Essays in International Trade and Growth

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Jose G. Asturias

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Timothy J. Kehoe and Fabrizio Perri, Advisors

August, 2013
Acknowledgements

I would like to thank my advisers Timothy J. Kehoe and Fabrizio Perri for their guidance and encouragement. I am also appreciative of Thomas Holmes for the help he gave me along the way. This dissertation would not exist without their help. I would also like to thank my coauthors Manuel Garcia-Santana, Sewon Hur, Scott Petty, Roberto Ramos, and Kim Ruhl for the long discussions that we had together. In addition, I am grateful to Terry Roe for his comments as well as for being a member of my dissertation committee. Many thanks to seminar participants of the International Workshop at the University of Minnesota, the Society for Economic Dynamics meeting in 2012, the Workshop on Dynamic Macroeconomics at the University of Vigo in 2012, Congressional Budget Office, U.S. Government Accountability Office, Davidson College, University of Missouri, and Georgetown University School of Foreign Service in Qatar.
Dedication

To my family and friends who made this possible.
Abstract

This thesis is composed of three separate essays.

In the first essay of this thesis, Scott Petty and I examine how the industry structure of the transportation industry affects trade. The containerized maritime transportation industry is characterized by oligopolistic competition and economies of scale. We build a model of endogenous transportation costs within a standard Melitz (2003) framework. Countries with large trade volumes face lower transportation costs because they can take advantage of economies of scale and competition in the transportation industry. We assemble a unique dataset on the containerized maritime transportation industry. The dataset includes the freight prices for transporting a container to a foreign port, the number of transportation firms operating between the US and foreign port, and the port-to-port trade flows. We document facts that are consistent with our theory. First, countries with larger trade volumes pay less in transportation costs. Second, countries with larger trade volumes have more and larger transportation firms. We then calibrate our model and estimate a transportation technology to evaluate trade reforms. Our results indicate that transportation costs fall more in smaller markets from tariff reductions. Models that do not consider these features of the containerized maritime transportation industry will fail to capture 60% of the benefits that our model generates.

In the second essay, Sewon Hur and I examine the role of entry barriers on firm entry and exit, aggregate productivity and output. Using cross-country data, we document that gross domestic product (GDP) per capita is positively correlated with firm entry rates, and that firm entry rates are positively correlated with barriers to firm entry. We develop a model, based on Asturias, Hur, Kehoe, and Ruhl (2012) where aggregate productivity growth is driven by the endogenous entry of productive firms and the endogenous exit of unproductive firms. Differences in entry policy lead to different levels of entry and output, while all economies grow at a balanced growth path with identical growth rates. In the quantitative extension, we show that reforms to entry costs can generate transition paths that resemble that of high-growth emerging economies.
In the third essay, Manuel Garcia-Santa, Scott Petty, and Roberto Ramos and I explore transportation costs and regional differences in scale of operation in India. The large variation in cross-country per capita income is a very important puzzle for economists. Recent works addressing this puzzle have focused on policy distortions that prevent the optimal allocation of resources: an inefficiently large amount of resources is allocated to less productive firms. The result is that we see too many small firms in poor countries. We build a new dataset of firm size distribution and prices by districts in India. Our findings indicate that in remote districts, firm size distribution is skewed towards small firms and prices are higher. We propose a method to see how large welfare gains would be if India had lower transportation costs.
Contents

Acknowledgements i
Dedication ii
Abstract iii
Contents v
List of Tables viii
List of Figures ix

1 Endogenous Transportation Costs 1
  1.1 Introduction ........................................ 1
  1.2 Related Literature .................................... 4
    1.2.1 Trade Costs ...................................... 4
    1.2.2 Transportation Costs .............................. 5
    1.2.3 Endogenous Transportation Costs .................. 6
  1.3 Empirical Analysis ..................................... 6
    1.3.1 Constructing the Dataset ......................... 7
    1.3.2 Describing the Data ................................ 8
    1.3.3 Documenting Empirical Relationships .............. 13
  1.4 Model ................................................ 18
    1.4.1 Consumer’s Problem ............................... 19
    1.4.2 Production Firm’s Problem ......................... 19
1.4.3 Transportation Industry .................................................. 21
1.4.4 Trade and Labor Clearing Conditions ............................... 22
1.4.5 Equilibrium ................................................................... 22
1.4.6 Characterizing the Transportation Industry Model ............. 23
1.5 Calibration ........................................................................ 24
  1.5.1 What is a Country? ...................................................... 24
  1.5.2 What is a Firm? .......................................................... 24
  1.5.3 Trade Model Parameters .............................................. 25
  1.5.4 Transportation Model Parameters ................................. 26
  1.5.5 Calibration of the Iceberg Model ................................. 27
1.6 Quantitative Results ............................................................ 27
  1.6.1 What do the Transportation Fixed Costs Look Like? ....... 27
  1.6.2 Simulate 1% Decrease in Tariffs .................................. 30
1.7 Conclusion ........................................................................ 33

2 Entry Barriers and Creative Destruction ................................. 34
  2.1 Introduction ..................................................................... 34
  2.2 Empirical Evidence ........................................................ 35
  2.3 Model ............................................................................. 38
    2.3.1 Households .............................................................. 38
    2.3.2 Producers ................................................................. 39
    2.3.3 Equilibrium ............................................................. 41
    2.3.4 Balanced Growth Path .............................................. 42
    2.3.5 Comparative Statics .................................................. 44
  2.4 Quantitative Analysis ....................................................... 45
  2.5 Conclusion ..................................................................... 46

3 Transportation Costs and Regional Differences in Scale of Operation 48
  3.1 Introduction ..................................................................... 48
  3.2 Data ............................................................................... 49
  3.3 Remoteness and Size Distribution ..................................... 51
  3.4 Remoteness and Prices .................................................... 52
  3.5 Conclusion ..................................................................... 53
References

4 Appendix to chapter 2

4.1 Proofs ................................................................. 64
List of Tables

1.1 Comparing Small and Large Transportation Markets ............... 17
1.2 Price and Containerized Trade Flows ........................... 18
1.3 Parameters for Every US-Foreign Port Combination ............... 25
1.4 Common Parameters ............................................. 26
1.5 Parameters for Transportation Industry for Every US/Foreign Port Combination ............................................. 26
1.6 Dispersion of Fixed Costs ........................................ 27
1.7 Size and Number of Transportation Firms ........................ 30
1.8 Transportation Prices and Trade Flows .......................... 31
1.9 Trade Flows After Liberalization ................................. 32
1.10 Real Income $f$ After Liberalization ......................... 32
1.11 Real Income $h$ After Liberalization ......................... 33
2.1 Calibration and parameters ...................................... 45
3.1 Descriptive Statistics: ASI & NSS Firms ...................... 55
3.2 Descriptive Statistics: Districts ................................. 56
3.3 Size Distribution and Transportation Costs .................. 57
3.4 Prices and Transportation Costs ............................... 58
List of Figures

1.1 Distribution of Freight Prices ........................................... 9
1.2 Distribution of Number of Competitors ................................ 10
1.3 Distribution of Total Container Traffic by Port ....................... 11
1.4 Distribution of Port-to-Port Containerized Trade Flows ............. 12
1.5 Distribution of Containerized Trade per Shipping Firm ............. 13
1.6 Price vs Distance ......................................................... 14
1.7 Freight Price and Competition ........................................... 15
1.8 Freight Price and Average Firm Size ................................... 16
1.9 log(MC) vs log(FC) ....................................................... 28
1.10 log(Distance) vs log(FC) ............................................... 29
1.11 log(Quality Port Infrastructure) vs log(FC) ......................... 29
2.1 Business entry and GDP per capita .................................... 36
2.2 Change in manufacturing firm age and GDP per capita .......... 37
2.3 Business entry and entry costs ......................................... 38
2.4 Transition ................................................................. 46
3.1 Share of Self-Employment Across Indian Districts .................. 54
Chapter 1

Endogenous Transportation Costs

1.1 Introduction

A commonly made observation in the containerized maritime transportation industry is that larger trade flows lead to lower transportation prices.\(^1\) This industry is characterized by large economies of scale and oligopolistic competition. As trade flows increase, prices decrease because of economies of scale and increased competition among transportation firms. Understanding how transportation costs respond to trade flows is important to policymakers if they want to evaluate the effects of a trade reform. Standard trade models do not consider the response of the transportation industry to changes in trade flows and thus underestimate the effect of trade reforms. This is especially important because transportation costs are an important barrier that firms face when they want to export their goods.\(^2\)

In this paper, we contribute to the understanding of how this industry structure affects trade in three ways. First, we assemble a unique dataset on the containerized maritime transportation industry.\(^3\) We have chosen to look at one particular subset

\(^1\)See Stopford (2009) and Rodrigue (2009) for example.

\(^2\)Anderson and van Wincoop (2004) indicate that the ad-valorem all-commodities arithmetic average is 10.7% for the United States.

\(^3\)The vast majority of internationally manufactured goods are transported in shipping containers. The shipping container is a standardized metal box that can be easily transported by boats, trains, or trucks. The fact that it is standardized lowers the cost of loading and unloading. Levinson (2008) describes how the adoption of the container lowered transportation costs and made service more reliable and Rua (2012) studies the diffusion of containerized trade. Hummels (2007) and Dalton (2011) find that around 70% of US imports of manufactured goods take place through maritime transport.
of the transportation industry and carefully collect industry data. Acquiring prices for transportation is difficult because of its confidential nature and accurate measures of transportation costs are rare.\textsuperscript{4,5} Our pricing data is a clean measure of port-to-port transportation costs due to the standardization of shipping containers and the fact that it was collected from a common source at a single moment.\textsuperscript{6,7} We document that small transportation markets pay 67\% more for transportation than do large markets. Second, we build a model with endogenous transportation costs in which firms in the transportation industry are oligopolistic competitors that enjoy economies of scale. The model generates predictions on how transportation prices respond to increases in trade flows. Lastly, using the assembled dataset on containerized maritime transportation industry, we calibrate the model and back out a transportation technology. We show that the traditional model can fail to capture significant portions of the gains that our model generates.\textsuperscript{8}

To introduce endogenous transportation costs, we embed a two-stage entry game of the transportation industry in a two-country Melitz (2003) framework. In the model, transportation firms pay a fixed cost in order to enter the market in the first stage and then compete a la Cournot in the second stage. Transportation firms operate a

\textsuperscript{4}It is a well known problem that trade-weighted aggregate transportation costs are biased downwards because consumers substitute into goods that are cheaper to import. Furthermore, Hummels and Lugovskyy (2006) show that bilateral CIF/FOB ratios data, a commonly used measure, is unreliable because of its many imputed values and measurement errors.

\textsuperscript{5}Most other paper use freight prices collected from customs data, such as Hummels (2001, 2007); Moreira, Volpe, and Blyde (2008). Limao and Venables (2001) assemble price quotes from a freight forwarder. Kleinert and Spies (2011) collect UPS pricing data.

\textsuperscript{6}The advantage of our pricing data over the data that comes from customs forms is that the latter data has heterogeneity across products, destination, and time that cannot easily be disentangled. In an attempt to recover port-to-port transportation costs from this data, we followed a procedure similar to Hummels (2007). We estimated a regression in which the dependent variable is ad-valorem freight costs and the independent variables are time dummies and foreign port destination dummies using the Waterborne Databanks for 2000-2005. The result is that foreign port destination dummies are not statistically significant on over half the ports in our sample.

\textsuperscript{7}The fact that our pricing data is a snapshot in time is especially important due to the volatility of worldwide transportation prices. The Baltic Dry, an index that tracks the price of moving commodities across major shipping lanes, declined from 2000 to 650 over the course of 2 months starting December 2011. The Shanghai Containerized Freight Index, which tracks containerized freight prices across major shipping lanes, also experiences large swings in prices.

\textsuperscript{8}Few papers deal with the endogeneity of transportation prices and the effects on trade. Francois and Wooton (2001) study theoretically the implications of collusion within the maritime transportation industry on freight prices. Hummels, Lugovskyy, and Skiba (2009); Hummels and Skiba (2002); Skiba (2007) empirically estimate how much transportation costs decline with trade volume.
technology that has increasing returns to scale. As market size grows, transportation firms grow in size and become more profitable because of the economies of scale. The increased profits induces new firms to enter reducing transportation prices. The model accommodates market structures ranging from a monopolist firm to perfect competition as a function of market size.

The introduction of the transportation industry into a model of trade results in several key implications for the relationship between market size and transportation prices. First, smaller markets have higher transportation costs, fewer and smaller transportation firms. Smaller markets will thus pay more because of the market power of transportation firms and the inability to take advantage of economies of scale. Second, how quickly transportation prices decline in relation to trade flows depends critically on the number of existing firms. Transportation markets with few firms will see large gains from increases in trade flows and markets with many firms will see relatively smaller gains. This indicates that the standard model can underpredict the effects in markets with a small number of transportation firms.

In the quantitative part of the paper, we calibrate the model and back out a transportation technology. We then simulate a small tariff reduction and show that the standard model can fail to capture 60% of the benefits that our model generates. We also show that the gains from endogenous transportation costs depend critically on the number of existing shippers in the market. As the number of shippers increases, the predictions of our model and the standard model converge.

