappa}^j_{dt} = (1-\sigma^d)(-\alpha_l\ln\alpha_l - \alpha_k\ln\alpha_k - \alpha_m\ln\alpha_m + \ln\frac{\sigma^d}{\sigma^d - 1})$.

Estimation of Eq.(31) requires data for factor prices, trade barriers, destination market's total demand for each industry's products and average prices for each of these industries in the target market.

[43] De Loecker (2011) discusses this issue extensively. Bias may arise if prices are correlated with inputs. Second, productivity estimates may reflect price and demand variation

[44] I'm ignoring possible capacity constraint issues.

[45] For estimating the production functions, in a way, I'm assuming that the demand reaches every firm in the same way. However, since the total demand is a composite of domestic and foreign demand for products of the industry, access of the demand to every firm is not the same. Hence, as a robustness check I estimate the production functions for different groups of firms in the industry- all firms, domestic firms, export intensive firms.

-$z_{j,t}$ and υ_{jt}, respectively- in the above expression. R_{mt} is an aggregate demand shifter.[46]

I estimate the production functions for each industry m separately. For estimating the above equation I'll use the methodology developed in Levinsohn and Petrin (2003) in which they use intermediate inputs to tackle the simultaneity problem stemming from the unobserved productivity and production factor allocations, which are used in the regression equation and most possibly correlated with the unobservable productivity.

B.2 Estimating the Choice and Revenue Equations

Let Θ be the vector of the parameters to be estimated. R_{djt}'s are the observed export revenues for the destination markets.

$$\mathcal{L}(R, G, z, f|\Theta) = \prod_t \prod_j \prod_d \mathcal{P}_d(R_{djt}, G_{djt}, z_{jt}, f_{jt}|\Theta)$$

For this section let θ be the vector of the parameters to be estimated for the revenue equations. All of the estimation procedures below are for a representative industry.

Parameters to be estimated semiparametically, $\Theta = \{\theta, \alpha_{dt}, \alpha_{zt}, \alpha_{G_{dt}}\}$.

$$P(R_{dj}, G_{dj}, f_j|\Theta) = P(R_{dj}|\Theta, f_j) \times P(G_{dj}|\Theta, f_j) \times P(f_j|\Theta)$$

Posterior distribution is

$$P(R_{dj}, G_{dj}, f_j, \Theta|Data) = \frac{P(R_{dj}, G_{dj}, f_j Data|\Theta) \times P(\Theta)}{P(Data)}$$

Choice

I use the algorithm proposed by Chib and Hamilton (2000 and 2002) and Rossi et al. (2005) for estimating the choice equation parameters.

$$\begin{aligned} p(D^d_{jt}|X_{jt}, Z_{djt}, f_j, \Theta) &= \prod_{d,j,t} p(D^d_{jt}|X_{jt}, Z_{djt}, f_j, , \Theta)^{D^d_{jt}} \times p(1 - D^d_{jt}|X_{jt}, Z_{djt}, f_j, \Theta)^{1-D^d_{jt}} \\ &= \prod_{d,j,t} p(G_{djt} \geqslant 0|X_{jt}, Z_{djt}, f_j, \Theta)^{D^d_{jt}} \times p(G_{djt} < 0|X_{jt}, Z_{djt}, f_j, \Theta)^{1-D^d_{jt}} \end{aligned}$$

Factors

Firm-specific factors show up in three types of equations: outcome (revenue), choice, and measurement (measured productivity) equations.

[46]In the estimating equation I'm controlling for the industry-wide aggregate demand shocks but for firm-specific demand shocks currently I don't have a proxy and an i.i.d assumption may be appropriate.

