Essays on International Trade

by

Sergii Meleshchuk

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Economics in the Graduate Division of the University of California, Berkeley

Committee in charge:

Professor Andrés Rodríguez-Clare, Chair
Assistant Professor Benjamin Faber
Assistant Professor Thibault Fally

Spring 2018
Essays on International Trade

Copyright 2018
by
Sergii Meleshchuk
Abstract

Essays on International Trade

by

Sergii Meleshchuk

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Andrés Rodríguez-Clare, Chair

This dissertation investigates the role of firms in international trade using the recently available comprehensive datasets on international trade transactions of firms in different countries. The first chapter investigates empirically second-degree price discrimination in business-to-business dealings. I use highly detailed transaction-level Colombian imports data to document the presence of quantity discounts. These data helps me identify firms on both ends of the transaction. I document a robust negative relationship between prices and quantities of the same good purchased from a given seller by different Colombian buyers. This relationship is not an artifact of the measurement error or the scope for the quality differentiation of particular goods. I also show that this result is not driven by the length of the relationship and between a given buyer and a given seller or the number of transactions between a given buyer and a given seller per year. Finally, the negative relationship between prices and quantities holds when I condition my estimation on the measure of buyers’ and sellers’ market power. I find that, on average, a 10% increase in the quantity purchased reduces the price charged per unit by 2% and conclude that this relationship is most likely consistent with price discrimination rather than alternative explanations.

To rationalize this empirical fact, the second chapter of this dissertation develops a tractable theoretical framework that embeds nonlinear pricing (second-degree price discrimination) into a standard international trade model and characterizes optimal policies. The model can be conveniently aggregated and yields a standard gravity equation of international trade flows. However, the model has several important implications that are absent from the traditional trade models. I show that welfare losses from second-degree price discrimination can be quite substantial. I also characterize optimal trade policy from the perspective of a social planner in a small open economy. I prove that optimal tariffs are higher when firms use non-linear prices as compared to standard models. In addition, if the policymaker sets tariffs that are optimal under linear pricing, but firms use second-degree price discrimination, this will lead to significant welfare losses.

The third chapter focuses on the the importance of the extensive margin (the number of firms exporting) for trade flows, which is highlighted by the Melitz model of international
trade. Using the World Bank’s *Exporter Dynamics Database* featuring firm-level exports from 50 countries, it documents that around 50% of variation in exports does occur on the extensive margin — a quantitative victory for the Melitz framework. The remaining 50% on the intensive margin (exports per exporting firm) contradicts a special case of Melitz with Pareto-distributed firm productivity, which has become a tractable benchmark. This benchmark model predicts that, conditional on the fixed costs of exporting, *all* variation in exports across trading partners will occur on the extensive margin. Combining Melitz with lognormally-distributed firm productivity and firm-destination fixed trade costs can explain the intensive margin seen in the EDD data. In the EDD, the importance of the intensive margin rises steadily when going from the smallest to largest exporting firms across source countries, as is also predicted by the Melitz model with lognormally-distributed productivity.
To my family.
Contents

1 Price Discrimination in International Trade: Empirical Evidence 1
   1.1 Introduction ................................................. 1
   1.2 Data .................................................. 3
   1.3 Detecting price discrimination in the data ..................... 5
   1.4 Conclusions .............................................. 12
   1.5 Figures ................................................ 13
   1.6 Tables ................................................. 14

2 Price discrimination in International Trade: Theory 19
   2.1 Introduction .............................................. 19
   2.2 Trade model with second-degree price discrimination .......... 21
   2.3 Quantification .......................................... 34
   2.4 Conclusions ............................................ 37
   2.5 Figures ............................................... 38
   2.6 Tables ............................................... 43

3 The Intensive Margin in Trade 44
   3.1 Introduction ............................................. 44
   3.2 The Intensive Margin in the Data .......................... 47
   3.3 The Intensive Margin in the Melitz-Pareto Model ............. 52
   3.4 The Intensive Margin in the Melitz-Lognormal Model ....... 63
   3.5 Conclusion ............................................. 74
   3.6 Figures ............................................... 76
   3.7 Tables ............................................... 91

A Appendices to Chapter 1 98
   A.1 Data description ......................................... 98
Appendices to Chapter 2
B.1 Solving the Problem of Intermediate Good Producers 102
B.2 Gravity 104
B.3 Gains from Trade 105
B.4 Welfare in a Closed Economy 107
B.5 Optimal policies 111
B.6 Hat algebra 116

Appendices to Chapter 3
C.1 Estimation of $\bar{\theta}$ 117
C.2 QQ-Estimation of $\sigma_\phi$ 118
C.3 Quasi-Bayesian Estimation for the full Melitz-lognormal model 119
C.4 MCMC algorithm 121
C.5 Convergence of the Monte Carlo Markov chains 123
C.6 Figures 124
C.7 Tables 158