What do our results imply for evaluating welfare gains from trade liberalizations in general? The answer will depend on the mode of transport primarily used between two countries. Neighboring countries mainly use land transportation but for non-contiguous countries maritime transportation is by far the most important method of moving goods. Excluding land trade, 40% of US imports are containerized maritime and 70% of US imports are waterborne. Evidence from the transportation literature indicates that economies of scale in non-containerized maritime shipping are also significant.9

This paper is structured as follows: Section 1.2 summarizes related literature. Section 1.3 describes the constructed dataset on transportation prices and the key features of the data. Section 1.4 lays out a model of endogenous transportation costs. Section 1.5

---

9See Christiansen, Fagerholt, Nygreen, and Ronen (2007).
details the calibration of the model. Section 1.6 reports the results from the quantitative exercise. Section 1.7 concludes.

1.2 Related Literature

1.2.1 Trade Costs

Trade costs consist of all costs associated with getting a good to a final user. These costs include the cost of transporting a good, the time spent in transit, insurance, policy barriers (tariffs), information costs (finding clients), contract enforcements, costs associated with the use of different currencies, legal and regulatory costs, and local distribution costs (wholesale and retail). There are a wide variety of frictions many of which are difficult to cleanly measure and disentangle.

Anderson and van Wincoop (2004) conducts a broad literature review of trade costs and concludes that trade costs are significant and there are still many open questions as to what determines them. They also point out that direct measures of trade costs are rare and often inaccurate. Indirect measures suggest that trade costs have declined over the last 50 years.

Some papers attempt measure to estimate various types of trade costs. Jacks, Meissner, and Novy (2008) measure trade costs between the periods of 1870-2000 for France, UK and United States. They find that trade costs have declined over this period. Waugh (2010) estimates trade frictions between rich and poor countries using bilateral trade volumes and price data. He finds that there are significant frictions between rich and poor countries. Allen (2011) attempts to understand why distance is important in gravity equations. He estimates a model with information and transportation costs using data from regional agricultural trade in the Philippines. He finds that search frictions can account for more than 90% of the decrease in trade as distance increases in the gravity equation.

Another set of papers investigates the impact of distance on trade flows and find that the role of distance is as important as ever, when measured as the coefficient of distance in the gravity equation. These papers include Berthelon and Freund (2008), Disdier and Head (2008), and Brun, Carrère, Guillaumont, and De Melo (2005). They find that distance remains important in determining international trade: a 1% increase
in distance is associated with a 1% decline in trade. These papers also find that the evolution of the distance coefficient in gravity equations has remained stable or may have risen through time. This fact is difficult to reconcile with empirical evidence that transportation costs have declined over this period.

1.2.2 Transportation Costs

Hummels (2001) and Hummels (2007) conduct surveys of the data on transportation costs. The first general conclusion that they make is that direct measures of transportation costs are difficult to find and are often error-prone. A commonly used measure of transportation prices is the differences between CIF (Cost, Insurance and Freight) and FOB (Free on Board) values from customs forms as published by the IMF. The difference between CIF and FOB is in theory the value of insurance and freight costs. However, Hummels and Lugovskyy (2006) show that there are a series of problem with this data. First, there are so many imputed values as well as measurement errors that it contains little useful information for evolution of costs over time or comparing costs across commodities. Another problem is that because these are aggregate transportation costs for all commodities, they can be significantly biased downward. This is because consumers will substitute away from goods that have high transportation costs. The second broad conclusion of these surveys is that transportation costs have declined over the last 50 years.

Other papers investigate the role of infrastructure in determining transportation costs and generally find that the quality of infrastructure is highly correlated to transportation costs and trade flows. These papers include Blonigen and Wilson (2006), Bougheas, Demetriades, and Morgenroth (1999), Limao and Venables (2001), Moreira, Volpe, and Blyde (2008), Wilmsmeier, Hoffmann, and Sanchez (2006), Clark, Dollar, and Micco (2004), Wilson, Mann, and Otsuki (2003).

FOB is a term of sale under which the price invoiced or quoted by a seller includes all charges up to placing the goods on board a ship at the port of departure specified by the buyer. Under CIF, the price invoiced or quoted by a seller includes insurance and all other charges up to the named port of destination.
1.2.3 Endogenous Transportation Costs

Few papers deal with the endogeneity of transportation prices and the effects on trade. Francois and Wooton (2001) study collusion within the maritime transportation industry and its effects on freight prices. Hummels, Lugovskyy, and Skiba (2009) empirically estimate a similar model using Latin American data and find that market power among transportation firms can play an important role in the transportation prices that countries face.

1.3 Empirical Analysis

In Section 1.3.1 we describe the construction of our dataset on the containerized liner shipping industry.\textsuperscript{11} In Section 1.3.2, we study the summary statistics of this data and find that there are large dispersions in freight prices, the number of firms that service destination ports, and measures of scale. In Section 1.3.3, we analyze how prices are correlated with distance, the number of active firms, and measures of scale. We summarize our findings by looking at the characteristics of small and large markets. We also use a methodology commonly used in the empirical literature to assess how transportation prices decline with trade flows.

\textsuperscript{11}The data is for the containerized liner shipping industry, which means that transportation firms make regularly scheduled stops at specified ports. The fact that it is regularly scheduled implies that there are larger entry costs because a firm must commit to regularly service a market. The alternative is tramp operations which do not have fixed schedules and are often hired to conduct a single voyage. This mode of transport is mainly used for non-containerized trade, which is transported unpackaged on the ship in large volumes (commonly commodities such as iron ore and coal). Hummels, Lugovskyy, and Skiba (2009) indicate that tramp ships transport less than 5% of US containerized import cargo by value.
1.3.1 Constructing the Dataset

Freight Prices

Data on freight prices were obtained from the freight forwarder\textsuperscript{12} Air Parcel Express (APX).\textsuperscript{13} The advantage of this data is that it is a snapshot of market prices taken at one point in time from a consistent source. The freight prices\textsuperscript{14} are for transporting a 20 foot container from the Port of Los Angeles/Long Beach, Baltimore, Charleston, New York/Newark, and Oakland to over 250 different destination ports in August 2012.\textsuperscript{15}

Route Structure and Number of Competitors

Data on the number of competing transportation firms was assembled using information from Containerisation International, a firm that collects and sells data specifically about containerized shipping to industry participants. From the yearbooks that this firm publishes, we collected data on every vessel engaged in container shipping globally. We also collected data on every container port that operates in the world. From these two sets of data, we identified every line on a global basis and constructed a map of network connectivity. This data constitutes the route structure of the containerized liner service industry. At the end of the process, we have data on each containerized shipping line, the ports on the line, the number of ships, the total capacity of these ships, and the company operating this line. We also have information about strategic alliances that operate shipping lines.\textsuperscript{16}

\textsuperscript{12}A freight forwarder is third party logistics provider. Freight forwarders advise exporters on freight costs and other fees (port charges, consular fees, costs of special documentation, insurance costs, and handling fees), as well about import rules and regulations, the methods of shipping, and the necessary documents related to foreign trade.

\textsuperscript{13}APX aggregates shipping quotes from a network of 800 freight forwarders in the Global MAX Shipping Network through their freight rate calculator.

\textsuperscript{14}The freight price includes the base freight price charged by the shipper and any additional surcharges. The most common surcharge is the Bunker Adjustment fee, which changes depending on the price of oil. Other charges, such as the transportation cost to the port, insurance, and the cost of completing paperwork (such as the Bill of Lading) were not included since these are not subject to the increasing returns to scale nature of the shipping technology.

\textsuperscript{15}Los Angeles and Long Beach, and New York and Newark are two separate ports that are adjacent to each other. We combine data from these ports.

\textsuperscript{16}A strategic alliance involves several firms jointly managing a shipping line in order to take advantage of economies of scale. These alliances have become increasingly important since the mid 1990s. Mega alliances such as the Grand Alliance, New World Alliance and CKYH Alliance control increasing levels of market share. See Panayides and Wiedmer (2011) for more information.
Total Container Traffic by Port

From the Containerisation International yearbooks, we also assemble data on the total container traffic for all the ports in the world of the recent past. For every port, we have data on the total number of 20 foot containers that were loaded and unloaded at the port each year.

Port-to-Port Containerized Trade Flows

Data on port-to-port containerized trade flows came from Waterborne Databanks issued by the US Maritime Administration. The dataset contains information about US international maritime trade on a port-to-port level. The data is broken down to the HS 6 product level and includes the cost of transportation and insurance, weight, and whether the shipment is containerized. The dataset also details whether the container was transshipped through a regional shipping hub. Transhipment occurs when there are no ships that travel directly from the origin to destination. For example, a container that goes from Los Angeles to Thailand may first go to Hong Kong.

Port-to-Port Distances

Distance data, measured in nautical miles, came from the online distance calculator offered by Sea Rates, a website that connects exporters to freight forwarders. The data is of the shortest navigable distance between two ports. For example, the calculator considers the fact that a boat going from Los Angeles to Europe can cross the Panama canal.

Containerized Trade per Shipping Firm

Using the port-to-port containerized trade flows and the number of competitors, we calculate the average trade flow per transportation firm. This gives us a measure of the average size of transportation firms between any two ports in our sample.

1.3.2 Describing the Data

In this section, we first show that there are large dispersions in transportation prices across destinations. We also that there are huge differences in measures of scale (both in
terms of total trade flows by destination port and trade flows per firm) and the number of active firms across these destinations.

**Dispersion of Freight Prices**

There are large differences in prices across destinations to transport a container. Figure 1.1 shows the entire distribution of freight prices in our dataset for which there is more than one shipper. The mean destination has a price of $1939 and the standard deviation is $570. The large number of destinations with prices around $1,000 are Asian destinations that enjoy low prices due to a large trade imbalance between the US and East Asia.\(^\text{17}\)

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{Distribution of Freight Prices}
\end{figure}

\(^{17}\)Trade imbalances have implications for transportation prices. For example, in 2008 the merchandise exports from the United States Asia used 5.6 million containers and merchandise exports from Asia to the United States used 14.5 million containers. Thus, 8.9 million containers need to be re-positioned. In fact, by 2005 70\% of the slots of containerships leaving the United States were empty. Thus, Asian exporters pay on average 50\% more in container shipping costs than American exporters.
Dispersion of Number of Competitors

Next, we will look at the number of competitors across destinations. The distribution of the number of competitors for destination ports that have at least one transportation firm serving the foreign port is shown in Figure 1.2. About 50% of our sample has monopolist shippers and about another 20% has two shippers. However, there are some destinations that have a very large number of firms. For example, Los Angeles to Shanghai has 12 shippers.

**Figure 1.2:** Distribution of Number of Competitors

Dispersion of Total Container Traffic by Port

The first measure of scale is total container traffic by port, shown in Figure 1.3. The distribution closely resembles log normal, which means that it is skewed in the right tail. Furthermore, this skewness is reflected in the fact that the mean of the distribution is 6.2 million containers and the median is 1.5 million containers. The differences across the size of ports are significant: the port in the 10th percentile has container traffic of 240,000 and the port in the 90th percentile has traffic of 5,400,000, which is 22 times the size of the smallest port.
Dispersion of Port-to-Port Containerized Trade Flows

We turn our attention to port-to-port containerized trade flows, as shown in Figure 1.3, as another measure of economies of scale. We find that the dispersion is an order of magnitude larger than the dispersion of total container traffic. The distribution also appears lognormal, and the mean of the distribution is $2,400 million and the median is $500 million. The differences across the port-to-port containerized trade flows are huge: the port in the 10th percentile has trade flows of $25 million and the port in the 90th percentile has trade flows of $5,800 million, which is 232 times the trade flows of the destination port with small trade flows.
Dispersion of Containerized Trade per Shipping Firm

We now look at the containerized trade per shipping firm as constructed in section 1.3.1 as our last measure of economies of scale. The distribution of the log containerized trade per shipping firm is shown in Figure 1.5. This distribution also seems lognormal, meaning that the distribution is highly skewed in the right tail. The mean of the distribution is $780 million and the median is $380 million. Like port sizes and port-to-port containerized trade flows, the differences across containerized trade per shipping firm are significant: the port in the 10th percentile has trade flows of $20 million and the port in the 90th percentile has trade flows of $1,800 million, which is 90 times the containerized trade per shipping firm for the small port.
1.3.3 Documenting Empirical Relationships

In this section, we investigate how prices are correlated with distance, the number of active firms, and measures of scale. We also compare the characteristics of small and large markets in terms of these measures. Lastly, we will use a methodology similar to the one used in the empirical literature to estimate how transportation costs decline with trade flows.

Freight Prices and Distance

A striking feature of the pricing data is the relationship with distance, as shown in Figure 1.6. The correlation is only around -0.08. It is interesting to note that the highest and lowest observations in the scatterplot are roughly at equal distance (around 5,000 miles).
Figure 1.6: Price vs Distance

Figure ?? shows the distribution of prices for Los Angeles across countries. The map shows the transportation price to foreign countries using the average price for ports located in that country. European countries pay roughly the same if not less than some Central American countries, even though the latter are much closer and are on the path most Europe-bound ships take. Africa also pays significantly more than Europe even though they are both at roughly the same distance.

**Freight Prices and Number of Transportation Firms**

As a next step we analyze how prices are related to the number of active transportation firms. Figure 1.7 is the same as Figure 1.6 except that destinations served by monopolist shippers are distinguished from the rest. The graph shows that destinations that have monopolist shippers generally have higher prices.
Freight Prices and Scale of Operation

Next, we study how the scale of operation is related to freight prices. Figure 1.8 shows the relationship between freight price and the containerized trade per shipping firm (as described in Section 1.3.2). There is a negative correlation of -0.46 between the two variables meaning that destinations that have larger firms also have lower transportation costs.
Characteristics of Small vs Large Transportation Markets

Next, we compare the difference between small and large markets by comparing the bottom 20% with the top 20% of our sample in terms of port-to-port containerized trade flows. The results are reported in Table 1.1.