In a similar production function estimation for the textile sector in Belgium De Loecker (2011) adds a multiplicative demand shock to the production function.

$$
\begin{aligned}
G_{djt} - \widetilde{\ln \pi_{djt}} - Z_{djt}\gamma &= \alpha'_{G_{dt}}\mathbf{f}_j + \eta_{G_{djt}} \\
R^d_{jt} - X_{jt}\theta &= \alpha'_{dt}\mathbf{f}_j + \eta_{djt} \\
z_{jt} &= \alpha_{zt}\mathbf{f}_j + \eta_{jt}
\end{aligned}
$$

After stacking factor loading and uniqueness terms in vectors $\mathcal{A}_t$ and $\mathcal{H}_t$, respectively, Eq.(35) implies an estimation of a regression.

$$
Y_{jt} = \mathcal{A}_t\mathbf{f}_j + \mathcal{H}_{jt} \tag{35}
$$

Idiosyncratic components are assumed to distribute normally, $\mathcal{H}_{jt} \sim N(0, \sum_H)$ where $\sum_H = Diag(\tau^2_{G_{dt}}, \tau^2_{dt}, \tau^2_{zt})$

MCMC estimation of a discrete choice model requires normalization of both the level and variance of the latent variables, G. I use the standard strategy of fixing the variance of the latent variable to 1. I also have to normalize the level of the latent variable so that the constants in the model are identified; hence the mean zero factors. I allow G to be in between -30 and 30. In the theoretical part I have assumed the firm-specific factors to be constant. For the sake of consistency I keep firm-specific factors constant over time. I determine the firm-specific factors for the firms that are active in the sample beginning in 2003. When it comes to 2004, I draw firm-specific factors only for the firms that join the sample in that year. For the ones that were active in 2003, their factors are inherited from the 2003 draws. I keep following the same strategy for the newly added firms in the following years. However, keeping these factors constant over time is not necessary for the identification of these firm-specific factors.

B.3 More About Identification

For this section let's suppress the index, j, and also time index, t.

$$
\begin{aligned}
\ln R_d &= \mu_d(X) + \ln u_d \\
\ln \psi_d &= \mu_{\psi_d}(Z) + \ln \xi_d
\end{aligned}
$$

G_d can be rewritten as

$$
G_d = \mu_{G_d}(X, Z) + u_{G_d}, \quad d \in \{1, ..N\} \tag{36}
$$

where $\mu_{G_d} = \mu_d(X) - \mu_\psi(Z)$ and $u_{G_d} = \ln u_d + \ln \xi_d + \ln\, \mathcal{O}_d$

$$
u_{G_d} = \alpha'_{G_d} f + \eta_{G_d}
$$

and $var(\eta_{G_d}) = \tau^2_{G_d}$.

The joint distributions (u_d, G_d), can be identified up to scale τ_{G_d}, where τ_{G_d}'s are the standard deviations for G's. Then, if we define $\bar{G}_d = \frac{G_d}{\tau_{G_d}}$ we can identify the joint distributions $(u_d, \bar{G}_d)$ given X, Z. After introducing the factor structure into the unobservables, I'm able to recover these joint distributions. I assume that the entire relationship between u_d, ξ_d and $\mathcal{O}_d$ stems from the factor structure captured by the vectors, f. Hence, all the η's are mutually independent.

The expression in Eq.(36) is the augmented latent component of the analysis that proxies for the expected discounted value of a firm deciding to be an incumbent for the current starting period minus deciding to be a non-incumbent for the same period.

Factor Loadings

Using country-specific revenue equations and the choice equations, the identification of factor loadings is possible with some normalizations. Adding productivity estimates as a measurement equation where I introduce a firm-specific factor structure on the estimated productivities does not add much to the identification of the model.

My approach for estimating the factor loading matrix is to normalize some of the factor loadings so that $\mathcal{A}_t$ is block lower triangular and fully identified as in Lopes (2000).[47]

$$\mathcal{A}_t = \begin{pmatrix} \alpha_{1,1} & 0 & 0 & 0 \\ \alpha_{2,1} & \alpha_{2,2} & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ \alpha_{k,1} & \alpha_{k+1,2} & \cdots & \alpha_{k,k} \\ \alpha_{k+1,1} & \alpha_{k+1,2} & \cdots & \alpha_{k+1,k} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,k} \end{pmatrix} \tag{37}$$

where k is the number of firm specific factors and m is the number of equations for a firm at a given year. Since the factor structure absorbs all the covariance across the estimating equations, the remaining covariance matrix for the idiosyncratic shocks, Σ, is diagonal. After the factor loadings are identified this covariance matrix is also identified.