Bibliography 164
# List of Figures

1.1 Non-parametric estimate of conditional expectation function of log prices . . . . 13  
2.1 Relative Welfare in Autarky . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38  
2.2 Optimal taxes in autarky . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39  
2.3 Gains from changing taxes from 0 to the optimal level in autarky . . . . . . . 40  
2.4 Optimal taxes in a small open economy . . . . . . . . . . . . . . . . . . . . . . . 41  
2.5 Gains from changing taxes from 0 to the optimal level . . . . . . . . . . . . . 42  
3.1 Intensive and Extensive margins of exporting . . . . . . . . . . . . . . . . . . . . 76  
3.2 Intensive and Extensive margins of exporting, by industry . . . . . . . . . . . . 77  
3.3 IME for each percentile, data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 78  
3.4 Manufacturing absorption and averaged exports per firm (destination fixed effects) 78  
3.5 Exports to largest destination and market entry . . . . . . . . . . . . . . . . . . . 79  
3.6 Model-implied fixed and variable trade costs and distance . . . . . . . . . . . . 80  
3.7 Fixed product-level trade costs and distance . . . . . . . . . . . . . . . . . . . . . 81  
3.8 IME for each percentile, Pareto and granularity . . . . . . . . . . . . . . . . . . . 82  
3.9 Number of firms and population . . . . . . . . . . . . . . . . . . . . . . . . . . . 83  
3.10 IME for each percentile, lognormal . . . . . . . . . . . . . . . . . . . . . . . . . 84  
3.11 Fixed and variable trade costs and distance, lognormal . . . . . . . . . . . . . 85  
3.12 Full lognormal model, pdf of log sales . . . . . . . . . . . . . . . . . . . . . . . . 86  
3.13 Share of firms selling to destination X but not to destination Y . . . . . . . . . 86  
3.14 Correlation between log exports to top destinations . . . . . . . . . . . . . . . 87  
3.15 IME for each percentile, data and full lognormal model . . . . . . . . . . . . . 87  
3.16 Gains from trade liberalization . . . . . . . . . . . . . . . . . . . . . . . . . . . . 88  
3.17 Counterfactual changes in trade flows . . . . . . . . . . . . . . . . . . . . . . . . 89  
3.18 Counterfactual changes in trade flows, $\lambda_{ij} > 1\%$ . . . . . . . . . . . . 90  
C.1 Summary plots for QBE chains: $d_{i1}$ . . . . . . . . . . . . . . . . . . . . . . . . . 124  
C.2 Summary plots for QBE chains: $d_{i2}$ . . . . . . . . . . . . . . . . . . . . . . . . . 125  
C.3 Summary plots for QBE chains: $d_{i3}$ . . . . . . . . . . . . . . . . . . . . . . . . . 126  
C.4 Summary plots for QBE chains: $d_{i4}$ . . . . . . . . . . . . . . . . . . . . . . . . . 127  
C.5 Summary plots for QBE chains: $d_{i5}$ . . . . . . . . . . . . . . . . . . . . . . . . . 128
C.6 Summary plots for QBE chains: $d_{i6}$ ................................. 129
C.7 Summary plots for QBE chains: $d_{i7}$ ................................. 130
C.8 Summary plots for QBE chains: $d_{i8}$ ................................. 131
C.9 Summary plots for QBE chains: $d_{i9}$ ................................. 132
C.10 Summary plots for QBE chains: $d_{i10}$ .............................. 133
C.11 Summary plots for QBE chains: $d_{i11}$ ............................. 134
C.12 Summary plots for QBE chains: $d_{i12}$ ............................. 135
C.13 Summary plots for QBE chains: $d_{i13}$ ............................. 136
C.14 Summary plots for QBE chains: $d_{i14}$ ............................ 137
C.15 Summary plots for QBE chains: $d_{i15}$ ............................ 138
C.16 Summary plots for QBE chains: $\mu_{f,i1}$ .......................... 139
C.17 Summary plots for QBE chains: $\mu_{f,i2}$ .......................... 140
C.18 Summary plots for QBE chains: $\mu_{f,i3}$ .......................... 141
C.19 Summary plots for QBE chains: $\mu_{f,i4}$ .......................... 142
C.20 Summary plots for QBE chains: $\mu_{f,i5}$ .......................... 143
C.21 Summary plots for QBE chains: $\mu_{f,i6}$ .......................... 144
C.22 Summary plots for QBE chains: $\mu_{f,i7}$ .......................... 145
C.23 Summary plots for QBE chains: $\mu_{f,i8}$ .......................... 146
C.24 Summary plots for QBE chains: $\mu_{f,i9}$ .......................... 147
C.25 Summary plots for QBE chains: $\mu_{f,i10}$ .......................... 148
C.26 Summary plots for QBE chains: $\mu_{f,i11}$ ........................ 149
C.27 Summary plots for QBE chains: $\mu_{f,i12}$ ........................ 150
C.28 Summary plots for QBE chains: $\mu_{f,i13}$ ........................ 151
C.29 Summary plots for QBE chains: $\mu_{f,i14}$ ........................ 152
C.30 Summary plots for QBE chains: $\mu_{f,i15}$ ........................ 153
C.31 Summary plots for QBE chains: $\bar{\sigma}_{\varphi,i}$ .................. 154
C.32 Summary plots for QBE chains: $\sigma_{\alpha,i}$ ...................... 155
C.33 Summary plots for QBE chains: $\sigma_{f,i}$ ......................... 156
C.34 Summary plots for QBE chains: $\rho_i$ .............................. 157
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Descriptive statistics: full sample</td>
<td>14</td>
</tr>
<tr>
<td>1.2</td>
<td>Descriptive statistics: baseline sample</td>
<td>14</td>
</tr>
<tr>
<td>1.3</td>
<td>OLS results</td>
<td>15</td>
</tr>
<tr>
<td>1.4</td>
<td>IV results: $\ln q_{ieg,t}$ instrumented with $\ln q_{ieg,t-1}$</td>
<td>15</td>
</tr>
<tr>
<td>1.5</td>
<td>IV results: $\ln q_{ieg,t}$ instrumented with $\ln q_{ieg,t+1}$</td>
<td>16</td>
</tr>
<tr>
<td>1.6</td>
<td>OLS results, controlling for expenditure and revenue shares</td>
<td>16</td>
</tr>
<tr>
<td>1.7</td>
<td>Unit values and size of import</td>
<td>17</td>
</tr>
<tr>
<td>1.8</td>
<td>OLS results: interactions with measures of quality differentiation and market organization</td>
<td>18</td>
</tr>
<tr>
<td>2.1</td>
<td>Allocation under social planner and decentralized equilibrium</td>
<td>43</td>
</tr>
<tr>
<td>3.1</td>
<td>Core Sample of EDD countries+China, years firm-level data is available</td>
<td>91</td>
</tr>
<tr>
<td>3.2</td>
<td>IME regressions, core sample</td>
<td>92</td>
</tr>
<tr>
<td>3.3</td>
<td>IME regression, small firms excluded, core sample</td>
<td>93</td>
</tr>
<tr>
<td>3.4</td>
<td>IME regression, disaggregated within manufacturing, core sample</td>
<td>94</td>
</tr>
<tr>
<td>3.5</td>
<td>Product-level IME regression, core sample</td>
<td>94</td>
</tr>
<tr>
<td>3.6</td>
<td>Margins of trade and distance</td>
<td>95</td>
</tr>
<tr>
<td>3.7</td>
<td>Trade costs and distance</td>
<td>95</td>
</tr>
<tr>
<td>3.8</td>
<td>Fixed trade costs distance elasticity and granularity</td>
<td>95</td>
</tr>
<tr>
<td>3.9</td>
<td>IME under granularity</td>
<td>96</td>
</tr>
<tr>
<td>3.10</td>
<td>Number of firms and population</td>
<td>96</td>
</tr>
<tr>
<td>3.11</td>
<td>Trade costs and distance, lognormal</td>
<td>96</td>
</tr>
<tr>
<td>3.12</td>
<td>Estimates of dispersion, full lognormal model</td>
<td>97</td>
</tr>
<tr>
<td>3.13</td>
<td>Implied IME in full lognormal model</td>
<td>97</td>
</tr>
<tr>
<td>3.14</td>
<td>Implied trade costs in full lognormal model</td>
<td>97</td>
</tr>
<tr>
<td>A.1</td>
<td>OLS results, sample from Table 1.4</td>
<td>100</td>
</tr>
<tr>
<td>A.2</td>
<td>OLS results, sample from Table 1.5</td>
<td>100</td>
</tr>
<tr>
<td>A.3</td>
<td>Additional OLS results in different subsamples</td>
<td>101</td>
</tr>
<tr>
<td>C.1</td>
<td>Additional EDD countries and years in the Extended Sample (only EDD statistics available)</td>
<td>158</td>
</tr>
</tbody>
</table>
C.2 IME regressions, extended sample ...................................................... 159
C.3 IME regression by income group ............................................................ 160
C.4 IME regression, small firms excluded, extended sample ....................... 161
C.5 IME regressions, IV .............................................................................. 161
C.6 IME regression, disaggregated by HS2 product, extended sample .......... 162
C.7 IME regression, dropping industries with large share of intermediaries .... 162
C.8 Estimates of $\bar{\theta}$ ........................................................................... 163
C.9 QQ estimates of $\bar{\sigma}_\phi$ ................................................................. 163
Acknowledgments

I am grateful to Andres Rodriguez-Clare, Benjamin Faber, Cecile Gaubert for their support, encouragement and invaluable criticism. Thanks also to Yuriy Gorodnichenko, Thibault Fally, Nick Sander, Matthias Hoelzlein, and Marc Dordal Carreras for valuable comments, as well as seminar participants at Berkeley and NBER ITI. Support from the Clausen Center and the Berkeley Graduate Division is gratefully acknowledged.
Chapter 1

Price Discrimination in International Trade: Empirical Evidence

1.1 Introduction

Most models in contemporary international economics and macroeconomic literature make the assumption that sellers charge all buyers the same price for the same good in a specific geographic market. In reality, however, a lot of business-to-business transactions occur in decentralized markets.\textsuperscript{1} Sellers of intermediate inputs have the power to vary prices across their buyers even if marginal costs of production are the same. In other words, sellers price-discriminate buyers. Anecdotally, quantity discounts (or second-degree price discrimination) are a common form of price discrimination in business-to-business transactions.\textsuperscript{2} However, welfare consequences of quantity discounts in general equilibrium models, as well as optimal policies in this setting remain largely understudied.