The differences across market size is dramatic. First, large markets are considerably larger than the small markets: the average large market is 374 times the size of a small market. Furthermore, the trade flows per firm in the large markets is 108 times that of the small market. Large markets also have 3 times as many transportation firms. Large markets also enjoy prices that are 40% lower than those of the small market. Lastly, the large markets are 44% further away than the small markets.
Table 1.1: Comparing Small and Large Transportation Markets

<table>
<thead>
<tr>
<th></th>
<th>Bottom 20%</th>
<th>Top 20%</th>
<th>Top/ Bottom 20 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price ($)</td>
<td>2,330</td>
<td>1,390</td>
<td>0.60</td>
</tr>
<tr>
<td>Number of firms</td>
<td>1.55</td>
<td>4.95</td>
<td>3.19</td>
</tr>
<tr>
<td>Containerized trade flows / firm (millions $)</td>
<td>23</td>
<td>2,493</td>
<td>108.39</td>
</tr>
<tr>
<td>Containerized trade flows (millions $)</td>
<td>27</td>
<td>10,120</td>
<td>374.81</td>
</tr>
<tr>
<td>Distance (naut. miles)</td>
<td>3,990</td>
<td>5,730</td>
<td>1.44</td>
</tr>
</tbody>
</table>

Freight Prices and Trade Volumes

We will now follow the empirical literature as a baseline to see how freight price are related to trade volumes. We now estimate the following regression

\[
\log(Price_{ij}) = \alpha_1 \log(TradeVolume_{ij}) + \alpha_2 \log(X_{ij}) + \alpha_3 + \epsilon_{ij},
\]

where \( j \) is the foreign port, \( i \) is the US port, and \( X \) is a set of controls.

In estimating this regression we face an issue with the endogeneity of trade volumes. This issue is similar to trying to identify a supply curve and simply regressing price on quantity supplied. We will follow Hummels and Skiba (2002) and instrument for trade volume using the population of the country.\(^{18}\) To use population as an instrument, it must be relevant (correlated to trade volumes) and valid (not correlated to the error term). The correlation between trade volume and population is 0.41, which means that the instrument is relevant. The instrument is valid as changes in population do not affect the supply side. This is equivalent to a demand shifter in trying to identify the supply curve.

We control for fixed effects that include an origin port dummy and a foreign region dummy. Table 1.2 shows the results. First, the estimates on the coefficient of trade volume change dramatically when we include instrument variable, which confirms that trade volume is endogenous. Column 3 indicates that a 1% increase in trade volume leads to a 0.191% decline in transportation costs once we control for observed factors.

\(^{18}\)Using population as an instrument is common in the empirical trade literature. See Frankel, Romer, and Cyrus (1996).
and fixed effects. In the quantitative section, we will compare our results from our calibrated model with those of the reduced form estimates.

**Table 1.2: Price and Containerized Trade Flows**

<table>
<thead>
<tr>
<th></th>
<th>log($Price_{ij}$)</th>
<th>log($Price_{ij}$)</th>
<th>log($Price_{ij}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log($TradeVolume_{ij}$)</td>
<td>-0.083***</td>
<td>-0.065***</td>
<td>-0.191***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.009)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>Distance</td>
<td>0.233***</td>
<td>0.122**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.059)</td>
<td></td>
</tr>
<tr>
<td>Total container traffic foreign port</td>
<td>-0.037***</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td>GDP per capita</td>
<td>0.070***</td>
<td>0.083***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.027)</td>
<td></td>
</tr>
<tr>
<td>Foreign port quality</td>
<td>0.081</td>
<td>-0.061</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.106)</td>
<td></td>
</tr>
<tr>
<td>English official language</td>
<td>0.122**</td>
<td>0.086</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.060)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Fixed effect controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.35</td>
<td>0.63</td>
<td>0.25</td>
</tr>
<tr>
<td>obs</td>
<td>206</td>
<td>206</td>
<td>206</td>
</tr>
</tbody>
</table>

***, **, * indicate statistical significance of 1%, 5%, 10% respectively

### 1.4 Model

We describe a Melitz (2003)\textsuperscript{19} model in which there are two countries, home $h$ and foreign $f$. The transportation industry is modeled as a two-stage entry game with Cournot competition. Transportation firms charge production firms $\tilde{p}$ in order to transport one unit of consumption good to the other country.

\textsuperscript{19}Our model of the transportation industry can also be embedded in other models of trade. We chose Melitz (2003) since Ruhl (2008) has shown that this class of models can replicate the large increases in trade flows observed after trade liberalizations.
1.4.1 Consumer’s Problem

Country $h$ is populated by identical consumers of measure $\bar{\ell}_h$. Each agent supplies one unit of labor to the market and spends his income on a continuum of domestic and imported goods indexed by $v$. The representative consumer of country $h$ takes prices $p_h^h(v)$, $p_f^h(v)$, and wage $w_h$ as given and chooses consumption of home and foreign goods, $c_h^h(v)$ and $c_f^h(v)$ respectively, to solve

$$\max_{c_h^h(v), c_f^h(v)} \left( \frac{1}{\sigma} \int_{v \in \Omega^h} c_h^h(v) \frac{\sigma - 1}{\sigma} dv + (1 - \alpha_h) \frac{1}{\sigma} \int_{v \in \Omega_f^h} c_f^h(v) \frac{\sigma - 1}{\sigma} dv \right)^{\frac{\sigma}{\sigma - 1}} \quad (1.1)$$

subject to

$$\int_{v \in \Omega^h} p_h^h(v)c_h^h(v)dv + \int_{v \in \Omega_f^h} \tau_f^h p_f^h(v)c_f^h(v)dv = w_h \bar{\ell}_h + T_h,$$

where $\Omega^j_i$ is the set of goods from country $j$ available in country $i$, $\sigma$ is the elasticity of substitution, $\alpha_h$ are expenditure weights on goods produced at home, $\tau_f^h$ is the tariff charged by country $h$ and is paid by consumers in country $h$.

The solution to the consumer’s problem in country $h$ for goods from country $f$ is

$$c_f^h(v) = \left( 1 - \alpha_h \right) w_h \bar{\ell}_h \left( \frac{\tau_f^h p_f^h(v)}{P_h^{1-\sigma}} \right)^{\frac{1}{\sigma}} \quad (1.2)$$

$$P_h^{1-\sigma} = \int_{v \in \Omega^h} p_h^h(v)^{1-\sigma} dv + \int_{v \in \Omega_f^h} \left( \frac{\tau_f^h p_f^h(v)}{P_h} \right)^{1-\sigma} dv, \quad (1.3)$$

where $P_h$ is the price index of country $h$.

1.4.2 Production Firm’s Problem

There is an infinite mass of potential entrants. A potential entrant in country $j$ pays fixed cost $w_j F$ to enter the market and draws a productivity $\phi$ from Pareto distribution with CDF $G(\phi) = 1 - \left( \frac{b}{\phi} \right)^\gamma$, where $\phi \geq b$ ($b$ is the minimum draw) and $\gamma > \sigma$. A firm with productivity level $\phi$ has a labor requirement of $1/\phi$ to produce one unit of consumption good.

Firms that only produce in the domestic market do not pay any additional fixed costs. Exporting firms must pay fixed cost $w_h F_{exp}$ to access the foreign market, and it must pay transportation cost $\tilde{p}$ for each unit of consumption good exported.
Pricing Decision

Firm $v$ in country $h$ that enters market $i = h$, $f$ takes aggregate prices, wages, and transportation prices $\tilde{p}_i^h$ as given and solves

$$\pi^i_h(v) = \max_{p_i^h(v)}\left(p_i^h(v)c_i^h(v) - c_i^h(v)\frac{w_h}{\phi(v)} - \tilde{p}_i^h c_i^h(v)\right).$$

(1.4)

There are no domestic transportation costs, $\tilde{p}_h^h = 0$ and $\tilde{p}_h^h = \tilde{p}$. The firms that enter market $i$ will set

$$p_i^h(v) = \frac{\sigma}{\sigma - 1}\left(\frac{w_h}{\phi(v)} + \tilde{p}_i^h\right).$$

(1.5)

The pricing rule, profitability of the firm, and the consumer’s consumption decision only depend on the firm’s productivity. In what follows we will no longer characterize a good by its label $v$ but by the productivity $\phi$ of the firm.

Exporting Decision

Since $\pi^f_h(\phi)$ is increasing in $\phi$, there is a minimum productivity cutoff $\hat{\phi}^f_h$ to export characterized by

$$\pi^i_h(\hat{\phi}^f_h) = w_h F_{exp}.$$  

(1.6)

Free Entry Conditions

The number of firms that draw and operate domestically, $M_h$, is determined by the free entry condition

$$\bar{\pi}_h = w_h F.$$

(1.7)

The expected profit is

$$\bar{\pi}_h = \int_b^\infty \pi^h_\phi dG(\phi) + \int_{\hat{\phi}_h^f}^\infty \pi^f_\phi dG(\phi).$$

The mass of exporters in the home country, $M_{exp,h}$, is determined by

$$M_{exp,h} = \left(1 - G(\hat{\phi}^f_h)\right) M_h.$$

(1.8)
1.4.3 Transportation Industry

The transportation industry is a two-stage entry game. In the first stage, $M$ shippers enter the market to transport goods between countries $i$ and $j$ by paying fixed cost $F^T$. In the second stage, firms compete a la Cournot. The transportation industry is owned and operated by country $h$.\(^{20}\)

**Demand for Transportation**

The demand for transportation is

$$Q = \int_{v \in \Omega^f_h} c^f_h(v) dv + \int_{v \in \Omega^f_j} c^f_j(v) dv. \quad (1.9)$$

Note that $Q$ is strictly decreasing in $\tilde{p}$ since $c^f_j(v) = (1 - \alpha_h w_h \ell_h) p^h_j(v)$ and $p^h_j(v) \sigma$ and $P_h^{1-\sigma}$ are strictly increasing in $\tilde{p}$. Let $\tilde{p}(Q)$ be the indirect demand function.

**Second Stage Transportation Firm Problem**

We work backwards through the entry game and start in the second stage. Suppose that $M$ shippers have entered the market. Firms have transportation technology

$$T = \phi^T \ell^T,$$

where $T$ is the quantity of consumption good transported.\(^{21}\)

Shipping firm $m$ takes competitors quantity supplied $Q_{-m}$ as given and chooses $Q_m$ and solves

$$\max_{Q_m} Q_m \left( \tilde{p}(Q_m + Q_{-m}) - \frac{w_h}{\phi^T} \right). \quad (1.10)$$

There is a unique symmetric equilibrium, so we can write the profit as a function of the number of transportation firms

$$\pi^M = \frac{Q}{M} \left( \tilde{p} - \frac{w_h}{\phi^T} \right) - w_h F^T. \quad (1.11)$$

\(^{20}\)The fact that only country $h$ labor is used ensures that transportation industry parameters are comparable in Section 1.6.

\(^{21}\)Note that firms have increasing returns to scale because in the first stage they paid a fixed cost. Thus, the average cost curve is declining with quantity.
First Stage Transportation Firm Problem

In the first stage, the number of shippers that enter the market $M$ is determined by the free entry condition

$$\pi^M = 0,$$

(1.12)

where $M$ is a continuous variable.

1.4.4 Trade and Labor Clearing Conditions

The trade condition between $h$ and $f$ is

$$M_f \int_{\phi_f^h}^{\phi_f^h} p^h_J(\phi) c^f_J(\phi) dG(\phi) = M_h \int_{\phi_f^h}^{\phi_f^h} p^h_J(\phi) c^f_J(\phi) dG(\phi) + \tilde{p} M_f \int_{v \in \Omega_f^q} \phi_f v dG(\phi),$$

(1.13)

where $\tilde{p} M_f \int_{v \in \Omega_f^q} \phi_f v dG(\phi)$ is the value of transportation services provided by country $h$ to $f$.

The labor clearing condition in country $h$ is

$$M_h \int_{\phi}^{\phi} \phi_f(v) dG(\phi) + M_h \int_{\phi_f^h}^{\phi_f^h} \phi_f(v) dG(\phi) + F M + F_\exp M_{\exp,h} + \frac{Q}{\phi} + M F_T = \ell_h,$$

(1.14)

$$M_f \int_{\phi}^{\phi} \phi_f(v) dG(\phi) + M_f \int_{\phi_f^h}^{\phi_f^h} \phi_f(v) dG(\phi) + F M + F_\exp M_{\exp,f} = \ell_f.$$

(1.15)

1.4.5 Equilibrium

Equilibrium. For $i, j = h, f$, an equilibrium is a set of allocations of consumption good $\{c^i_j(\phi)\}$, export thresholds $\{\phi^i_j\}$, prices for the production good firm $\{p^i_j(\phi)\}$, price for transporting one unit of consumption good $\tilde{p}$, quantity of goods transported $Q_m$ and $Q$, and aggregate variables $\{w_i, P_i\}$ such that

1. Given $\{w_i, p^i_j(\phi)\}$, $\{c^i_j(\phi)\}$ is given by 1.2 and solves the consumer’s problem in 1.1.

2. Given $\{w_i, P_i, \tilde{p}\}$ and $\{c^i_j(\phi)\}$ as given by 1.2, $\{p^i_h(\phi)\}$ is given by 1.5 and solves
the problem of the firm with productivity $\phi$ in 1.4 for all $\phi$ when $i = h$ and for all $\phi \geq \hat{\phi}_h^i$ when $i = f$.