Distributions to be estimated nonparametrically

Theorem (Kotlarski 1967):Let $\mathcal{V}_1, \mathcal{V}_2$ and Θ be three independent real random variables, and let $\mathcal{Y}_1 = \mathcal{V}_1 + \Theta$ and$\mathcal{Y}_2 = \mathcal{V}_2 + \Theta$. If the characteristic function of the pair $(\mathcal{Y}_1, \mathcal{Y}_2)$ does not vanish, then the distribution of $(\mathcal{Y}_1, \mathcal{Y}_2)$ determines the distributions of $\mathcal{V}_1, \mathcal{V}_2, \Theta$ up to a change in location.

To build the similarity between our structure and the theorem we can state our structural equations in the following manner:

$$\begin{aligned} [\bar{G}_d - \mu_{\bar{G}_d}(X, Z)](\alpha_{\bar{G}_d})^{-1} &= f + \eta_{\bar{G}_d}(\alpha_{\bar{G}_d})^{-1} \\ (R^d - \mu_d(X))(\alpha_d)^{-1} &= f + \eta_d(\alpha_d)^{-1} \\ z(\alpha_z)^{-1} &= f + \eta_z(\alpha_z)^{-1} \end{aligned}$$

Kotlarski's theorem tells us that distributions of $f, \eta_{\bar{G}_d}(\alpha_{\bar{G}_d})^{-1}$,and $\eta_d(\alpha_d)^{-1}$are identified. Since we know α's, η's are also identified.

The joint distribution of outcomes is

$$F(R^d|X) = \int F(R^d|X, f) \cdot g_f(f)$$

$g_f(f)$ have already been defined. Owing to the factor structure we have defined, $F(R^d|X, f)$ are independent across d.

[47]Geweke and Zhou (1996) and Aguilar and West (2000) follow similar strategies.

B.4 Discussion About the MCMC Procedures

At time t , we observe $N = 27$ entry decisions for a firm. The decisions are represented by the latent terms G_{jt}'s. Corresependingly, we do observe the revenue equations R_{jt}'s. If a firm has not entered into any market, we do not have any revenue equations for that firms during the estimation.

$$J_{j,t} = \begin{pmatrix} G_{jt}^{d=1} \\ G_{jt}^{d=2} \\ \vdots \\ G_{jt}^{d=N} \\ R_{j,t}^{d=1} \\ \vdots \\ R_{jt}^{d=N} \end{pmatrix} \sim N \left(\begin{pmatrix} x_{G,j,t}^{d=1}\theta_G + \alpha'_{G_{d,t}} f \\ \vdots \\ x_{G,j,t}^{d=M}\theta_G + \alpha'_{G_{d,t}} f \\ x_{R,j,t}^{d=1}\theta_R + \alpha'_{dt} f_j \\ \vdots \\ x_{R,j,t}^{d=M}\theta_R + \alpha'_{dt} f_j \end{pmatrix}, \begin{matrix} I_N & 0 \\ 0 & T_{\eta,N\times N} \end{matrix} \right)$$

$J_{j,t} \sim N\left(X_{j,t}\theta + f_j\alpha_t +, \sum_{Ht}\right)$ where $X_{jt} = diag(x_{G,j,t}^{d=1}, x_{R,j,t}^{d=1}, ...x_{G,j,t}^{d=N}, x_{R,j,t}^{d=N})$, $\theta' = (\underbrace{\theta'_G, ..\theta'_G}_{N \text{ times}}, \underbrace{\theta'_R ... \theta'_R}_{N \text{ times}})$.

Then stacking $J_{j,t}$'s over individual firms we have $J_t = X\theta + \eta$, and $\sum_{Ht}$ is orthogonal to $\sum_{Ht'}$ where $t' \neq t$. If there are m firms in the sample in the SUR regressions J_t contains at most $2 \times m \times 2N$ observations. When we stack J_t's in J, we have at most $6 \times 2 \times m \times 2N$ observations. While estimating θ, I use these observations, and for recovering the factors I also add the measured productivities to this set of observations. Table 10 gives us the prior distribution used in the estimation procedures.