In an attempt to shed light on this matter, I make three contributions. First, I provide evidence of price discrimination in the transaction-level Colombian imports data. Second, I develop a tractable theoretical framework that embeds second-degree price discrimination into a widely-used class of gravity models in international trade and characterize optimal

\textsuperscript{1}Rauch (1999) finds that firms in at least 2/3 of the US industries do not sell their output on organized exchanges or have reference prices for their goods

\textsuperscript{2} There is ample anecdotal evidence that these discounts are omnipresent in business-to-business transactions. In a field study of 49 companies, Munson and Rosenblatt (1998) find that over 80% of buyers receive quantity discounts on most of the goods that they purchase. Over half of managers in supplier firms report that they use quantity discounts to boost annual demand for their goods. Jackson Jr (2015) gets similar results in another field study. Best Buy, Costco, Home Depot, Orchard Supply Hardware, and many other chains have quantity discount programs for large orders. Recognizing the importance of the practice, Amazon allows the sellers to use tiered pricing discounts on its Amazon Business platform. Leading consulting firms highlight the importance of quantity discounts in their reports. According to Boston Consulting Group:

“In most B2B industries, discounts represent a company’s largest marketing investment, often amounting to 30% or more of list-price sales. [...] flexible discounters vary discounts on the basis of common characteristics, such as sales or order volume, channel, or customer segment”
policies. Third, I calibrate the model using Colombian transaction-level data and use the calibrated model for welfare calculations. The first contribution is analyzed in Chapter 1 of the dissertation, while the second and the third ones are discussed in Chapter 2.

In this Chapter, I use Colombian transaction-level microdata to document a new set of stylized facts about price discrimination in international trade. The main goal of my empirical exercise is to estimate the supply-side elasticity of prices with respect to quantities of goods purchased. Using the unique richness of my data, I estimate this elasticity within narrowly defined exporting firm × good cells.

In my baseline specification, the elasticity of prices with respect to quantities is approximately -0.2. That is, a 10% increase in quantity purchased reduces the price of a good produced by a particular exporter by 2%. The value of the estimated coefficient is robust across subsamples. These findings are consistent with second-degree price discrimination, a practice that allows firms to charge lower price when the volume of the purchase is higher. In a number of robustness checks, I rule out several competing explanations of the negative elasticity, such as non-classical measurement error and quality differentiation within narrowly defined exporter-goods categories.

This paper is related to several strands of research. First, this paper is related to the literature modeling the welfare effects of price discrimination in input markets. Traditionally, welfare implications of price discrimination in intermediate inputs have been viewed through the prism of third-degree price discrimination in industries with a fixed number of firms, as in Katz (1987), DeGraba (1990), and Yoshida (2000). In contrast, I model endogenous entry and production decision in both sectors and show that these decisions are distorted in the presence of price discrimination in an open economy.

Second, this paper contributes to a rapidly developing literature on how import status influences firms (Halpern, Miklós Koren, and Szeidl (2015), Blaum, Lelarge, and Peters (2015), Antrás, Fort, and Tintelnat (2017), Bernard, Moxnes, and Ulltveit-Moe (2017)). This literature assumes uniform (linear) pricing. In contrast, I extend the model by allowing intermediate good producers to engage in second-degree price discrimination.

Third, similar to Kugler and Verhoogen (2011) and Bastos, J. Silva, and Verhoogen (2016), I document heterogeneity in input prices across importers. These papers find a positive relationship between input prices and the size of an importer and interpret price differences as a difference in the quality of the intermediate inputs. I replicate these results using total imports as a proxy for the size of the firm. Relative to these papers, I use much more detailed data on the identities of the exporters of intermediate inputs. Controlling for exporter-good fixed effects allows me to soak up variation in quality and marginal costs of production of the good at a very granular level. I find a negative elasticity of prices

---

3While second-degree price discrimination has been well documented in numerous industries producing consumer goods, the empirical literature on price discrimination in business-to-business relationships is relatively small due to the lack of reliable and detailed data. Grennan (2013) documents patterns of price discrimination in the market for coronary stents in the US and Beckert, Smith, and Takahashi (2015) do it for the brick industry in the UK.
with respect to quantities and interpret this relationship from the perspective of *quantity discounts*.

Fourth, the aggregate effects of trade in my model are in line with the literature on quantitative trade models. The model features gravity equation governing trade flows and gains from trade formula that depends on the share of expenditures on domestic inputs and the trade elasticity, as in Arkolakis, Costinot, and Rodríguez-Clare (2012). In the spirit of some results in Arkolakis, Costinot, Donaldson, et al. (2017), the distortion introduced by second-degree price discrimination doesn’t affect the expression for the gains from trade. However, it does influence optimal policy. Following Demidova and Rodríguez-Clare (2009) and Costinot, Rodríguez-Clare, and Werning (2016), I use a primal approach to derive the optimal policy for the settings with different market structures.

Finally, in regards to data, this paper is closely related to Bernard and Dhingra (2015). They use Colombian transaction-level data to study optimal contracts between importers and exporters and develop a model that also features buyer-specific prices. My paper differs from theirs in several respects. First, I provide direct evidence of the relationship between prices and quantities in the data. Second, I study aggregate gains from trade, rather than the division of gains between importers and exporters. Third, I characterize optimal policies.

### 1.2 Data

I examine the relationship between prices charged by exporters and quantity purchased by importers using transaction-level customs data from Colombia. The dataset spans nine years between 2007 and 2015. Each record in the dataset corresponds to the individual transaction and contains information on the date of the transaction, total value in US dollars, quantity of goods purchased, a 10-digit good code, the name and tax number of the importer, the name and the country of the exporter. In addition, each transaction records the name of importer industry, that can be matched to the 4-digit ISIC Rev. 3.1 and ISIC Rev. 4 industry classifications. Using this dataset, I construct the following variables at the importer-exporter-HS10 good-year (*iegτ*) level:

- **Year**: year of transaction.
- **Importer**: Colombian tax number.

---

---

4 Eaton, Jinkins, et al. (2016) use similar dataset to explore sellers’ search of potential buyers

5 The data is provided by private company Datamyne. This company has also provided Argentinian customs-level data to Gopinath and Neiman (2014).

6 The dataset used in the analysis represents all of the imports transactions for the years 2007-2015. Annual aggregate imports (CIF value), calculated for each of the HS6 goods tracks total imports reported in the COMTRADE database extremely closely. The regression of annual imports in the COMTRADE data against total imports in the transaction-level dataset has the slope coefficient of 0.992 and the $R^2$ above 0.99.