3. Given $\{w_i, P_i\}$, $\{c_j^i(\phi)\}$ as given by 1.2, and $\{p_h^i(\phi)\}$ as given by 1.5, $Q_m$ solves the transportation firm’s problem in 1.10, $\tilde{p}$ satisfies 1.9, and $Q = Q_mM$.

4. Given $\{\pi^M\}$ in 1.11, $M$ satisfies the free entry condition in 1.12.

5. $\{\hat{\phi}_j^i\}$ satisfies 1.6.


7. Trade flows satisfy 1.13.

1.4.6 Characterizing the Transportation Industry Model

We now discuss why increases in market size lead to declines in transportation prices as well as more and larger transportation firms. Suppose that there is an increase in market size. The existing firms in the market grow in size and become more profitable due to the economies of scale of the transportation technology. The increased profit induces new firms to enter the market, reducing transportation prices.

The Cournot structure of our model accommodates a wide range of market structures. Suppose that there is a market that is only big enough for a monopolist shipper. This monopolist will accordingly charge monopolist markups. As market size grows, the number of firms in the market will grow and markups will decline. As market size continues to grow and more firms enter, our model will approach perfect competition where all firms charge marginal cost.\footnote{These results for a Cournot setup are discussed by Sutton (1991).}

Another important implication of the Cournot setup is that the rate at which transportation prices decline is not uniform. For example, an increase in trade is going to lead to a much larger decline in transportation prices in monopolistic markets than in markets with many firms. This is because the monopolistic market has not taken as much advantage of the reductions in market power and economies of scale. This result will play an important role in the next section because it predicts that the gains from lower transportation prices will be larger in ports with few competitors.
1.5 Calibration

1.5.1 What is a Country?

A country in our model corresponds to the entire area serviced by a port in the data. For example, the country of Los Angeles is the entire area that is serviced by this port. To determine the GDP of the country, we use data on port-level inbound and outbound container traffic from Containerisation International. The GDP of a port is the GDP of the country the port is located in multiplied by the fraction of the country’s containers that go through that port.

Although our data is on a port-to-port basis, it is largely consistent with country-to-country gravity equation estimates. The gravity equation is an extremely robust empirical fact that bilateral trade between two countries is proportional to their respective sizes, measured by their GDP, and inversely proportional to the geographic distance between them. A gravity style regression using the new definition of a country results in coefficients on foreign GDP of 0.98 and on distance of -0.72. Disdier and Head (2008) conducted a meta-analysis of the empirical gravity equation literature and find that the average coefficient on distance is around -0.90 and on GDP around 1.

1.5.2 What is a Firm?

A transportation firm in our model corresponds to the number of firms that transport goods between two given ports. A transportation firm is not an entire shipping firm that operates on a global basis, but simply the decision of the firm to make a stop at the destination port. One important feature of the industry is that there are important economies of scope in offering transportation services over a network of ports. For example, vessels traveling between the West Coast of the United States and East Asia will often stop at Los Angeles and Oakland before going to Asia where it will stop at series of ports. Thus, the fixed cost of entering a foreign destination is the cost for a ship to make the additional stop at that foreign destination given the present network structure of the industry.
1.5.3 Trade Model Parameters

The parameters and target statistics that vary by US/foreign port combination are reported in Table 1.3. The labor endowment of the US port $\bar{\ell}_h$ is normalized to 1. The labor endowment of the foreign port $\bar{\ell}_f$ is set to match the GDP of the foreign port.\(^{23}\) The home bias parameters $\alpha_h$ and $\alpha_f$ are set to match the openness of the two ports.\(^{24}\) Tariffs are set to the bilateral trade-weighted averages in the data as reported by World Bank WITS.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Symbol</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor endowment $h$</td>
<td>$\bar{\ell}_h$</td>
<td>Normalized to 1</td>
</tr>
<tr>
<td>Labor endowment $f$</td>
<td>$\bar{\ell}_f$</td>
<td>GDP of foreign port</td>
</tr>
<tr>
<td>Home bias</td>
<td>$\alpha_h, \alpha_f$</td>
<td>Value trade / GDP of both ports</td>
</tr>
<tr>
<td>Tariffs</td>
<td>$G_h, G_f$</td>
<td>Trade-weighted tariffs</td>
</tr>
</tbody>
</table>

Parameters and target statistics that are common across all ports are reported in Table 1.4. $F_{exp}$ is set to match the fact that 18% of US manufacturing firms export. For each US port we take the foreign port in which the maximum percentage of firms exported. Then, the sum (weighted by GDP) over US ports is set to 18%.

$\gamma$ is chosen to match the fact that exporting firms have 26% higher value-added per worker as reported by Bernard, Jensen, Redding, and Schott (2007). For each US port we find the the average exporter value-added premium per worker over all destination ports. The sum (weighted by GDP) over the US ports is set to 26%.

The elasticity of substitution is set to $\sigma = 2.9$ following Feenstra (2010).\(^{25}\) Finally, the fixed cost to operate $F$ is normalized to 1.\(^{26}\)

\(^{23}\)The GDP of a port is defined to be the sum of the value of consumption, investment, and net exports. As in Gibson (2007), the sunk cost of entry $F$ is thought of as the purchase of fixed capital and is thus part of investment. Net exports is characterized by equation 1.13.

\(^{24}\)Trade costs that we do not consider are “soaked up” by the home bias parameter. Obstfeld and Rogoff (2000) discuss that home bias can be mapped isomorphically to trade costs.

\(^{25}\)Feenstra (2010) takes the median value of $\sigma$ across the 4-digit SITC industries imported by a country. He reports that this median estimate is tightly distributed around 2.9.

\(^{26}\)\(F_{exp}\) is expressed in terms of $F$. 
<table>
<thead>
<tr>
<th>Fixed cost to export</th>
<th>$F_{exp}$</th>
<th>18% of plants export</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tail parameter</td>
<td>$\gamma$</td>
<td>Exporters' value-added per worker 26% higher</td>
</tr>
<tr>
<td>Minimum draw</td>
<td>$b$</td>
<td>Normalized to 1</td>
</tr>
<tr>
<td>Elasticity of substitution</td>
<td>$\sigma$</td>
<td>2.9</td>
</tr>
<tr>
<td>Fixed cost to enter</td>
<td>$F$</td>
<td>Normalized to 1</td>
</tr>
</tbody>
</table>

### 1.5.4 Transportation Model Parameters

The parameters for the transportation model are described in Table 1.5. For each US-foreign port pair, we must find the fixed cost $F^T$ and the unit labor requirement $\frac{1}{\sigma^T}$. The fixed cost is set to match the observed number of firms that are operating between the two ports. There are many fixed costs that are consistent with the number of firms in the data, so we choose the fixed cost so that there are $\frac{M+(M+1)}{2}$ firms operating.

The unit labor requirement is set to match a target $\tilde{p}$. We set the ratio of prices across all foreign destinations to match the price dispersion in the assembled data. The level of prices is set to match the trade weighted transportation of that port to the world.\(^{27}\)

| Parameters for Transportation Industry for Every US/Foreign Port Combination |
|------------------------------------------|---------------------------------|----------------|
| Parameters                              | Symbol                          | Target         |
| Fixed cost                              | $F^T$                           | $\pi \frac{M+(M+1)}{2} = 0$, M is shippers in data |
| Unit labor requirement                   | $\frac{1}{\sigma^T}$           | Target $\tilde{p}$ |

The intuition behind the identification of the transportation model parameters is simple. Suppose that there is a destination that has very large trade flows and has few shippers. The model rationalizes the data by assigning it a relatively large fixed cost.

\(^{27}\)The trade-weighted transportation costs are calculated from the 2005 Waterborne Databanks.
Suppose that for a given number of shippers we observe a high price in the market, then the model rationalizes the data by assigning it a high unit labor requirement.

### 1.5.5 Calibration of the Iceberg Model

We use the parameters in Tables 1.3 and 1.4. The $\tau$ is set to match the trade-weighted transportation costs that we observe in the data for that port.

### 1.6 Quantitative Results

#### 1.6.1 What do the Transportation Fixed Costs Look Like?

There is a very large dispersion in the fixed costs across destinations. In Table 1.6, we compare the destination in the 90th vs the 10th percentile in terms of fixed cost. We compare each port separately because the units of the fixed cost are in terms of the labor endowment of the US port, which was normalized to 1 in the calibration. We find large dispersions in fixed costs across destinations. In the average US port in our sample, the destination in the 90th percentile is 73x larger than the destination in the 10th percentile. We look at three possible factors that can be correlated to these fixed costs: distance, marginal cost, and the quality of the foreign port.

<table>
<thead>
<tr>
<th>Port</th>
<th>obs</th>
<th>10% percentile</th>
<th>90% percentile</th>
<th>90/10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charleston</td>
<td>28</td>
<td>0.001123</td>
<td>0.027640</td>
<td>22</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>46</td>
<td>0.000275</td>
<td>0.016243</td>
<td>59</td>
</tr>
<tr>
<td>New York</td>
<td>35</td>
<td>0.000236</td>
<td>0.021466</td>
<td>90</td>
</tr>
<tr>
<td>Oakland</td>
<td>31</td>
<td>0.000074</td>
<td>0.018515</td>
<td>124</td>
</tr>
</tbody>
</table>

We first analyze the correlation between the fixed cost and marginal cost, as shown in Figure 1.9. We see that there is a strong negative correlation between the fixed cost and the marginal cost. Across the US ports that we studied, the average correlation between $\log \frac{1}{\phi T}$ and $F^T$ was -0.57, suggesting that there is a technological
choice that shippers face. Next, we analyze the correlation between distance and the fixed costs, as shown in Figure 1.10. Across the US ports that we studied, the average correlation between $\log dist$ and $\log F^T$ was 0.48. Lastly, we find in Figure 1.11 that the quality of the foreign port is not highly correlated to the fixed cost.

**Figure 1.9:** $\log(MC)$ vs $\log(FC)$
Figure 1.10: \( \log(\text{Distance}) \) vs \( \log(\text{FC}) \)

![Graphs by origin](Image)

Figure 1.11: \( \log(\text{Quality Port Infrastructure}) \) vs \( \log(\text{FC}) \)

![Graphs by origin](Image)
1.6.2 Simulate 1% Decrease in Tariffs

We now simulate a 1% decline in tariffs and analyze the impacts on the cross-section of foreign ports. The results will be broken down by the number of existing firms in the market before the liberalization because this is a critical factor in how responsive price change will be to increased trade flows.

**More and Larger Firms Transportation after Trade Reform**

First we confirm the predictions in Section 1.4.6 that after a liberalization there are more transportation firms and they are larger. Table 1.7 shows the results from the simulation. The percentage increase in the the number of firms was highest in the case of one firm and drops quickly with two shippers. The third column indicates that the average firm size also increased the most when there was a monopolist firm.

<table>
<thead>
<tr>
<th>Initial # of Firms</th>
<th># of Firms (% change)</th>
<th>Average Firm Size (% change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.47</td>
<td>3.45</td>
</tr>
<tr>
<td>2</td>
<td>2.79</td>
<td>2.96</td>
</tr>
<tr>
<td>3</td>
<td>2.53</td>
<td>2.75</td>
</tr>
<tr>
<td>4+</td>
<td>2.29</td>
<td>2.50</td>
</tr>
</tbody>
</table>

**Empirical Literature as a Baseline**

We now compare our results with the standard methodologies used in the empirical literature as reported in Section 1.3.3. We used population as an instrument variable for the demand for transportation. The second column of Table 1.8 shows the declines in $\tilde{p}$ generated by our model. Again, the largest decline takes place when there is one existing firm. As the number of existing shippers increases, the decline in transportation prices drops by more than half from 1.74 to 0.82%. In the third column we also see that the increase in trade flows also declines dramatically from 7.16 to 4.83%. The fourth column indicates that the implied elasticity of change $\tilde{p}$ with respect to trade flows from
Our results are consistent with the strategy adopted by the empirical literature but add some additional insight. The reduced form results indicate that a 1% increase in trade leads to a 0.19% decline in prices. The results from the model show that the decline in transportation prices is not a constant 0.19% but depend critically on the number of existing firms. Thus, the elasticity implied by our model can be 25% higher or 10% lower. Furthermore, the elasticity captured by the reduced form estimates are closer to those ports with 3 or 4 firms, which is a very small portion of the ports in our sample. As shown in Figure 1.2, over 70% of ports in our sample have less than 2 transportation firms.

<table>
<thead>
<tr>
<th>Initial # of Firms</th>
<th>% of Ports</th>
<th>$\tilde{p}$ (% chg)</th>
<th>Trade flows (% chg)</th>
<th>Elasticity Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>-1.74</td>
<td>7.16</td>
<td>0.24</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>-1.43</td>
<td>5.84</td>
<td>0.24</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>-1.19</td>
<td>5.38</td>
<td>0.22</td>
</tr>
<tr>
<td>4+</td>
<td>18</td>
<td>-0.82</td>
<td>4.83</td>
<td>0.17</td>
</tr>
</tbody>
</table>

**Table 1.8: Transportation Prices and Trade Flows**

**Iceberg Model as a Baseline**

Next, we compare the predictions our model with those of models with iceberg transportation costs. Table 1.9 compares the percent change in trade flows. The iceberg model predicts that trade flows will increase by almost the same percent regardless of market structure. The model with endogenous transportation costs predicts that markets with 1 shipper will see far greater increases than those with more shippers. In a monopolistic market, the standard model will miss a little under 40% of the increases in trade flows generated by our model. As the number of shippers increases, the model with iceberg costs converges to our model. Thus, the model with iceberg transportation costs correctly predicts trade flows for markets with a large number of existing firms and falls short in markets with one existing firm.