Factor loadings

Priors for the factor loadings are distributed Normal, $\alpha_{ij} \sim N(\mu_0 = 0, A_0)$ for $i \neq j$, and $\alpha_{ij} \sim N(\mu_0 = 0, A_0)1(\alpha_{ii} > 0)$ for $i = 1, ..., k$.

For the first k rows of $\mathcal{A}_t$, $\alpha_i \sim N(\mu_i, A_i)1(\alpha_{ii} > 0)$ where $\mu_i = A_i(A_0^{-1}\mu_0 1_i + \sigma_i^{-2} f'_i y_t)$ and $A_i^{-1} = A_0^{-1}1_i + \sigma_i^{-2} f'_i f_i$. For the last $m - k$ rows of $\mathcal{A}_t$, $\alpha_i \sim N(\mu_i, A_i)1(\alpha_{ii} > 0)$ where $\mu_i = A_i(A_0^{-1}\mu_0 1_i + \sigma_i^{-2} f'_i y_t)$ and $A_i^{-1} = A_0^{-1}1_i + \sigma_i^{-2} f' f$.

Factor Scores

The posterior distribution for the firm-specific factors, $\mathbf{f}_j$, is as given below

$$p(f_j|\Theta, Data) \propto \exp\left(-\frac{1}{2}(Y_{jt} - \mathcal{A}_t\mathbf{f}_j)' \sum\nolimits_H^{-1} (Y_{jt} - \mathcal{A}_t\mathbf{f}_j)\right) p(f_j)$$

$$p(\mathbf{f}_j|\Theta, Data) \propto N(C^{-1}D, C^{-1}) \text{ where } C^{-1} = (\mathcal{A}_t)' \sum\nolimits_H^{-1} (\mathcal{A}_t) + (\tau_f^2)^{-1}$$
$$D = (\tau_f^2)^{-1}\mu_f + (\mathcal{A}_t)' \sum\nolimits_H^{-1} (Y_{jt} - \mathcal{A}_t\mathbf{f}_j)$$

$Y_j = \mathcal{A}_t\mathbf{f}_j + \mathcal{H}_{jt}$ where Y_{jt} is $6 \times 2 \times m \times (2N+1) \times 1$, $\mathcal{A}_t$ is $6 \times 2 \times m \times (2N+1) \times k$, and $\mathbf{f}_j$ is $k \times 1$.

C DATA Preparation

I used the Worldbank WITS website to get 4 digit HS Rev.2 tariff data. I use "Effectively Applied Rates (AHS)" for tariff data. WITS uses the concept of effectively applied tariff rates which is defined as he lowest available tariff. There were missing values in the time series tariff data and in such cases if there was a missing value for a product-country pair, I imputed that value with the closest available data in the time series for that product-country pair. I did not have the tariff data for Syria and to avoid losing all the observations about that country, I imputed the Syrian data with the average product tariff rates over all other countries yearly. I have imputed the missing tariff data with the average tariff for a product across 27 countries in the sample.

For calculating the real exchange rates, I used the Penn World Tables. I used nominal exchange rates and GDP deflators to obtain the real exchange rates.

In computing sectoral real wages, I deflated the nominal wage indices with sectoral CPI. Both data series were taken from TurkStat. Deposit interest rates were taken from the Turkish Central Bank. To compute the sectoral intermediate input price indices, I have weighted the sectoral output price indices with the weights from the sectoral use tables from 2002, and deflated them with a general manufacturing price index.

For building the Ext. Margin variable (number of countries that a firm is in during a given period), I calculated the number of countries from the set of the pre-selected 27 countries in which a firm sold its products during a period. If a firm does not show up in the industry Census in a year between two years in which it shows up, then I assumed that the firm was alive during the year in which it did not show up in the Census.