7 The first 6 digits identify the good according to the Harmonized Commodity Description and Coding System (HS), the next 2 digits identify the NANDINA subheading and the last two digits identify Colombian customs agency subheading.
CHAPTER 1. PRICE DISCRIMINATION IN INTERNATIONAL TRADE: EMPIRICAL EVIDENCE

- Quantity: total quantity over all transactions within $iegt$ cell.
- FOB unit values in US dollars (prices): total FOB value over all transactions divided by total quantity within $iegt$ cell.
- HS 10 good.
- Exporter: identified by name.
- Importer industry: ISIC rev 3.1 4-digit industry.

I provide more details on how I construct the variables in the Online Appendix A.1. Panel a) of Tables 1.1 and 1.2 reports descriptive statistics on the counts of importers, exporters, and goods in the full sample, and in the baseline sample that is restricted to importers that are not in retail or wholesale industries. In the full sample, there are over 30,000 different importers per year, which import from approximately 100,000 different exporters. The average number of different HS10 goods across years is around 6,400. My identification strategy will rely on the fact that are sometimes several buyers of a certain HS10 good produced by a certain exporter. Panel b) of Tables 1.1 and 1.2 reports statistics on the number of importers that buy a certain good from a certain exporter. The average number of importers per exporter-good-year cell is 1.3 in the full sample and 1.2 in the baseline sample. Most of the exporters sell their goods to a single importer and are thus omitted from the latter analysis.

For robustness checks I use the following set of controls from Kugler and Verhoogen (2011):

- RnD and advertising expenditures intensity. This variable measures a share of RnD and advertising expenditures in total sales based on the U.S. US Federal Trade Commission Line of Business Survey. It is constructed for a subset of 4-digit ISIC industries. The measure reflects the scope for vertical differentiation: a higher RnD and advertising intensity reflects a bigger scope for vertical differentiation.

- Gollop-Monahan input dissimilarity index. The larger the value of index - the more dispersion there is in input bundles across plants in a certain sector, and so potentially the more differentiated their output is. The measure is available at the level of 4-digit ISIC industries.

- Rauch index. Kugler and Verhoogen (2011) take Rauch (1999) “liberal” measure of the classification of industries at the SITC 4-digit level, and assign a value 0 if its output is traded on organized exchanges or possesses a “reference price” and assign a value 1 for all other industries. Then they match SITC industries to the ISIC rev. 2 4-digit industries using concordance from OECD.

I match these variables to the HS6 goods classification. The RnD and advertising expenditures intensity measure and the Gollop-Monahan index are normalized to be mean zero.
with unit standard deviation across 4-digit industries. Rauch index varies between 0 (more centralized industries) to 1 (less decentralized industries).

1.3 Detecting price discrimination in the data

In this section I estimate the supply-side elasticity of prices with respect to quantities. I sketch an empirical model in which prices and quantities are determined by a combination of demand, marginal costs, and quality shocks. I then discuss my strategy to identify this elasticity using variation in demand shocks.

Empirical model

To discipline the estimation, consider the following firm-level demand function for intermediate inputs:

\[ q_{iegt} = D(p_{iegt}, \nu_{iegt}, \epsilon_{iegt}), \quad (1.1) \]

where:

- \( i, e, g, t \) index importer, exporter, HS 10, and year respectively;
- \( q_{iegt} \) is the quantity of the good purchased;
- \( p_{iegt} \) is the free on board price paid by the importer \( i \) to the exporter \( e \) for the good \( g \) at time \( t \);
- \( \nu_{iegt} \) is the quality of the input;
- \( \epsilon_{iegt} \) is a demand shock.

Importers choose how many units of a good to buy based on its price and two shocks: the quality of the good, and the demand shock. The demand shock can incorporate a broad combination of variables, such as firm-level productivity, demand on its output, purchases of other intermediate goods, import tariffs, freight and insurance costs (which are not included into the FOB price), idiosyncratic preferences for certain goods produced by certain exporters (that are unrelated to quality), etc. I consider the following log-log approximation of equation (1.1):

\[ \ln q_{iegt} = \alpha \ln p_{iegt} + \ln \nu_{iegt} + \ln \epsilon_{iegt}. \quad (1.2) \]

This specification can arise, for example, in a setting with CES production function over intermediate inputs.

On the supply side, exporters charge a price which is defined to be a product of marginal costs and markups:

\[ p_{iegt} = MC(Q_{eg}, \nu_{iegt}, \varphi_{iegt}) \Psi(q_{iegt}), \quad (1.3) \]

where:
• \( MC \) stands for the marginal cost function;
• \( Q_{egt} \) is the exporter e total production of good g sold to Colombia in year t: \( Q_{egt} = \sum_i q_{iegt} \);
• \( \varphi_{egt} \) is the exporter-good-time cost shock;
• \( \Psi \) denotes the markup function;

In this general setup, marginal costs are allowed to vary with total quantity sold to all importers, quality of the good sold to a particular importer, and some exporter-good-time productivity (or cost) shock \( \varphi_{egt} \). Departing from the models with perfect, monopolistic, or oligopolistic competition, in which exporters do not vary their markups across imports, I allow markups to vary with the quantity sold to a specific importer.

Consider the following parametric specification of equation (1.3):

\[
\ln MC(Q_{egt}, \nu_{iegt}, \varphi_{egt}) \equiv \gamma \ln Q_{egt} + \rho \ln \nu_{iegt} + \ln \varphi_{egt} \quad (1.4)
\]

\[
\ln \Psi(q_{iegt}) \equiv \beta \ln q_{iegt} + \delta. \quad (1.5)
\]

This specification nests the predominant majority of trade models featuring monopolistic competition or oligopolistic market structure in which firms charge the same price for the same good they sell to different buyers. The demand-supply system of equations (1.1) and (1.3) can be expressed as:

\[
\ln q_{iegt} = -\alpha \ln p_{iegt} + \ln \nu_{iegt} + \ln \epsilon_{iegt} \quad (1.6)
\]

\[
\ln p_{iegt} = \delta + \beta \ln q_{iegt} + \gamma \ln Q_{egt} + \rho \ln \nu_{iegt} + \ln \varphi_{egt}, \quad (1.7)
\]

The main goal of this section is to estimate the parameter \( \beta \), which reflects the elasticity of markups with respect to the quantity of the goods sold to a specific buyer. In the models with competitive, monopolistic or oligopolistic markets, we expect \( \beta = 0 \) (i.e. the price of good is the same across buyers). For expositional purposes, it is useful to express the system of equations (1.6)-(1.7) in the 'reduced form':

\[
\ln q_{iegt} = \frac{-\alpha (\delta + \gamma \ln Q_{egt} + \ln \varphi_{egt}) + (1 - \alpha \rho) \ln \nu_{iegt} + \ln \epsilon_{iegt}}{1 + \alpha \beta} \quad (1.8)
\]

\[
\ln p_{iegt} = \frac{(\delta + \gamma \ln Q_{egt} + \ln \varphi_{egt}) + (\rho + \beta) \ln \nu_{iegt} + \beta \ln \epsilon_{iegt}}{1 + \alpha \beta} \quad (1.9)
\]

**Identification**

Identification of \( \beta \) is subject to classical endogeneity concerns. For example, a naive regression of log prices on log quantities trying to estimate \( \beta \) from equation (1.7) will yield a biased coefficient since log quantities are not independent of the error term. For example, as can be seen from the equation (1.8), quantity \( q_{iegt} \) is not independent of the quality shock that
will enter error term in this expression. Below, I explore several identifying strategies that deal with such concerns as simultaneity bias and non-classical measurement error.