---

28A change in tariffs can be thought of as a right shift in the demand curve. Thus, if we compare the before and after trade flows and transportation prices, we can trace out the industry average cost curve.
The model with iceberg costs fares worse in predicting welfare. Table 1.10 shows the average percent change in real income for the foreign country.\textsuperscript{29} Note that the welfare gains are small in percentage terms because this is the benefit due to decreasing tariffs by 1% with one foreign port. In destinations with one shipper, a model with iceberg costs will miss 62\% of the gains. Even with 4 and more shippers, the iceberg costs model still misses a substantial 29\% of the gains.

Table 1.10: Real Income $f$ After Liberalization

<table>
<thead>
<tr>
<th>Initial # of Firms</th>
<th>Iceberg (% change)</th>
<th>Endogenous $\bar{p}$ (% change)</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0008</td>
<td>0.0021</td>
<td>0.38</td>
</tr>
<tr>
<td>2</td>
<td>0.0008</td>
<td>0.0015</td>
<td>0.53</td>
</tr>
<tr>
<td>3</td>
<td>0.0018</td>
<td>0.0029</td>
<td>0.62</td>
</tr>
<tr>
<td>4+</td>
<td>0.0051</td>
<td>0.0071</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Table 1.11 shows the percent change in real income for the home country. The home country loses in this trade reform due to the fact that it is lowering tariffs below its optimal level.\textsuperscript{30} The iceberg transportation cost model always overpredicts the losses that country $h$ faces. When there is a monopolist, the iceberg transportation cost model predicts losses that are 365\% larger than those generated by the model with endogenous

\textsuperscript{29}Foreign ports with 4+ shippers gain more from the trade liberalization because they have relatively larger trade with the foreign port as a fraction of GDP than the other ports.

\textsuperscript{30}The optimal tariff for a country is increasing with its country size. In our average simulation, the home GDP is 14 times that of the foreign GDP. This implies that the optimal tariff for the home country will be very large. Irwin (1997) discusses the history of thought of the optimal tariff.
transportation costs. Even in the case with 4 or more shippers, predictions from losses are 36% greater.

**Table 1.11: Real Income $h$ After Liberalization**

<table>
<thead>
<tr>
<th>Initial # of Firms</th>
<th>Iceberg (% change)</th>
<th>Endogenous $\tilde{p}$ (% change)</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.00004</td>
<td>-0.0000086</td>
<td>4.65</td>
</tr>
<tr>
<td>2</td>
<td>-0.00019</td>
<td>-0.0000920</td>
<td>2.07</td>
</tr>
<tr>
<td>3</td>
<td>-0.00047</td>
<td>-0.0002900</td>
<td>1.62</td>
</tr>
<tr>
<td>4+</td>
<td>-0.00150</td>
<td>-0.0011000</td>
<td>1.36</td>
</tr>
</tbody>
</table>

1.7 Conclusion

The goal of this paper has been to show that features of the transportation industry have important implications for trade. Theoretically, we have embedded a model of the transportation industry in a model of trade. The model shows smaller transportation markets will have higher prices, fewer and smaller transportation firms. Furthermore, the model makes predictions about how quickly transportation prices decline as market size grows. Quantitatively, we have shown that a standard model with iceberg transportation costs can miss more than 60 percent of the gains from a trade reform.

This research has implications for policy: raising tariffs is particularly harmful for countries with small GDPs because they are penalized with higher transportation costs. Future research should focus on the implications of the large dispersion in freight prices that we observe that data. One interesting avenue is to think about how variations in prices account for the overall level of frictions that firms face when exporting.
Chapter 2

Entry Barriers and Creative Destruction

2.1 Introduction

Economist Joseph Schumpeter viewed capitalism as the process of industrial mutation that incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one. In other words, innovative entry by entrepreneurs is the disruptive force that sustains economic growth. Indeed, using cross-country data, we find that GDP per capita is positively correlated with firm entry rates. We also find that firm entry rates are positively correlated with various measures of barriers to firm entry.

We develop a model where aggregate productivity growth is driven by the continual entry of young, productive firms and the exit of old, unproductive firms. The model, based on Asturias, Hur, Kehoe, and Ruhl (2012), features a tractable balanced growth path where aggregate productivity, consumption, and output all grow at the same rate. Lower entry costs lead to a larger entry of new firms, and due to general equilibrium effects, lead to a larger exit of old, existing firms. Hence, differences in entry policy lead to different levels of entry and output, while all economies grow at a balanced growth path with identical growth rates. Moreover, in the quantitative extension, we show that reforms to entry costs can generate transition paths that resemble that of high-growth emerging economies.
Our work is closely related to Barseghyan and DiCecio (2011) who construct a model in which cross country differences in entry costs lead to differences in output and productivity. A key departure of our model is that we have an analytically tractable balanced growth path that features endogenous firm entry and exit. Aggregate productivity growth in our model is determined by entry and exit decisions of individual firms. Building up aggregate differences from firm level decisions, as we do, has been researched by several authors, such as Parente and Prescott (1999) and Herrendorf and Teixeira (2011). In these models a coalition secures monopoly rights that allow the coalition to block the adoption of better technologies. In our theory better technology is embodied in new firms and barriers to entry drive suboptimal levels of new firm entry. Our model of heterogeneous firms is also based on Hopenhayn (1992). We have abstracted from firm level uncertainty, which allows us to develop a framework that generates an extremely tractable balanced growth path, rather than the stationary distributions considered in other models of heterogeneous firms.

2.2 Empirical Evidence

In this section we document three patterns in the data: (i) firm entry rate is positively correlated with GDP per capita across countries, (ii) declining average manufacturing firm age is associated with higher growth rates, and (iii) the firm entry rate is decreasing in various measures of entry costs.

As in Asturias, Hur, Kehoe, and Ruhl (2012), we use the World Banks World Development Indicators business entry statistic, which measures the number of new firms officially registered per year, to construct the business entry rate as the number of newly registered firms per thousand working age people. We plot GDP per capita, measured at purchasing power parity, against firm entry rates in Figure 2.1. The data are for the year 2004. Firm entry rates are positively correlated with income per capita, with a correlation coefficient of 0.4.
The second fact that we document is that economies that are growing faster also have declining average firm ages for manufacturing firms. We use data from the World Bank Enterprise Surveys, a program that has collected firm level survey data in developing countries since 2002. Among many questions, the survey asks firms the age of their establishment. Typically 1200-1800 interviews are conducted in larger economies, 360 interviews are conducted in medium-sized economies, and 150 interviews are conducted in smaller economies. After removing resource rich countries, we find that declining average manufacturing firm age is associated with higher growth rates, as can be seen in Figure 2.2. This relationship is statistically significant at the 10
Finally, as in Asturias, Hur, Kehoe, and Ruhl (2012), we document that rate of firm entry are positively correlated with policy induced barriers to entry. In particular, we use the World Banks Doing Business surveys, to construct the measure of entry barriers which estimates the costs of starting up a representative industrial or commercial business. As can be seen in Figure 2.3, the relationship is clearly positive and significant.
2.3 Model

The model is closely related to the one developed in Asturias, Hur, Kehoe, and Ruhl (2012). A key departure from the original model is the abstraction from financial frictions. This allows us to look at not only the effects of entry barriers on firm entry, but also endogenous exit. The model is a closed economy general equilibrium model comprised of a representative household, a representative final good producer, and a continuum of monopolistically competitive intermediate goods producers.

2.3.1 Households

The representative household is endowed with labor, $L$, which is inelastically supplied to firms. The household chooses consumption of composite good, $C_t$, and bond holdings, $B_{t+1}$, to solve

$$\max_{c_t, B_{t+1}} \sum_{t=0}^{\infty} \beta^t \log C_t$$
\[ P_t C_t + q_{t+1} B_{t+1} = w_t L + \Pi_t + B_t \quad (2.1) \]

\[ C_t \geq 0, \quad B_t \geq -g^t B, \quad B_0 \text{ given} \]

where \( \beta \) is the time discount factor and \( \Pi_t \) is the aggregate dividends paid by firms in the economy. We normalize the wage so that \( w_t = 1 \) in each period. The constraint \( B_t \geq -g^t B \) rules out Ponzi schemes but we choose the constant \( B \) large enough so that the constraint does not otherwise bind in equilibrium; \( g \geq 1 \) is the growth factor for the economy.

### 2.3.2 Producers

The representative final good firm, which operates in perfect competition, purchases intermediate goods from intermediate firm \( z, y(z) \), to solve

\[
\min_{y(t)} \int_0^{\eta_t} p_t(z) y_t(z) \, dz
\]

\[ \text{s.t.} \left( \int_0^{\eta_t} y_t(z)^\rho \, dz \right)^{\frac{1}{\rho}} = Y_t \]

The measure of goods available, \( \eta_t \), is an important variable in this model. We show that entry costs and financial frictions decrease the measure of intermediate goods available. The elasticity of substitution between intermediate good varieties is \( \frac{1}{1-\rho} \); the varieties are close, but imperfect, substitutes, \( \rho > 1 \). Solving the final good producer’s problem, we obtain the demand function for good \( z \),

\[ y_t(z) = \frac{Y_t}{p_t(z)^{\frac{1}{1-\rho}} P_t^{\frac{\rho}{1-\rho}}} \quad (2.2) \]

with

\[ P_t = \left( \int_0^{\eta_t} p_t(z)^{-\frac{\rho}{1-\rho}} \, dz \right)^{\frac{1-\rho}{\rho}} \quad (2.3) \]

There is a continuum of intermediate good firms which live for a maximum of \( N \).
periods, unless they choose to exit the market earlier. The firm producing good $z$ that is of age $j$ produces according to

$$ y_t(z) = x(z)\ell_t(z) \quad (2.4) $$

where $x(z)$ is the marginal productivity of firm $z$. Production is subject to a fixed cost of operating, $\kappa_j$, which is denominated in units of labor, and may be conditional on the age of the firm, $j$. If a firm chooses not to pay $\kappa_j$ the firm exits the economy forever. Conditional on choosing to produce, the firm chooses $p_t(j)$ to maximize profit, taking as given the final good price, $P_t$, and aggregate output $Y_t$,

$$ \pi_{jt}(x(z), Y_t, P_t) = \max_{p_t(z)} p_{jt}(z) \frac{Y_t}{p_{jt}(z)^{1-\rho} P_t^{1-\rho}} - \frac{Y_t}{x(z)p_{jt}(z)^{1-\rho} P_t^{1-\rho}} - \kappa_j \quad (2.5) $$

The solution to this problem yields the standard markup over marginal cost pricing, $p_{jt}(z) = (\rho x(z))^{-1}$. We assume that firms can costlessly differentiate their products, so no two firms will choose to produce the same good $z$. Every firm with productivity $x$, however, chooses the same price and sells the same identical quantity. In what follows, we no longer characterize a good by its label, $z$, but by the productivity, $x$, of the firm that produces it.

Every period, a mass $\mu$ of potential entrants draw their productivities from a Pareto distribution that improves each period before deciding to produce—an entry—or to exit immediately. A potential entrant in period $t$ draws its productivity from the distribution

$$ F_t(x) = 1 - \left( \frac{x}{y_t} \right)^{-\gamma} $$

which is characterized by a mean that is improving at rate $g - 1$. Additionally, we require that $\gamma(1 - \rho) - \rho > 0$. We assume that new entrants pay a higher fixed cost than existing firms. In what follows, we consider a simple structure for the fixed costs of operating: a new entrant pays $\kappa$ and, regardless of a firm’s age, all continuing firms pay $f < \kappa$.

In equilibrium, the behavior of firm entry and exit can be summarized by simple cutoff rules. Let $\hat{x}_{jt}$ be the productivity of the firm which has the lowest productivity
of all firms of age $j$ that are operating in period $t$. Profits, $\pi(x)$, increase monotonically with productivity, so every firm of age $j$ with $x \geq \hat{x}_{jt}$ produces in period $t$. Since entering firms must finance any fixed costs from the operating profits during the period in which they enter, the productivity cutoff level of entering firms, denoted as $\hat{x}_{1t}$ is given by the period no-profit condition:

$$p(\hat{x}_{1t})y_t(\hat{x}_{1t}) - \frac{y_t(\hat{x}_{1t})}{\hat{x}_{1t}} - \kappa = 0$$  \hspace{1cm} (2.6)

For firms of age $j > 2$, the cutoff productivity $\hat{x}_{jt}$, in the case of endogenous exit, is determined by the no profit condition,

$$p(\hat{x}_{jt})y_t(\hat{x}_{jt}) - \frac{y_t(\hat{x}_{jt})}{\hat{x}_{jt}} - f = 0$$  \hspace{1cm} (2.7)

or in the case of exogenous exit, is given by the entry cutoff level,

$$\hat{x}_{jt} = \hat{x}_{1,t-j+1}.$$  \hspace{1cm} (2.8)

For the rest of this section, we assume \( \frac{\kappa}{j} < g^N \frac{e^a}{1-p} \), which ensures that firm exit is endogenous.

Given this notation, the mass of firms that are producing at time $t$ is

$$\eta_t = \mu \sum_{i=1}^{N} \left(1 - F_{t-i+1}(\hat{x}_{i,t-i+1})\right).$$  \hspace{1cm} (2.9)

### 2.3.3 Equilibrium

In this section, we focus on balanced growth paths. To define a balanced growth path, we first define an equilibrium. The definition is also useful for potential future research that studies transition paths of the model.

To specify the equilibrium, we need to provide initial conditions on the number of firms entering period 0 that have ages $j > 1$, given by the cutoff levels $\hat{x}_{j0}, j = 2, \ldots, N$, and the bond holdings by households, $B_0$.