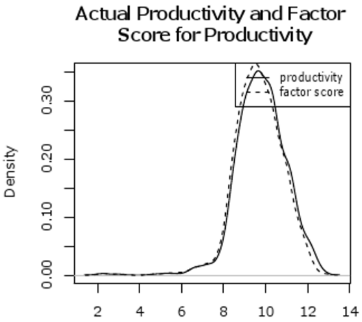
For different firms I determined different market potentials in the same country. First, I determined the 4 digit NACE sector that a firm operates in. Using the conversion tables between the Harmonized system product classification and the European PRODCOM product classification system, which defines products by adding digits to the 4-digit NACE codes calculating the total imports of goods produced by specific NACE industries of a country is possible. I calculated the market potential of a destination country for a domestic firm as the destination market's total exports under the corresponding HS-coded goods for each year. Knowing the sectors of domestic firms, I obtained product flows between countries in the Harmonized System from the COMTRADE dataset. To eliminate the trend effect on the market potential data, I divided the market potential data with the GDP in current dollars index for each country. I also used current USD values of trade transactions in the study. Data on distance in kilometers are from CEPII. I used real effective yearly exchange rate data extracted from the Turkish Central Bank website.

I used firms that employ 20+ employees in the analysis mainly due to the data collection policy of the Statistical Institute. It collects data on firms with 20+ workers every year. However, for smaller firms, the Institute changes the sample every year, which weakens the panel structure of the analysis when those firms are added.[48]

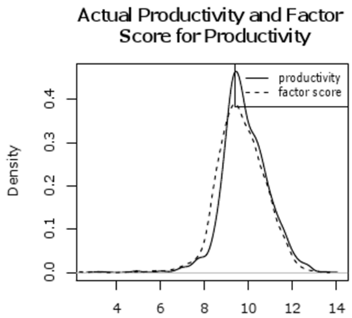
[48] Also, inclusion of those small firms increases time and CPU demand of the analysis.