Let

\[ q_{iegt} = \ln q_{iegt} - \tilde{E}[\ln q_{iegt}|X_{iegt}] \]  
\[ p_{iegt} = \ln p_{iegt} - \tilde{E}[\ln p_{iegt}|X_{iegt}], \]

where \( X_{iegt} \) is the matrix of control variables, and \( \tilde{E}[\cdot|\cdot] \) is a linear conditional expectation function. Define:

\[ \beta_{OLS} \equiv \frac{COV(\hat{q}_{iegt}, \hat{p}_{iegt})}{VAR(\hat{q}_{iegt})} \]

By the Frisch-Waugh-Lovell theorem, \( \hat{\beta} \) is the coefficient from a regression of \( \ln p_{iegt} \) on \( \ln q_{iegt} \) when controlling for \( X_{iegt} \).

To identify the parameter \( \beta \), I will rely on the following assumption:

\[ E[\nu_{iegt}^2|X_{iegt}] = 0. \]  
\[ (1.12) \]

Under identifying assumption (1.12):

\[ \beta_{OLS} = \beta \]

Selecting a matrix of control variables \( X_{iegt} \) is critical for this identification strategy. I explore the richness of the Colombian data to ensure that equation (1.12) holds.

The first-order concern that can lead to violation of the equation (1.12) is variation in marginal costs. Higher marginal costs increase price and reduce demand, thus leading to a downward bias in the OLS estimate. In the baseline specification, I deal with this issue by including a set of exporter-HS10 good-year fixed effects as an \( X_{iegt} \) control matrix. This matrix absorbs exporter-good marginal cost shocks at a very granular level. The identification of \( \beta \) in this specification will come from the cross-sectional variation in unit values and quantity within the fixed effects cells.

Variation in quality within exporter-good cells is another potential source of bias. The baseline specification allows quality to vary across different exporters within HS10 good categories, as well as within exporters across different HS10 categories, but not across different buyers within the HS10 good produced by a certain exporter. Although the baseline approach already controls for quality variation at a granular exporter-good level, OLS estimate is biased if there is potential for quality differentiation within exporter-good cells (if there are "hidden varieties" within exporter-HS10 good cells). To address this issue, I augment the matrix of controls with the importer-year and importer-exporter-good fixed effects. Both sets of fixed effects allow for some degree of quality differentiation within exporter-good cells without making OLS estimate biased. Importer-year fixed effects allow importers to purchase goods of higher quality from some exporter, as long as they also purchase goods of
higher quality from the other exporters. This deals with, for example, importer-level shocks that make them purchase goods of higher quality (this mechanism is present in Kugler and Verhoogen (2011) and I. Brambilla, Lederman, and Porto (2012)). Specification with the importer-exporter-good fixed effects assume that importers buy the same variety of the HS10 good produced by the same exporter across time. In this case identification comes from the panel nature of the data. Loosely speaking, time-series variation allows me to purge the data of importer-exporter-good fixed effects. After that, $\beta$ is identified from the cross-sectional variation within exporter-good-year cells. The combination of exporter-good-year, importer-exporter, and importer-exporter-good fixed effects allows me to control for most sources of quality heterogeneity studied in the previous literature. Including additional sets fixed effects, however, poses a trade-off. On the one hand, they relax the assumptions on the degree of quality differentiation within varieties under which OLS estimates are unbiased. On the other hand, controlling for importer-level fixed effects also removes some of the variation in the demand shocks that is necessary to identify $\beta$. Moreover, due to the sparse nature of the data, additional fixed effects can considerably reduce the sample size.

The direction of the bias in the OLS estimate induced by hidden varieties is not obvious: depending on the underlying parameters it can go either way. Most of the papers in trade literature that focus on the quality of intermediate inputs find that quantity and quality of intermediate inputs comove across different importers: bigger firms on average buy more inputs and those inputs are of higher quality (see, for example, Kugler and Verhoogen (2011), Bastos, J. Silva, and Verhoogen (2016), I. Brambilla, Lederman, and Porto (2012), D. L. Brambilla I. and Porto (2016)). One of the ways to rationalize this relationship theoretically is through a model in which quantity and quality of intermediate inputs are complements in the production function. Extrapolating this logic to quality differentiation within exporter-good cells, one can conclude that omitted quality differentiation is likely to bias the OLS coefficient up. Thus, if I find that the OLS estimate is negative, and there are still concerns that the fixed effects fail to control for quality shocks, conventional wisdom would imply that the true elasticity is a more negative number. In the robustness subsection below I explore several indirect ways to estimate the direction of the bias.

**Robustness checks**

**Measurement error.** Since the customs data do not provide an independent measure of prices, they have to be inferred from unit values. Unit values are constructed as the ratio of total import value and quantity. If there is measurement error in quantity, it will appear in the unit values as well by construction. Let $\ln q_{ietg}$ and $\ln p_{ietg}$ stand for the observed quantities and unit values, and let

\[
\ln q'_{ietg} = \ln q_{ietg} + \xi_{ietg} \\
\ln p'_{ietg} = \ln p_{ietg} - \xi_{ietg}
\]

where $\xi_{ietg}$ is a classical measurement error that is independent of other shocks and is serially uncorrelated. Under these conditions, even if identifying assumption (1.12) holds, the OLS
coefficient from a regression of observed log unit values on log quantities, controlling for exporter-good-year fixed effects will be biased downwards:

$$\beta_{OLS}' = \beta - \frac{\sigma_{\xi \xi}}{\sigma_{\epsilon \epsilon} \left(\frac{1}{1+\alpha \beta}\right)^2 + \sigma_{\xi \xi}} < \beta,$$

where $\sigma_{\xi \xi}$ is the variance of the measurement error and $\sigma_{\epsilon \epsilon}$ is the variance of the demand shocks after controlling for fixed effects. To check whether measurement error is really an issue, I instrument $\ln q_{iegt}$ with a lagged value $\ln q_{iegt-1}$ or a lead value $\ln q_{iegt+1}$. These instruments satisfy the exclusion restriction if measurement error is serially uncorrelated.\footnote{I assume there is no measurement error in total payment $\ln p_{iegt} + \ln q_{iegt}$. If in fact I observe $\ln p_{iegt}' + \ln q_{iegt}' = \ln p_{iegt} + \ln q_{iegt} + \chi_{iegt}$ then the IV coefficient will be unbiased if $\chi_{iegt}$ is orthogonal to $\xi_{iegt}$.} This instrumental variable strategy can also considerably reduce the estimation sample due to the requirement that each exporter-importer-good pair shall be observed in the two consequent years.

**Market power.** Halpern and Miklos Koren (2007) develop a theory of pricing-to-firm in which exporters use third-degree price discrimination and charge different prices for their products depending on the demand elasticity of their buyers. If this story is true and quantity of goods purchased is correlated with demand elasticity of the buyer, then the estimated OLS coefficient will not truly reflect the elasticity of markups with respect to quantity, but rather the elasticity of markup with respect to the demand elasticity. Consider the model with oligopolistic competition as in Atkeson and Burstein (2008) but in which exporters can segment their markets such that a 'market' in this setting is an individual importer. If importers combine inputs in a CES fashion, their price elasticity of demand for an input will decrease in the share of expenditures on that input and so they will be charged a higher markup. In this case, if quantity and share of expenditures are positively correlated, there will be a positive relationship between prices and quantity. Thus, the OLS estimate of $\beta$ will be biased upward. This upward bias is not a general results and in some other settings if demand elasticity is falling with expenditure share, markups will be falling as well. I control for the market power story with logs of expenditure shares, defined as:

$$\ln sh_{iegt}^i = \ln \frac{p_{iegt}q_{iegt}}{\sum_{eg} p_{iegt}q_{iegt}}$$

Similarly, if importers have market power over exporters, there can be variable markdowns. To control for the markdown story, I construct logs of exporter revenue shares defined as

$$\ln sh_{iegt}^e = \ln \frac{p_{iegt}q_{iegt}}{\sum_{eg} p_{iegt}q_{iegt}},$$

and include them as control variables in some specifications.