**Definition 2.1.** Given these initial conditions, an equilibrium is sequences of entry-exit threshold values, \( \{\hat{x}_{jt}\}_{t=0}^{\infty}, \ j = 1, \ldots, N \), prices and allocations for intermediate
firms, \{p_t(x), y_t(x), \ell_t(x)\}_{t=0}^{\infty}, \text{ for all } x \geq \hat{x}_{jt}, \ j = 1, \ldots, N, \text{ bond prices and bond holdings, } \{q_{t+1}, B_{t+1}\}_{t=0}^{\infty}, \text{ aggregate dividends and final output, } \{D_t, Y_t\}_{t=0}^{\infty}, \text{ and household consumption, } \{C_t\}_{t=0}^{\infty}, \text{ such that:}

1. Given \{P_t, D_t, q_{t+1}\}_{t=0}^{\infty}, \{C_t\}_{t=0}^{\infty} \text{ and } \{B_{t+1}\}_{t=0}^{\infty} \text{ solve the household’s problem in 2.1.}

2. Given \{P_t, Y_t, q_{t+1}\}_{t=0}^{\infty}, \{p_t(x), \ell_t(x)\}_{t=0}^{\infty} \text{ solves the problem of the firm with productivity } x \text{ and age } j \text{ in 2.5, for all } x \geq \hat{x}_{jt}.

3. The labor market clears for all \ t \geq 0,

\[
L = \mu \sum_{i=1}^{N} \int_{\hat{x}_{jt}}^{\infty} (\ell_t(x) + \kappa_i) \, dF_{t-i+1}(x) \, dx
\]  

4. Entry and exit cutoffs satisfy conditions 2.6, 2.7, and 2.8 for all \ j = 1, \ldots, N, \text{ and all } t \geq 0.

5. The bond market clears for all \ t \geq 0,

\[
B_{t+1} = 0.
\]  

6. Dividend payments satisfy for all \ t \geq 0,

\[
\Pi_t = \sum_{i=1}^{N} \int_{\hat{x}_{it}}^{\infty} \tilde{\pi}_i(x) \, dF_{t-i+1}(x) \, dx.
\]

\textbf{2.3.4 Balanced Growth Path}

In this section, we prove that the model has a balanced growth path, and we characterize the behavior of the key variables along the balanced growth path.

\textbf{Definition 2.2.} A \textit{balanced growth path} is a path of entry-exit threshold values, \{\hat{x}_{jt}\}_{t=0}^{\infty}, \ j = 1, \ldots, N, \text{ real aggregate profits, } \{\tilde{\Pi}_t\}_{t=0}^{\infty} \text{ where } \tilde{\Pi}_t = \frac{\Pi_t}{P_t}, \text{ household consumption, } \{C_t\}_{t=0}^{\infty}, \text{ and final good output, } \{Y_t\}_{t=0}^{\infty}, \text{ such that } \hat{x}_{jt}, \ j = 1, \ldots, N, \text{ and, } \tilde{\Pi}_t,
$C_t, Y_t$, grow at the same rate

\[
\frac{\hat{x}_{jt+1}}{\hat{x}_{jt}} = \frac{\hat{\Pi}_{t+1}}{\hat{\Pi}_t} = \frac{C_{t+1}}{C_t} = \frac{Y_{t+1}}{Y_t} = g, \ j = 1, \ldots, N.
\]

We will denote $\hat{x}_{1t}$ as a potential entrant’s minimum productivity necessary to operate and $\hat{x}_{jt}$ as the minimum productivity necessary for an incumbent with age $j \geq 2$ to remain in the market. Furthermore, let $n_t(x)$ be the last age at which a firm of productivity $x$ born at time $t$ remains in the market.

**Proposition 2.1.** Let $\frac{\hat{x}}{j} < \frac{g}{1-\rho}$. In any balanced growth path, the cutoffs $\hat{x}_{jt}$ are characterized by

\[
\hat{x}_{1t} = \kappa \left[ (1 - \rho) (L + \Pi_t) \right]^{1-\rho} P_t^{-1} \rho^{-1}
\]

\[
\hat{x}_{jt} = f \left[ (1 - \rho) (L + \Pi_t) \right]^{1-\rho} P_t^{-1} \rho^{-1}
\]

(2.13) \hspace{1cm} (2.14)

The proof of Proposition 2.1 can be found in the appendix. We can now state Proposition 1 which establishes a balanced growth equilibrium.

**Proposition 2.2.** The economy has a balanced growth.

The proof of Proposition 2.2 is similar to that in Asturias, Hur, Kehoe, and Ruhl (2012), and can be found in the appendix.

We can now characterize . An existing firm of age will remain in the market as long as it remains profitable. This implies that

\[
\hat{x}_{j,t+j-1} = \max \left\{ \kappa \left[ (1 - \rho) (L + \Pi_t) \right]^{1-\rho} P_t^{-1} \rho^{-1} , \ f \left[ (1 - \rho) (L + \Pi_{t+j-1}) \right]^{1-\rho} P_{t+j-1}^{-1} \rho^{-1} \right\}
\]

(2.15)

The first term is the cutoff productivity for firm age $j$ and the second term is the minimum productivity to be profitable that period. Using the balanced growth path conditions that $\Pi_{t+j-1} = \Pi_t$ and $P_{t+j-1} = g^{1-j} P_t$, we get that if $n_t(\hat{x}_{1t}) < N$ then the following condition must be satisfied

\[
g^{n_t(\hat{x}_{1t})-1} \frac{\rho}{1-\rho} \leq \frac{\kappa}{\hat{x}} < g^{n_t(\hat{x}_{1t})} \frac{\rho}{1-\rho}
\]

(2.16)
and if \( n_t(\hat{x}_{1t}) = N \) then
\[
\frac{\kappa}{f} > g^{(N-1)}\frac{e^\theta}{1-\rho}.
\] (2.17)

Note that does not depend on in the balanced growth path. Thus, we will denote
\( \hat{n}(\kappa) = n_t(\hat{x}_{1t}; \kappa) \).

**Proposition 2.3.** If \( g > 1 \), then in any balanced growth path, \( \hat{n}(\kappa) \) is an increasing step function of \( \kappa \).

**Proposition 2.4.** Continuity of the aggregate variables

1. \( P_t(\kappa) \) is continuous in \( \kappa \)
2. \( Y_t(\kappa) \) is continuous in \( \kappa \)
3. \( \Pi(\kappa) \) is continuous in \( \kappa \)
4. \( M(\kappa) \) is continuous in \( \kappa \).

The proofs of Propositions 2.3 and 2.4 are similar to that in Asturias, Hur, Kehoe, and Ruhl (2012), and can be found in the appendix. Proposition 2.4 implies that the aggregate variables are continuous even though \( \hat{n}(\kappa) \) is not continuous. This comes from the fact that when there is a jump in \( \hat{n}(\kappa) \), only an infinitesimally small number of firms change the age at which they exit. This is extremely useful for comparative statics. Since \( Y_t(\kappa) \) and \( M(\kappa) \) are continuous, and differentiable everywhere except at the jump points of \( \hat{n}(\kappa) \), it suffices to consider only the points at which \( Y_t(\kappa, \theta) \) and \( M(\kappa, \theta) \) are differentiable in the comparative statics analysis.

### 2.3.5 Comparative Statics

In this section, we prove that firm entry and output are decreasing in entry costs.

**Proposition 2.5.** Firm entry and output are decreasing in entry costs, i.e.

1. \( \log Y_t(\kappa) \) is decreasing in \( \kappa \), and
2. \( \log M(\kappa) \) is decreasing in \( \kappa \).
Proof. From Proposition 2.4, we know that $Y_t(\kappa)$ and $M(\kappa)$ is continuous in $\kappa$. It suffices to show that $\frac{d \log Y_t(\kappa)}{d \kappa} < 0$ and $\frac{d \log M(\kappa)}{d \kappa} < 0$ for $\kappa \neq \kappa(m), \forall m = 2, ..., N$. We can show these conditions are true. Refer to the technical appendix of Asturias, Hur, Kehoe, and Ruhl (2012) for details. \qed

2.4 Quantitative Analysis

We present here the quantitative results of the baseline model. We choose standard values for the curvature parameter $\gamma$ and elasticity of substitution $\sigma$. $g$ is set to match the productivity growth of the United States, and firms are assumed to live up to 30 years. The parameters used are summarized below in Table 2.1.

<table>
<thead>
<tr>
<th>Table 2.1: Calibration and parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>$\gamma$</td>
</tr>
<tr>
<td>$\sigma$</td>
</tr>
<tr>
<td>$g$</td>
</tr>
<tr>
<td>$\frac{\kappa}{t}$</td>
</tr>
<tr>
<td>$t$</td>
</tr>
<tr>
<td>$n$</td>
</tr>
</tbody>
</table>

As can be seen in Figure 2.4 below, the growth rate of real manufactures consumption is 2 percent in steady state. As the rate of entry cost relative to operational cost decreases gradually from 4 to 1, the economy exhibits rapid growth, averaging 1.5 to 2 percentage points faster than the steady state 2 percent rate of growth.
In this paper, we have developed a model where aggregate productivity growth is driven by the continual entry of young, productive firms and the exit of old, unproductive firms. The model features a balanced growth path where aggregate productivity, consumption, and output all grow at the same rate. We show that lower entry costs lead to a larger entry of new and productive firms, a larger exit of old and productive firms, which leads to greater productivity, output, and consumption.

In the calibrated version of the model, we find that cross-country differences in entry barriers can account for 30% of the differences in GDP per capita. We also show that reforms to entry costs generate periods of high growth before the economy converges to a new balanced growth path, at a higher level of income. Lower entry barriers induce more entry and exit along the transition, which is consistent with the facts we documented in the empirical section.

This paper focused on the effect of entry costs on firm entry, exit, and growth. Asturias, Hur, Kehoe, and Ruhl (2012) explores the interaction between entry costs
and financial frictions. We have abstracted from other potentially important barriers to growth such as flexible labor markets, rule of law, and good institutions. We leave this for future research.
Chapter 3

Transportation Costs and Regional Differences in Scale of Operation

3.1 Introduction

There is a large variation in cross-country per capita income. Much of this variation has been attributed to differences in total factor productivity (TFP): rich countries use resources more efficiently than poor countries.\(^1\) Recent work has quantified the marginal effects of different distortions on TFP within the context of heterogeneous production units.\(^2\) In this literature, the presence of distortions prevents an optimal allocation of resources. In particular, an inefficiently large amount of resources is allocated to small less productive firms, implying big output and welfare losses. This literature emphasizes that due to these distortions poor countries have “too many” small firms.\(^3\)

In this project, we conduct a related empirical exercise using firm-level manufacturing data from India. We combine two sets of data in order to create a representative

\(^1\)See for instance Caselli (2005) and Hall and Jones (1999).

\(^2\)Some examples are Guner, Ventura, and Yi (2008), Restuccia and Rogerson (2008), and Hsieh and Klenow (2009).

\(^3\)Garcia-Santana and Ramos (2012) find that the percentage of employment accounted by small firms is consistently higher in poor countries, and in countries with larger economic distortions such as financial frictions, entry barriers, etc.
sample of Indian firms broken down by districts. Our initial findings show that there is a wide dispersion of firm size distributions across districts in India. Furthermore, we find that in remote districts firm size distribution is skewed towards small firms and prices are higher. These relationships are robust to various specifications.

These empirical findings suggest the need to build a model of trade between Indian districts that incorporates transportation costs as well as market power among firms. The mechanism of the model would be as follows: i) In the presence of high transportation costs, the most productive firms do not find it worthwhile to supply their production in remote areas. ii) This provides monopoly power to less productive entrepreneurs, allowing them to operate and supply goods locally. iii) Misallocation arises because the marginal productivity of large productive firms becomes greater than the marginal productivity of small unproductive ones.

Thus, the model would endogenously generate firm size distribution and price levels at the district level. The model would be calibrated using plant level data on production inputs, prices, and transportation infrastructure at the district-level. The calibrated model would allow us to quantify the effect of transportation costs on aggregate productivity, welfare, and measures of misallocation often used by researchers.

This project is related to other papers that investigate the role of improved infrastructure in India as well as papers such as Donaldson (2010) and Herrendorf, Schmitz, and Teixeira (2012), which show the potentially large gains from infrastructure projects.

### 3.2 Data

We append two complimentary micro datasets of Indian firms: the Annual Survey of Industries (ASI) and the National Sample Survey (NSS) covering roughly the registered and unregistered manufacturing sector, respectively. Both databases correspond to the fiscal year 2005-2006.

The Annual Survey of Industries has been collected by the Indian Government’s Central Statistical Organization since the late 1960’s. It targets all registered factories

---


5 According to the 2-digits NIC, the manufacturing sector corresponds to the codes 15 to 37.
employing 10 or more workers using power and those employing 20 or more workers without using power. The sampling design is twofold. Units with 100 or more workers are categorized as a census sector and the rest of units are treated as a sample sector. In the sample sector, units are stratified according to the 4-digit level of NIC-04 in each State and 1/5th of each strata are selected for coverage, subject to a minimum sample size of 6 units in each stratum. Information collected on units is divided into blocks covering identification, fixed assets, working capital and loans, employment and labor costs, other expenses, other incomes, input items and products and by-products.