Figure 5: Productivity in Food industry over time

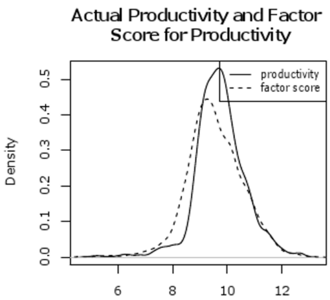
(a) Food industry in 2003



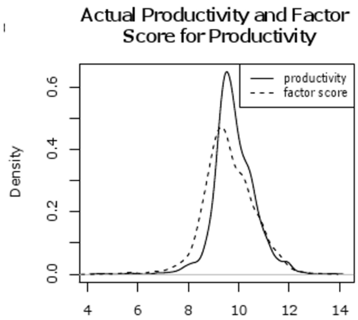
(b) Food industry in 2004



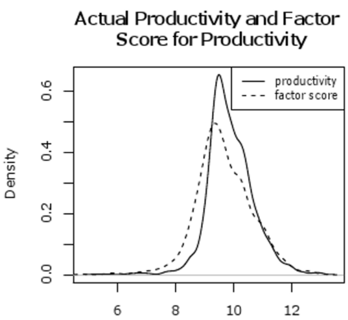
(c) Food industry in 2005



(d) Food industry in 2006



(e) Food industry in 2007



(f) Food industry in 2008

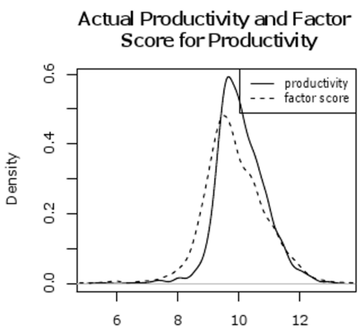
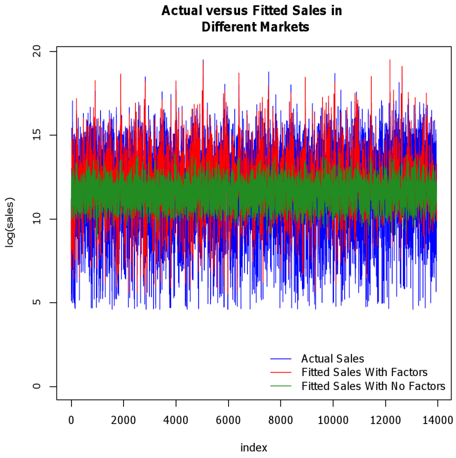
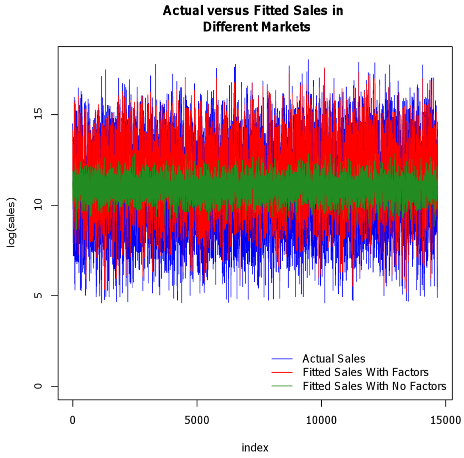