**Quality differentiation.** To check whether quality differentiation plays a role after a very granular set of fixed effects is included as controls, I add an interaction term between
log quantity and a measure of the scope for quality differentiation discussed in the previous section. If dispersion of prices across buyers is driven by variation in quality, I expect the coefficient on the interaction between log quantity and measures of the scope for quality differentiation to be of the same sign as the coefficient on log quantity in the regression without interactions. Since in the baseline specification I find negative elasticity, I expect the coefficient on the interaction terms to be negative as well.

In addition, I use the Rauch index that measures the degree to which goods are traded on organized exchanges or have reference prices. If a good is traded on an organized exchange, or has a reference price, it reduces the power of firms to charge their buyers different prices. The value of index is 1 for decentralized industries and 0 for those with organized exchanges or reference prices. Hence, if dispersion of prices across buyers is caused by price-discrimination, I expect the interaction coefficient between the index and log quantity to have the same sign as a coefficient on log quantities.

Main findings

Firms that act as trade intermediaries, such as wholesalers, can potentially behave differently from firms that use international trade transactions to source intermediate inputs in production. Hence, the baseline sample excludes firms from retail and wholesale industries, as well as private persons. Figure 1.1 plots the non-parametric conditional expectation of log prices as a function of log-quantities for the baseline sample, controlling for exporter-HS10 good-year, importer-year, and importer-exporter-HS0 good fixed effects. The data are divided into 20 equally sized bins based on log quantities, each dot corresponds to the averages of log price and log quantity within each bin. The relationship between log prices and log quantities is downward-sloping and is well approximated by a linear functional form.

To quantify the strength of the relationship between prices and quantities, I start with estimating the \( \beta \) coefficient from the regression of the form:

\[
\ln p_{iegt} = \beta \ln q_{iegt} + \Gamma X_{iegt} + u_{iegt} \tag{1.13}
\]

where \( X_{iegt} \) is a vector of controls. Depending on the specification, the vector of controls includes exporter-HS10 good-year, importer-year, or importer-exporter-HS10 good fixed effects.

Baseline results. Table 1.3 reports estimation results of ((1.13)) for the three sets of fixed effects. In the first column I control for the exporter-HS10 good-year fixed effects. The estimated coefficient is -0.22 implying that a one percentage point increase in quantity is associated with a 0.22 percentage points reduction in unit value (doubling the quantity is associated with a 20% discount). Columns 2 and 3 add importer-year and importer-exporter-good fixed effects. The estimated coefficient remains largely unchanged.

Measurement error. Tables 1.4 and 1.5 report the IV estimates of \( \beta \) from the regression (1.13), where \( q_{iegt} \) is instrumented with a lag, \( q_{iegt,t-1} \) and a lead, \( q_{iegt,t+1} \) respectively. The set of fixed effects is the same as in Table 1.3; however, the number of observations is 2-3
times smaller. The coefficient in the first two columns is slightly smaller (around -0.17 to -0.19) in absolute terms than the OLS coefficient, which is consistent with the idea that the relationship is partially driven by the measurement error in quantities. The coefficient still remains negative and statistically significant, and the first stage is very strong (first stage F-stat is above 900 in all cases). The coefficient in the third column, which is estimated with the richest set of fixed effects, is higher than the OLS estimate in absolute terms (-0.41 when instrumented with a lag and -0.38 when instrumented with a lead). However, the first stage is much weaker (F-stat is approximately 14-19) and the precision of the estimate is much smaller (OLS estimate is still in the 95% confidence interval of the IV estimate). Hence, I conclude that the negative relationship between log prices and log quantities in the data is not driven by the measurement error.

**Market power.** Table 1.6 reports OLS estimates with logs of expenditure and revenue shares included as controls. The coefficient on log quantities becomes even more negative – in all of the three specifications it is in the ballpark of -0.3. Thus, I don’t find evidence that the negative elasticity of prices with respect to quantities is driven by the omitted expenditure and revenue shares.

**Quality differentiation.** Kugler and Verhoogen (2011) find that bigger firms buy intermediate inputs at higher unit values. According to their interpretation, bigger firms buy inputs of higher quality. I cannot test this mechanism directly using customs data as it does not have traditional information on the firm size, such as revenue or employment. I proxy the size of the firm with its total expenditures on foreign goods in a given year. Table 1.7 reports the estimates from a regression of log unit values of HS6 goods purchased by a firm on its total import. Depending on the specification, I control for the year, the industry of an importer, and the HS6 good fixed effects. In all of the specifications the estimated coefficient is positive and significant, thus I find support for the Kugler and Verhoogen (2011) hypothesis. Note that in this analysis goods are not disaggregated by exporters. If the same mechanism works when firms choose specific providers of their inputs, I expect the OLS estimate of $\beta$ in equation (1.13) to be biased up – firms that choose to buy more would also buy higher quality goods. Table 1.8 reports the estimates for the regressions in which log quantities are interacted with the Rauch measure of market organization and the two measures of the scope for quality differentiation: the RnD and advertising intensity and the Gollop-Monahan index. The difference in estimated coefficient between goods that are traded on organized exchanges or have reference prices (Rauch index = 0) and other goods (Rauch index = 1) is -0.08 to -0.1, implying that the elasticity is larger in absolute terms for the goods traded on decentralized markets. Estimated coefficients on the interactions with measures of the scope for quality differentiation are positive, and not very economically significant. These interaction coefficients reflect by how much the elasticity $\beta$ changes in response to an increase in the scope for quality differentiation by 1 standard deviation.

9The difference between the OLS and the IV coefficients in the third column is not driven by the difference in the samples. Online Appendix Tables A.1 and A.2 reestimate equation (1.13) for the samples used in Tables 1.4 and 1.5 respectively. The difference from the estimates in Table 1.3 is not larger than 0.03 in all cases.
Those coefficients are close to 0.02 for Gollop-Monahan measure and -0.006 to 0.36 for the RnD and advertising intensity. The sign of the coefficient is largely the one predicted by the literature on quality differentiation. I conclude that there is no evidence that negative elasticity of prices with respect to unit values is driven by hidden varieties of different quality within exporter-HS10 cells.

**Additional results.** I reestimate equation (1.13) for several different samples, results are reported in the Online Appendix A.2. Quantitative results hold when I: restrict the sample to the manufacturing firms only; include firms in retail and wholesale industries; drop 1% of outliers with the largest OLS residuals; restrict the smallest amount of transactions to $5,000 per year; restrict the sample to intermediate inputs only; restrict the sample to good produced in manufacturing industries only.

I conclude that there exists a negative relationship between unit values and quantities of the goods purchased by Colombian importers. The relationship is driven by the variation in markups that is consistent with quantity discounts (second-degree price discrimination).