The National Sample Survey covers all manufacturing enterprises not targeted by the ASI. Therefore, it includes the unregistered manufacturing sector, which accounts for about one-third of the total contribution of the manufacturing sector to GDP. We use the 62nd round of this survey, covering the fiscal year 2005-2006. In this round, a dual frame approach was used for sampling purposes. First, a list of 8,000 big non-ASI manufacturing enterprises for the urban sector were identified and considered in the database, without resorting to sampling. Second, for the coverage of all other unorganized manufacturing enterprises an area frame approach was followed. Information on surveyed units is divided into blocks covering: identification, operational characteristics, operating expenses, receipts, gross value added, employment, compensation to workers, fixed assets and outstanding loans.

Table 3.1 shows descriptive statistics of both databases. Overall, there are more than 130 thousand observations representing more than 17 million Indian firms. Enterprises from the ASI database are bigger, more productive and pay higher wages. Virtually all firms from the NSS dataset employ less than 20 employees, whereas half of the ASI firms are small. Self-employment accounts for around 47 percent of NSS firms and just 2 percent of ASI firms. 71 percent of NSS firms are located in rural areas. 39 percent of ASI firms are placed in rural locations.

In our empirical specifications, we explore the relationship between remoteness of the Indian districts and the size distribution as well as prices. Therefore, we aggregate the micro data by district. Table 3.2 shows descriptive statistics of the 595 Indian districts in which our firms are located. Remoteness of each district is computed by means of the share of employment accounted by firms located in rural areas and from the percentage of towns in each district located within 10 and 25 miles from a major highway. We
compute statistics of the size distribution namely, the share of employment accounted by small plants and the share of self-employment. There is also wide variation in the size of each district in terms of number of firms and total employment.

3.3 Remoteness and Size Distribution

In this section we show that there is wide variation in the size distribution of Indian firms across Indian districts and that this variation can be accounted by the remoteness of those districts.

There is Wide Variation in Size Distribution Measures in India  Figure 3.1 shows the distribution of self-employment across India. There is wide variation in the size distribution across Indian districts. For the 25 percentile, self-employment accounts for 11.2 per cent of total employment in the district. This figure increases to 19.3 per cent in the median and to 30 per cent in the 75 percentile.

Remoteness is Correlated with Size Distribution  In order to understand the relationship between connectivity and the size distribution of firms, we use our micro data to compute statistics of the size distribution of Indian districts. We then correlate these variables with transportation costs, proxied by the share of employment in rural areas and percentage of towns within 10 miles from a major highway.

Our regressions take the following form:

$$\text{Size Distribution}_{ij} = \text{Remoteness}_{ij} + \log \text{Number of Employees}_{ij} + \text{State}_j$$  (3.1)

Size Distribution$_{ij}$ corresponds to statistics of the size distribution of firms in district $i$ and state $j$. They can be the share of employment that is self-employment, the share of employment in small firms and the log average firm size. Remoteness$_{ij}$ proxies for transportation costs in each district. We use two variables namely, the share of employment in rural locations within the district and the percentage of towns within 10 miles from a major highway. In our regressions we control for size by including the log total number of employees in the district and include state fixed effects to account
for unobservables such as different regulations at the state level. We cluster standard errors at the state level.

Table 3.3 shows the results. Our proxies for transportation costs significantly (negatively) correlate to the size distribution of firms across districts. Columns (1) and (2) show that a 1 percentage point increase in the share of employment in rural locations is associated to .13 percentage points increase in the share of self-employment and .22 percentage points increase in the share of employment accounted by small firms. Note that this result takes into account different regulatory environments across states by introducing fixed effects. Also, column (3) shows that rural districts are also significantly associated to firms of lower average number of employees.

Columns (4) to (6) proxy transportation costs at the district level with the percentage of towns in the district within 10 miles from a major highway. The qualitative results are preserved. Districts that are better connected are associated to firms of larger size. The coefficients are somewhat lower though and even the one associated to the share of self-employment is marginally insignificant (p-value is 10.3).

### 3.4 Remoteness and Prices

In this section, we show that prices of a given product tend to be higher in those districts in which connectivity is more limited. To show this result, we compute average log prices across districts of 4,042 products and regress it against our proxies for transportation costs. Specifically we run regressions of the following form:

\[
\text{Average Log Price}_{pij} = \text{Remoteness}_{ij} + \log \text{Number of Employees}_{ij} \\
+ \log \text{Average Log Wages}_{ij} + \log \text{Number of Firms Selling}_{pij} + \text{State}_j + \text{Product}_p
\]

Average Log Price\(_{pij}\) is the average log price of the product \(p\) located in district \(i\) and state \(j\). Remoteness\(_{ij}\) are proxies of transportation costs in district \(i\). They include the

---

6These are identified by 5-digit ASICC codes. Examples of products are: bran, rice (12801); rice, parboiled (12311); rice, broken (12317); trousers / pants, cotton (63434). We drop those generic products identified as 'other' and 'not elsewhere classified'
share of employment allocated to rural firms, the share of rural firms and the percentage of towns in the district within 10 miles from a major highway. We include additional controls at the district level, namely, the log total number of employees and average log wages in the district, to proxy for input costs. We also include the log number of firms selling the product in the district to account for competition effects. Finally, we include state and product fixed effects.

Table 3.3 shows the results. Column (1) shows that the prices of products produced in districts characterized by a higher share of employment allocated to firms in rural locations are significantly associated to higher prices. This result is quantitatively larger when we include several controls, such as size of district, average wages and the number of firms in the district selling the product. A 10 percentage points increase in the share of employment working in rural locations is associated to an increase of 3.3 percent of prices.

Results are very similar when we proxy transportation costs by the share of firms in rural locations, which we do in columns (3) and (4). In columns (5) and (6) we proxy transportation costs by the percentage of towns in the district that are located within 10 miles from a major Indian highway. In the specification in which we include covariates column (6), we find a negative, although marginally insignificant (p-value is .108). Overall, we find strong evidence that higher prices are observed in those districts in which connectivity is more limited.

3.5 Conclusion

In this project, we have assembled a firm-level dataset that is representative of the Indian manufacturing sector in each district of India. The empirical evidence that we document suggests the need to build a model of trade among Indian districts that incorporate both transportation costs and market power among firms. A calibrated version of the model would allow for quantitative exercises to investigate how changes in transportation costs affect how welfare and measures of misallocation commonly used by researchers.
Figure 3.1: Share of Self-Employment Across Indian Districts
**Table 3.1: Descriptive Statistics: ASI & NSS Firms**

<table>
<thead>
<tr>
<th>Panel A: Annual Survey of Industries</th>
<th>Mean</th>
<th>25 perc.</th>
<th>50 perc.</th>
<th>75 perc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees</td>
<td>70.01</td>
<td>10</td>
<td>20</td>
<td>49</td>
</tr>
<tr>
<td>Gross Value Added per Worker</td>
<td>112.14</td>
<td>12.07</td>
<td>57.22</td>
<td>135.83</td>
</tr>
<tr>
<td>Total Compensation per Employee</td>
<td>54.57</td>
<td>24.48</td>
<td>40.51</td>
<td>64.80</td>
</tr>
<tr>
<td>Total Sales</td>
<td>175726</td>
<td>3682</td>
<td>13098</td>
<td>56480</td>
</tr>
<tr>
<td>Self-Employment</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Small Firm</td>
<td>0.50</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Rural Firm</td>
<td>0.39</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Number of Products</td>
<td>1.56</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

| Panel B: National Sample Survey     |        |          |          |          |
| Number of Employees                 | 2.14   | 1        | 2        | 2        |
| Gross Value Added per Worker        | 2.12   | 0.40     | 0.80     | 1.73     |
| Total Compensation per Employee     | 2.30   | 0.45     | 0.90     | 1.56     |
| Total Sales                         | 143    | 1        | 4        | 14       |
| Self-Employment                     | 0.47   | 0        | 0        | 1        |
| Small Firm                          | 1.00   | 1        | 1        | 1        |
| Rural Firm                          | 0.71   | 0        | 1        | 1        |
| Number of Products                  | 1.05   | 1        | 1        | 1        |

Table 3.1 shows descriptive statistics of Indian firms for the fiscal year 2005-2006. Panel A shows statistics of firms in the Annual Survey of Industries (ASI). Panel B shows statistics of the National Sample Survey (NSS). Gross value added per worker and total compensation per employee are measured in thousand rupees per worker. Total sales is measured in thousand rupees. See Section 3.2 for details of variables.
Table 3.2: Descriptive Statistics: Districts

<table>
<thead>
<tr>
<th></th>
<th>Mean (1)</th>
<th>25 perc. (2)</th>
<th>50 perc. (3)</th>
<th>75 perc. (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of Employment in Small Firms</td>
<td>0.81</td>
<td>0.73</td>
<td>0.89</td>
<td>0.97</td>
</tr>
<tr>
<td>Share of Self-Employment</td>
<td>0.22</td>
<td>0.11</td>
<td>0.19</td>
<td>0.30</td>
</tr>
<tr>
<td>Share of Employment in Rural Firms</td>
<td>0.66</td>
<td>0.49</td>
<td>0.71</td>
<td>0.88</td>
</tr>
<tr>
<td>% of Towns within 10 Miles</td>
<td>0.37</td>
<td>0.00</td>
<td>0.33</td>
<td>0.63</td>
</tr>
<tr>
<td>% of Towns within 25 Miles</td>
<td>0.66</td>
<td>0.33</td>
<td>0.82</td>
<td>1.00</td>
</tr>
<tr>
<td>Total Employment</td>
<td>76.65</td>
<td>15.90</td>
<td>39.27</td>
<td>85.29</td>
</tr>
<tr>
<td>Number of Firms</td>
<td>28.98</td>
<td>6.85</td>
<td>15.17</td>
<td>34.22</td>
</tr>
</tbody>
</table>

Table 3.2 shows descriptive statistics of Indian districts for the fiscal year 2005-2006 constructed from the micro databases ASI and NSS and GIS. Total employment and number of firms are measured in thousands of workers and firms, respectively. See Section 3.2 for details of variables.
Table 3.3: Size Distribution and Transportation Costs

<table>
<thead>
<tr>
<th></th>
<th>Share of Self-Empl.</th>
<th>Share of Empl. in Small Firms</th>
<th>Log Average Firm Size</th>
<th>Share of Self-Empl.</th>
<th>Share of Empl. in Small Firms</th>
<th>Log Average Firm Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Share of Empl. in Rural Locations</td>
<td>0.1391***</td>
<td>0.2203***</td>
<td>-0.7153***</td>
<td>-0.0379</td>
<td>-0.0725**</td>
<td>0.1965***</td>
</tr>
<tr>
<td>Perc. of Towns within 10 Miles</td>
<td>(0.0238)</td>
<td>(0.0506)</td>
<td>(0.1654)</td>
<td>(0.0226)</td>
<td>(0.0269)</td>
<td>(0.0690)</td>
</tr>
<tr>
<td>Log Total Empl.</td>
<td>-0.0242**</td>
<td>-0.0220**</td>
<td>-0.0143</td>
<td>-0.0389***</td>
<td>-0.0298***</td>
<td>0.0868***</td>
</tr>
<tr>
<td></td>
<td>(0.0097)</td>
<td>(0.0087)</td>
<td>(0.0619)</td>
<td>(0.0072)</td>
<td>(0.0080)</td>
<td>(0.0388)</td>
</tr>
<tr>
<td>Observations</td>
<td>592</td>
<td>588</td>
<td>592</td>
<td>568</td>
<td>566</td>
<td>568</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.338</td>
<td>0.36</td>
<td>0.385</td>
<td>0.322</td>
<td>0.322</td>
<td>0.366</td>
</tr>
</tbody>
</table>

Table 3.3 shows the estimation of equation (3.1). The dependent variable is the share of self-employment (1 employee) in the district -columns (1) and (4)-; share of employment in small firms (less than 20 employees) -columns (2) and (5)-; and log average firm size -columns (3) and (6)-. Covariates include the share of employment accounted by firms located in rural areas; percentage of towns in the district within 10 miles from a major Indian highway; and log total employment of the district. State fixed effects are included in all specifications. Robust standard errors are in parenthesis. Significance levels: *: 10%; **: 5%; ***: 1%. See Section 3.2 for details of variables.
Table 3.4: Prices and Transportation Costs

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of Empl. in Rural Locations</td>
<td>0.1636***</td>
<td>0.3324***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0619)</td>
<td>(0.0614)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Firms in Rural Locations</td>
<td>0.2629***</td>
<td>0.3542***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0620)</td>
<td>(0.0611)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perc. of Towns within 10 Miles</td>
<td>0.0392</td>
<td>-0.0743</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0468)</td>
<td>(0.0463)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Total Empl.</td>
<td>0.0670***</td>
<td>0.0686***</td>
<td>0.0623***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0156)</td>
<td>(0.0156)</td>
<td>(0.0165)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Log Wages per Worker</td>
<td>0.0018</td>
<td>-0.0069</td>
<td>-0.0422</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0192)</td>
<td>(0.0188)</td>
<td>(0.0204)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Number of Firms Selling Product</td>
<td>-0.395***</td>
<td>-0.3939***</td>
<td>-0.395***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0192)</td>
<td>(0.0188)</td>
<td>(0.0204)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Product Dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>42,518</td>
<td>42,511</td>
<td>42,518</td>
<td>42,511</td>
<td>40,011</td>
<td>40,007</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.01</td>
<td>0.06</td>
<td>0.01</td>
<td>0.06</td>
<td>0.01</td>
<td>0.06</td>
</tr>
<tr>
<td>Number of asicc</td>
<td>4,034</td>
<td>4,034</td>
<td>4,034</td>
<td>4,034</td>
<td>3,960</td>
<td>3,960</td>
</tr>
</tbody>
</table>

Table 3.4 shows the estimation of equation (3.2). The dependent variable is the average log price of the product in the district. Specifications across columns differ by the covariates included. Columns (1) and (2) include the share of employment accounted by firms located in rural areas as a proxy for transportation costs. Columns (2) and (3) include the share of firms in rural locations. Columns (5) and (6) include the percentage of towns in the district within 10 miles from a major Indian highway. State and product fixed effects are included in all specifications. Robust standard errors are in parenthesis. Significance levels: "*: 10%; **: 5%; ***: 1%. See Section 3.2 for further details.
References


MARITIME TRANSPORTATION,” Handbooks in operations research and management science, 14, 189–284.