Figure 6: Goodness of Fit

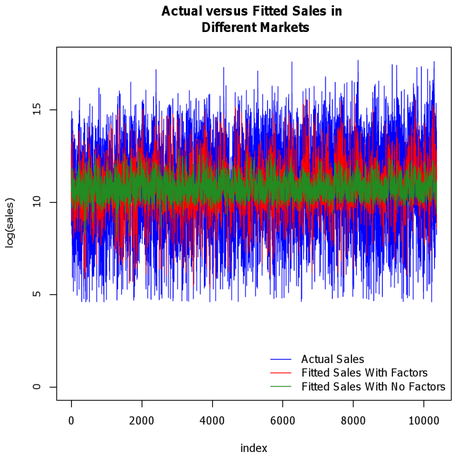
(a) Food Industry



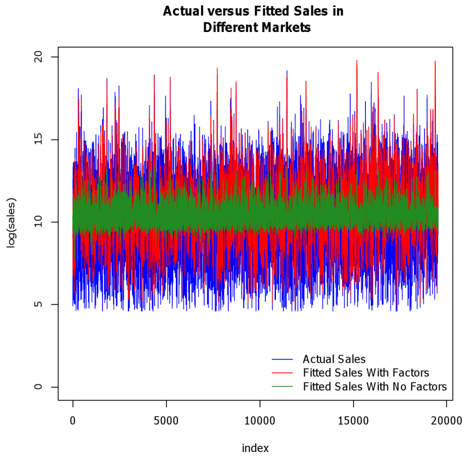
(b) Apparel Industry



(c) Metals Industry



(d) Machinery Industry



(e) Automotive Industry

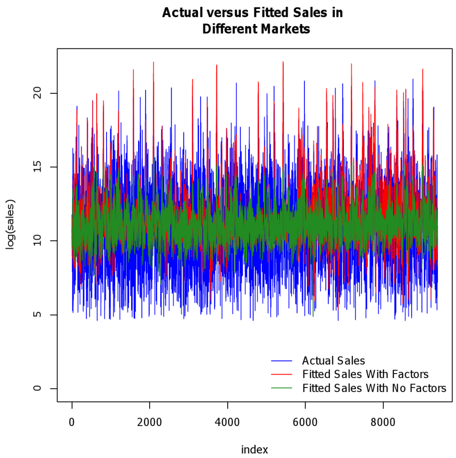


Table 9: Fractions of Selection Behavior Captured Correctly

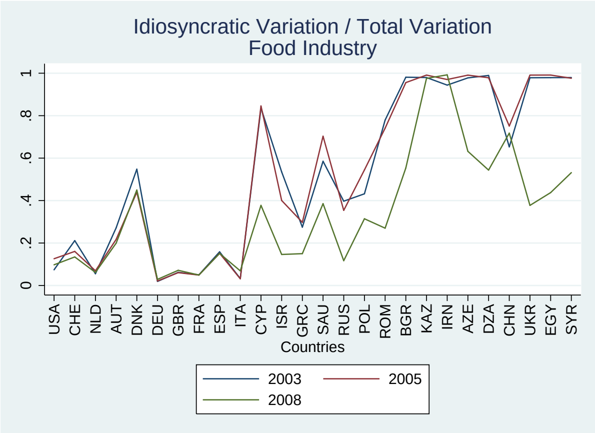
	(NF)		(F)		(F+)	
	Entry	Total	Entry	Total	Entry	Total
Food	0.67	0.92	0.67	0.93	0.73	0.95
Apparel	0.52	0.89	0.52	0.9	0.77	0.92
Metals	0.58	0.88	0.59	0.9	0.58	0.91
Machinery	0.59	0.9	0.58	0.88	0.6	0.9
Automotive	0.71	0.91	0.71	0.87	0.86	0.89

Table 10: Priors

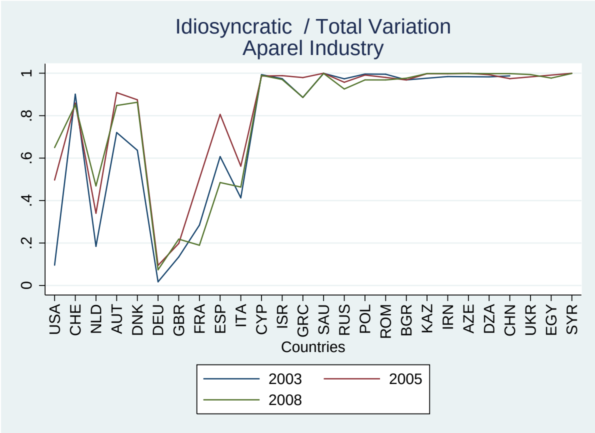
Variable	Disribution	1. parameter	2.Parameter
θ	N(μ_θ,T_θ)	0	$diag(0.02)$
$\sum_H$	$IG(\nu/2, \nu s^2/2)$	4	$\begin{pmatrix} I_N & 0 \\ 0 & diag_N(6.5) \end{pmatrix}$
α	N($\mu_0 = 0$,A_0)	0	1

Figure 7: Variance Decomposition in Export Revenues Across Countries

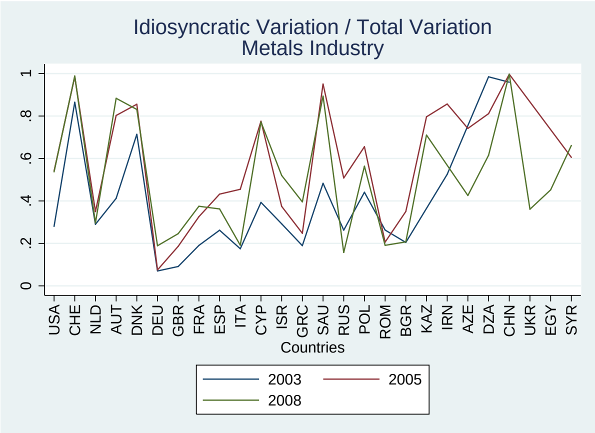
(a) Food Industry



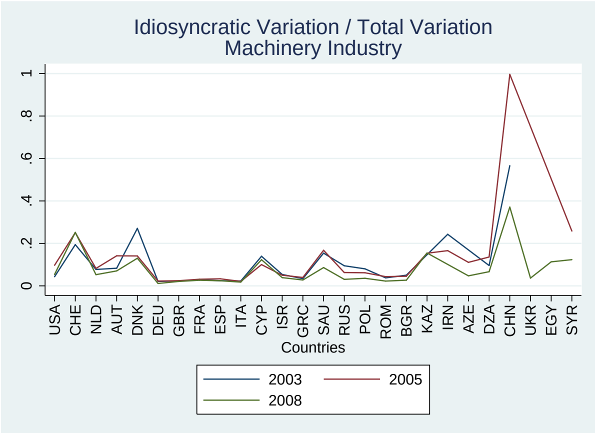
(b) Apparel Industry



(c) Metals Industry



(d) Machinery Industry



(e) Automotive Industry