### 1.4 Conclusions

In this Chapter I investigated second-degree price discrimination in business-to-business transactions both empirically and theoretically. Using a highly detailed transaction-level Colombian imports data, I document the negative elasticity of prices with respect to quantities, even within narrowly defined exporter-good categories. I show that such potential explanations as variation in marginal costs, quality differentiation, or measurement error, are not the main drivers of this relationship. The evidence is most consistent with non-linear pricing story, according to which firms provide quantity discounts to big buyers.
Figure 1.1: Non-parametric estimate of conditional expectation function of log prices

Note: the horizontal axis represents log quantities, the vertical axis represents log prices. The data was grouped into 20 equal-sized bins based on log quantity. Each dot represents average log price and log quantity within each bin. The red line represents linear fit. The data is demeaned by exporter-HS10 good-year and importer-exporter-HS10 good fixed effects. The data is restricted to firms in all industries but wholesale and retail trade.
**Chapter 1. Price Discrimination in International Trade: Empirical Evidence**

### 1.6 Tables

**Table 1.1: Descriptive statistics: full sample**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>s. d.</th>
<th>5th %-ile</th>
<th>95th %-ile</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel a: statistics across years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of importers</td>
<td>32,283</td>
<td>33,322</td>
<td>3,239</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of exporters</td>
<td>109,894</td>
<td>112,764</td>
<td>9,311</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of HS10 products</td>
<td>6,413</td>
<td>6,389</td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of HS6 products</td>
<td>4,613</td>
<td>4,592</td>
<td>98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of exporter-HS10 product pairs</td>
<td>541,867</td>
<td>566,328</td>
<td>65,473</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel b: statistics within and across years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of products per importer</td>
<td>12</td>
<td>3</td>
<td>33</td>
<td>1</td>
<td>52</td>
</tr>
<tr>
<td>Number of products per exporter</td>
<td>5</td>
<td>1</td>
<td>16</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Number of importers per exporter</td>
<td>1.5</td>
<td>1</td>
<td>2.6</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Importers per exporter-HS10 product</td>
<td>1.3</td>
<td>1</td>
<td>1.1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Panel a reports mean, median, and standard deviation for totals of variables listed in the first column across 2007-2015 years. Panel b) reports statistics calculated for the pooled dataset.

**Table 1.2: Descriptive statistics: baseline sample**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>s. d.</th>
<th>5th %-ile</th>
<th>95th %-ile</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel a: statistics across years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of importers</td>
<td>11,580</td>
<td>11,769</td>
<td>1,515</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of exporters</td>
<td>53,255</td>
<td>54,104</td>
<td>3,159</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of HS10 products</td>
<td>5,917</td>
<td>5,911</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of HS6 products</td>
<td>4,316</td>
<td>4,308</td>
<td>43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of exporter-HS10 product pairs</td>
<td>214,519</td>
<td>199,439</td>
<td>20,462</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel b: statistics within and across years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of products per importer</td>
<td>12</td>
<td>3</td>
<td>34</td>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>Number of products per exporter</td>
<td>4</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Number of importers per exporter</td>
<td>1.4</td>
<td>1</td>
<td>2.2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Importers per exporter-HS10 product</td>
<td>1.2</td>
<td>1</td>
<td>1.8</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Panel a reports mean, median, and standard deviation for totals of variables listed in the first column across 2007-2015 years. Panel b) reports statistics calculated for the pooled dataset. The sample is restricted to importers who reported their industry in the customs declaration, firms in retail and wholesale trade are omitted.
Table 1.3: OLS results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\ln p</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\ln p</td>
<td>-0.226***</td>
<td>-0.233***</td>
<td>-0.236***</td>
</tr>
<tr>
<td>[0.007]</td>
<td>[0.007]</td>
<td>[0.009]</td>
<td></td>
</tr>
</tbody>
</table>

Year-Exporter-HS10 FE
Yes    Yes    Yes

Year-Importer FE
No    Yes    Yes

Importer-Exporter-HS10 FE
No    No    Yes

# of clusters i
16,059 7,811 3,013

# of clusters e
20,434 16,885 5,510

Observations 327,108 298,493 149,740

Standard errors clustered at importer and exporter level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***.

Table 1.4: IV results: \( \ln q_{igt} \) instrumented with \( \ln q_{igt,t-1} \)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\ln p</td>
<td></td>
<td>\ln p</td>
<td>\ln p</td>
</tr>
<tr>
<td>\ln p</td>
<td>-0.176***</td>
<td>-0.189***</td>
<td>-0.410***</td>
</tr>
<tr>
<td>[0.010]</td>
<td>[0.011]</td>
<td>[0.094]</td>
<td></td>
</tr>
</tbody>
</table>

Year-Exporter-HS10 FE
Yes    Yes    Yes

Year-Importer FE
No    Yes    Yes

Importer-Exporter-HS10 FE
No    No    Yes

# of clusters i
6,164 3,209 1,893

# of clusters e
7,350 5,960 2,895

First stage F-stat
2,642.3 946.9 14.8

Observations
118,589 105,491 70,688

Standard errors clustered at importer and exporter level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***.
### Table 1.5: IV results: $\ln q_{\text{iegt}}$ instrumented with $\ln q_{\text{iegt},t+1}$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln p$</td>
<td>-0.177***</td>
<td>-0.186***</td>
<td>-0.388***</td>
</tr>
<tr>
<td></td>
<td>[0.010]</td>
<td>[0.012]</td>
<td>[0.089]</td>
</tr>
<tr>
<td>Year-Exporter-HS10 FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year-Importer FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Importer-Exporter-HS10 FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td># of clusters $i$</td>
<td>6,154</td>
<td>3,238</td>
<td>1,869</td>
</tr>
<tr>
<td># of clusters $e$</td>
<td>7,323</td>
<td>5,963</td>
<td>2,870</td>
</tr>
<tr>
<td>First stage F-stat</td>
<td>2,658.6</td>
<td>932.9</td>
<td>18.7</td>
</tr>
<tr>
<td>Observations</td>
<td>118,319</td>
<td>105,166</td>
<td>70,491</td>
</tr>
</tbody>
</table>

Standard errors in brackets
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Standard errors clustered at importer and exporter level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***.

### Table 1.6: OLS results, controlling for expenditure and revenue shares

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln p$</td>
<td>-0.301***</td>
<td>-0.309***</td>
<td>-0.298***</td>
</tr>
<tr>
<td></td>
<td>[0.009]</td>
<td>[0.009]</td>
<td>[0.012]</td>
</tr>
<tr>
<td>log import share per firm</td>
<td>0.200***</td>
<td>1.416***</td>
<td>1.552***</td>
</tr>
<tr>
<td></td>
<td>[0.028]</td>
<td>[0.049]</td>
<td>[0.079]</td>
</tr>
<tr>
<td>log export share per firm</td>
<td>2.134***</td>
<td>1.845***</td>
<td>1.780***</td>
</tr>
<tr>
<td></td>
<td>[0.054]</td>
<td>[0.052]</td>
<td>[0.068]</td>
</tr>
<tr>
<td>Year-Exporter-HS10 FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year-Importer FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Importer-Exporter-HS10 FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td># of clusters $i$</td>
<td>16,059</td>
<td>7,811</td>
<td>3,013</td>
</tr>
<tr>
<td># of clusters $e$</td>
<td>20,434</td>
<td>16,885</td>
<td>5,510</td>
</tr>
<tr>
<td>Observations</td>
<td>327108</td>
<td>298493</td>
<td>149740</td>
</tr>
</tbody>
</table>