“EXPLAINING THE GROWTH IN MANUFACTURING TRADE,” Discussion paper.


Product variety and the gains from international trade. MIT Press.


Chapter 4

Appendix to chapter 2

4.1 Proofs

Proof of Proposition 2.1. From equation 2.6,

\[ p(\hat{x}_{1t})y_t(\hat{x}_{1t}) - \frac{y_t(\hat{x}_{1t})}{\hat{x}_{1t}} - \kappa = 0 \]

Substituting the solution from equation 2.5,

\[ (1 - \rho)(L + \Pi_t)P^t_i \hat{x}^{\rho}_t (\hat{x}_\rho) - \kappa = 0 \]

Applying the balanced growth path conditions \((P_t = g^{-1}P_{t-1}, \Pi_t = \Pi_{t-1})\)

\[ (1 - \rho)(L + \Pi_t)P^t_i \hat{x}^{\rho}_t (\hat{x}_\rho) - \kappa = 0 \]

Rearranging terms yields the desired result.

From equation 2.7,

\[ p(\hat{x}_{jt})y_t(\hat{x}_{jt}) - \frac{y_t(\hat{x}_{jt})}{\hat{x}_{jt}} - f = 0 \]

Substituting the solution from equation 2.5,

\[ (1 - \rho)(L + \Pi_t)P^t_i \hat{x}^{\rho}_t (\hat{x}_\rho) - f = 0 \]

64
Applying the balanced growth path conditions \( P_t = g^{-1}P_{t-1}, \Pi_t = \Pi_{t-1} \), and rearranging terms yields the desired result.

It remains to show that \( \hat{x}_{jt} \geq \hat{x}_{1,t-j+1} \) for \( j = 2, \ldots, N \).

\[
\hat{x}_{jt} = f^{\frac{1-\rho}{\rho}}[[1-\rho](L + \Pi_t)]^{\frac{1-\rho}{\rho}} P_{t-1}^{-1} \rho^{-1}
\]

\[
\hat{x}_{1,t-j+1} = \kappa^{\frac{1-\rho}{\rho}} [[(1-\rho)(L + \Pi_{t-j+1})]^{\frac{1-\rho}{\rho}} P_{t-j+1}^{-1} \rho^{-1}
\]

Applying the balanced growth path conditions \( P_t = g^{-1}P_{t-1}, \Pi_t = \Pi_{t-1}, \beta = q_t \)

\[
\hat{x}_{1,t-j+1} = \kappa^{\frac{1-\rho}{\rho}} [[(1-\rho)(L + \Pi_t)]^{\frac{1-\rho}{\rho}} (P_t g^{j-1})^{-\frac{1}{\rho}} \rho^{-1}
\]

Hence,

\[
\frac{\hat{x}_{1,t-j+1}}{\hat{x}_{jt}} = \frac{\kappa^{\frac{1-\rho}{\rho}} [[(1-\rho)(L + \Pi_t)]^{\frac{1-\rho}{\rho}} (P_t g^{j-1})^{-\frac{1}{\rho}} \rho^{-1}}{f^{\frac{1-\rho}{\rho}}[[1-\rho](L + \Pi_t)]^{\frac{1-\rho}{\rho}} P_{t-1}^{-1} \rho^{-1}}
\]

\[
= \frac{\kappa^{\frac{1-\rho}{\rho}} [((1-\rho)(L + \Pi_{t-j+1})]^{\frac{1-\rho}{\rho}} (P_t g^{j-1})^{-\frac{1}{\rho}} \rho^{-1}}{f^{\frac{1-\rho}{\rho}}[[1-\rho](L + \Pi_t)]^{\frac{1-\rho}{\rho}} P_{t-j+1}^{-1} \rho^{-1}}
\]

\[
= \kappa^{\frac{1-\rho}{\rho}} (g^{j-1})^{-\frac{1}{\rho}} f^{\frac{1-\rho}{\rho}}\]

\[
< 1
\]

where the last inequality follows from our assumption that \( \frac{\kappa}{f} < g^N \frac{\rho}{1-\rho} \).

\[\square\]

**Proof of Proposition 2.2.** The proof of Proposition 2.2 involves guessing and verifying the existence of an equilibrium with a balanced growth path.

The expression for aggregate profits is given by
\[ \Pi_t = \sum_{i=1}^{N} \mu \int_{\hat{x}_{it}}^{\infty} \pi_{it}(x) \, dF_{t-i+1}(x) \]

\[ = \sum_{i=1}^{N} \mu \left\{ (1 - \rho)(L + \Pi_t) \frac{1}{\rho} P_t^{1-\rho} \frac{\rho^{\gamma(t-i+1)}(1 - \rho)}{\gamma(1 - \rho) - \rho} \frac{\rho^{\gamma(1-\rho)}}{\hat{x}_{it}^{1-\rho}} \right\} \]

The expression for the intermediate good price is given by

\[ P_t = \sum_{i=1}^{N} \mu \int_{\hat{x}_{it}}^{\infty} p(x) \, dF_{t-i+1}(x) \]

\[ = \rho \frac{\rho^{\gamma(1 - \rho)}}{\gamma(1 - \rho) - \rho} \sum_{i=1}^{N} g^{\gamma(t-i+1)} \frac{\rho^{\gamma(1-\rho)}}{\hat{x}_{it}^{1-\rho}} \]

Using the expression for cutoffs from Proposition 2.1 which are given by,

\[ \hat{x}_{it} = \kappa^{1-\rho} \left[ (1 - \rho)(L + \Pi_t) \right]^{1-\rho} P_t^{1-\rho} \]

\[ \hat{x}_{jt} = f^{1-\rho} \left[ (1 - \rho)(L + \Pi_t) \right]^{1-\rho} P_t^{1-\rho} \]

the balanced growth path conditions \((P_t = g^{-1} P_{t-1}, \Pi_t = \Pi_{t-1})\) and solving for \(P_t\) and \(\Pi_t\) yields

\[ P_t(\kappa)^{-\gamma} = g^{\gamma t} \left[ (1 - \rho)(L + \Pi(\kappa)) \right]^{1-\rho} \frac{\rho^{\gamma(1-\rho)}}{\gamma(1 - \rho) - \rho} \mu \omega(\kappa) \]

\[ \Pi(\kappa) = \frac{L \xi(\kappa)}{1 - \xi(\kappa)} \]

where

\[ \xi(\kappa) = 1 - \rho - \frac{\gamma(1 - \rho)}{\gamma \omega(\kappa)} \left( \kappa^{1-\rho} \gamma \sum_{i=1}^{\hat{n}(\kappa)} \kappa_i + f^{1-\rho} \gamma \sum_{i=\hat{n}(\kappa)+1}^{N} g^{\gamma(1-i)} \right) \]

\[ \omega(\kappa) = \kappa^{1-\rho} \gamma \sum_{i=1}^{\hat{n}(\kappa)} g^{(1-i)} \frac{\rho}{\gamma} + f^{1-\rho} \gamma \sum_{i=\hat{n}(\kappa)+1}^{N} g^{\gamma(1-i)} \]
Real output is given by
\[ Y_t(\kappa) = \frac{L + \Pi(\kappa)}{P_t(\kappa)} \]

Thus, our guess has been verified.

\[ \square \]

Proof of Proposition 2.3. We will show that there are \( N - 1 \) jump points. For \( m \) such that \( 2 \leq m \leq N \), define \( \kappa(m) \) to be \( \kappa(m) = fg^{(m-1)\frac{\rho}{1-\rho}} \). We show that \( \kappa(m) \) are the jump points of the step function. We show this in two steps. First, it is clear that for small \( \epsilon > 0 \),
\[
\frac{\kappa(m) - \epsilon}{f} < g^{(m-1)\frac{\rho}{1-\rho}}
\]
which violates the condition that the marginal entrant lives to the age of \( m \).

Second, it is clear that for small \( \epsilon > 0 \),
\[
g^{(m-2)\frac{\rho}{1-\rho}} \leq \frac{\kappa(m) + \epsilon}{f} < g^{(m-1)\frac{\rho}{1-\rho}}
\]
which satisfies the condition that the marginal entrant lives to the age of \( m - 1 \).

Finally, it is clear that for small \( \epsilon > 0 \),
\[
g^{(m-1)\frac{\rho}{1-\rho}} \leq \frac{\kappa(m) + \epsilon}{f} < g^{m\frac{\rho}{1-\rho}}
\]
which satisfies the condition that the marginal entrant lives to the age of \( m \). Hence, \( \hat{n}(\kappa) \) is an increasing step function of \( \kappa \).

\[ \square \]

Proof of Proposition 2.4. We know that
\[
P_t(\kappa)^{-\gamma} = g^{\gamma t[(1 - \rho)(L + \Pi(\kappa))]^{\frac{\rho(1 - \rho)}{\rho - \rho}}\rho^{\gamma} \frac{\gamma(1 - \rho)}{\gamma(1 - \rho) - \rho}}\mu^{\omega(\kappa)}
\]
\[
M(\kappa) = k^{\frac{\gamma(1 - \rho)}{\rho}} [L + \Pi(\kappa)] \frac{\gamma(1 - \rho) - \rho}{\gamma} \mu^{-1}\omega(\kappa)^{-1}
\]
\[
Y_t(\kappa) = [L + \Pi(\kappa)] P_t(\kappa, \theta)^{-1}
\]
\[
\Pi(\kappa) = \frac{L \xi(\kappa)}{1 - \xi(\kappa)}
\]
where

\[
\xi(\kappa) = 1 - \rho - \frac{\gamma(1 - \rho) - \rho}{\gamma \omega(\kappa)} \left( \kappa^{\frac{1 - \rho}{\rho} \gamma} \sum_{i=1}^{\hat{n}(\kappa)} \kappa_i + f^{\frac{\rho - \gamma(1 - \rho)}{\rho}} \sum_{i=\hat{n}(\kappa)+1}^{N} g^{\gamma(1-i)} \right)
\]

\[
\omega(\kappa) = \kappa^{\frac{1 - \rho}{\rho} \gamma} \sum_{i=1}^{\hat{n}(\kappa)} g^{(1-i) \frac{\rho}{1-\rho}} + f^{1 - \frac{1 - \rho}{\rho} \gamma} \sum_{i=\hat{n}(\kappa)+1}^{N} g^{\gamma(1-i)}
\]

It suffices to show that \(\xi(\kappa)\) and \(\omega(\kappa)\) are continuous in \(\kappa\) at the jump points.

For \(\omega(\kappa)\), we know that \(\omega(\kappa) = \lim_{\kappa \to 0_+} \omega(\kappa(m) - \epsilon), \forall n = 2, \ldots, N\) since

\[
\omega(\kappa(m)) = \kappa(m)^{\frac{1 - \rho}{\rho} \gamma} \sum_{i=1}^{m} g^{(1-i) \frac{\rho}{1-\rho}} + f^{1 - \frac{1 - \rho}{\rho} \gamma} \sum_{i=m+1}^{N} g^{\gamma(1-i)}
\]

\[
= \kappa(m)^{\frac{1 - \rho}{\rho} \gamma} \sum_{i=1}^{m-1} g^{(1-i) \frac{\rho}{1-\rho}} + f^{1 - \frac{1 - \rho}{\rho} \gamma} \sum_{i=m}^{N} g^{\gamma(1-i)}
\]

\[
= \lim_{\epsilon \to 0_+} \omega(\kappa(m) - \epsilon)
\]

For \(\xi(\kappa)\), we know that \(\xi(\kappa) = \lim_{\kappa \to 0_+} \xi(\kappa(m) - \epsilon), \forall n = 2, \ldots, N\) since

\[
\xi(\kappa(m)) = 1 - \rho - \frac{\gamma(1 - \rho) - \rho}{\gamma \omega(\kappa(m))} \left( \kappa(m)^{\frac{1 - \rho}{\rho} \gamma} \sum_{i=1}^{m} \kappa_i + f^{\frac{\rho - \gamma(1 - \rho)}{\rho}} \sum_{i=m+1}^{N} g^{\gamma(1-i)} \right)
\]

\[
= 1 - \rho - \frac{\gamma(1 - \rho) - \rho}{\gamma \omega(\kappa(m))} \left( \kappa(m)^{\frac{1 - \rho}{\rho} \gamma} \sum_{i=1}^{m-1} \kappa_i + g^{\gamma(1-m)} f^{\frac{\rho - \gamma(1 - \rho)}{\rho}} \sum_{i=m+1}^{N} g^{\gamma(1-i)} \right)
\]

\[
= 1 - \rho - \frac{\gamma(1 - \rho) - \rho}{\gamma \omega(\kappa(m))} \left( \kappa(m)^{\frac{1 - \rho}{\rho} \gamma} \sum_{i=1}^{m-1} \kappa_i + f^{\frac{\rho - \gamma(1 - \rho)}{\rho}} \sum_{i=m}^{N} g^{\gamma(1-i)} \right)
\]

\[
= \lim_{\epsilon \to 0_+} \xi(\kappa(m) - \epsilon)
\]

\(\square\)