Standard errors clustered at importer and exporter level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***.
Table 1.7: Unit values and size of import

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln $p$</td>
<td>0.106***</td>
<td>0.079***</td>
<td>0.097***</td>
</tr>
<tr>
<td>ln $p$</td>
<td>[0.013]</td>
<td>[0.009]</td>
<td>[0.006]</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>HS6 FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,126,847</td>
<td>1,126,843</td>
<td>1,126,705</td>
</tr>
</tbody>
</table>

Standard errors clustered at importer level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***. ln $p$ is average unit value of HS6 good purchased by a firm in a given year. ln $import$ is total import by a firm in a given year. Industry is a 4-digit ISIC industry of an importer.
Table 1.8: OLS results: interactions with measures of quality differentiation and market organization

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ln $p$</td>
<td>ln $p$</td>
<td>ln $p$</td>
</tr>
<tr>
<td>ln $q$</td>
<td>-0.181***</td>
<td>-0.220***</td>
<td>-0.224*</td>
</tr>
<tr>
<td></td>
<td>[0.046]</td>
<td>[0.057]</td>
<td>[0.092]</td>
</tr>
<tr>
<td>ln $q$ X Rauch index</td>
<td>-0.118***</td>
<td>-0.108***</td>
<td>-0.078**</td>
</tr>
<tr>
<td></td>
<td>[0.011]</td>
<td>[0.011]</td>
<td>[0.030]</td>
</tr>
<tr>
<td>ln $q$ X Gollop-Monahan index</td>
<td>0.025*</td>
<td>0.027*</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>[0.011]</td>
<td>[0.013]</td>
<td>[0.018]</td>
</tr>
<tr>
<td>ln $q$ X RnD and ad intensity</td>
<td>0.018</td>
<td>0.036***</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>[0.013]</td>
<td>[0.008]</td>
<td>[0.018]</td>
</tr>
<tr>
<td>Year-Exporter-HS10 FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year-Importer FE</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Importer-Exporter-HS10 FE</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td># of clusters $i$</td>
<td>10,514</td>
<td>5,034</td>
<td>875</td>
</tr>
<tr>
<td># of clusters $e$</td>
<td>13,570</td>
<td>1,028</td>
<td>5,510</td>
</tr>
<tr>
<td>Observations</td>
<td>106,703</td>
<td>78,184</td>
<td>16,733</td>
</tr>
</tbody>
</table>

Standard errors clustered at importer and exporter level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***.
Chapter 2

Price discrimination in International Trade: Theory

2.1 Introduction

Based on the empirical results discussed in Chapter 1, in Chapter 2 I analyze the importance of the second-degree price discrimination from the theoretical perspective in two steps. In the first step, I rationalize these findings through the lens of a theoretical framework. Specifically, I develop a tractable general equilibrium model featuring two sectors with heterogeneous firms producing nontradable final and tradable intermediate goods. Departing from previous trade models, I assume that firms in the final good sector can only purchase inputs directly from their producers, reselling of inputs is not allowed. Intermediate good producers do not observe the type (productivity) of their buyers, but they know the distribution of those types across buyers. This leads firms that produce intermediate inputs to construct an optimal price-quantity schedule for their output, rather than to quote a single price. Intermediate good producers decide to have a quantity-payment schedule that is equivalent to a two-part tariff: a sum of a fixed payment and a part that is a product of quantity, a constant marginal cost, and a constant markup. As a result, the unit price is decreasing in quantity. To understand the rationale behind these quantity discounts, consider the case in which intermediate good producers have perfect information about the type of the buyers they deal with. If buyers have CES demand, sellers are able to extract all of the buyers’ surplus with payment-quantity bundles featuring constant unit prices. If sellers do not observe buyers’ types, high-demand buyers can chose a bundle designed for low-demand buyers and get strictly positive surplus. Sellers thus have to provide quantity discounts to induce buyers to reveal their types truthfully.

The model yields a gravity equation and the same gains from trade formula as in Arkolakis, Costinot, and Rodríguez-Clare (2012), thus making the analysis of welfare and optimal policies very tractable. To understand the mechanism of how second-degree price discrimination affects the economy, I study the autarkic equilibrium first. I compare welfare in my
model with two benchmarks. The first benchmark is the social planner’s allocation (first-best allocation). The second benchmark features monopolistic competition in the market for inputs. The presence of second-degree price discrimination creates a distortion. The fixed part of the payment schedule increases the fixed costs of production for final good producers. Higher fixed costs increase the productivity cutoff of active firms. Hence, the fraction of entrants into the final good sector who decide to produce is lower than under the social planner. This distortion also encourages excess entry into the intermediate goods sector relative to the first-best allocation. Another part of the distortion is the markup present in the variable part of the payment. Since fixed production and entry costs in the final good sector are paid in labor purchased on a competitive market, this markup discourages final good producers from buying intermediate inputs and encourages them to invest more into entry compared to the first-best allocation. One of the ways to decentralize the first-best allocation in autarky is to provide a lump-sum production subsidy to firms in the final good sector and an ad-valorem subsidy for the purchases of intermediate inputs. The former will undo the effect of the fixed payment distortion, while the latter will cancel the impact of the markup distortion.

I then study optimal policies from the perspective of a small open economy, where both importers and domestic intermediate good producers use second-degree price discrimination. Since lump-sum subsidies are not a popular policy instrument, I restrict my welfare analysis to the problem of a policymaker who can only levy ad-valorem taxes on the purchases of intermediate inputs. I provide closed-form expressions for the optimal tariffs chosen by this constrained policymaker in two settings: first, in which intermediate good producers compete monopolistically, and second, in which they are able to price-discriminate. I show several important differences between these settings. The first difference is that optimal tax on domestic inputs is higher under second-degree price discrimination. Under monopolistic competition, a policymaker provides a subsidy for the purchases of domestic inputs to neutralize the effect of the markup charged by the intermediate good producers. Two forces lead to a higher tax under second-degree price discrimination. First, the markup in a variable part of the payment is lower than in the standard uniform pricing case. Second, a policymaker wants to levy an additional tax on intermediate inputs to address excess entry in the intermediate good sector. The second difference is that the import tariff is also higher under second-degree price discrimination. In this setting, a higher tariff reduces fixed payment charged by foreign firms. This leads to incomplete pass through of the tariff to import prices and thus increases the optimal level of import tax.

In the second step, I calibrate the model using the microdata and quantify the welfare effects of optimal policies. The model has four main parameters. Two of them are calibrated using the distribution of imports by Colombian firms and the distribution of exports by foreign firms. The other two parameters are taken from the literature on the trade elasticity and the elasticity of substitution in consumption. Using these parameter values, I perform three exercises. First, I quantify the welfare losses driven by second-degree price discrimination. For the baseline calibration these losses are equal to 5% in terms of real consumption. Second, I quantify the ratio of welfare under the optimal policies relative to welfare in the
benchmark case with no taxes. I find that in a small open economy, optimal policy can increase welfare by around 2% in settings with monopolistic competition and second-degree price discrimination. Third, I quantify the costs of a 'policy mistake' – a case in which firms use second-degree price discrimination but a policymaker adopts taxes that are optimal under monopolistic competition. In a small open economy this mistake can lead to welfare losses of around 0.7%.

2.2 Trade model with second-degree price discrimination

In this section, I rationalize the empirical evidence on the negative relationship between prices and quantities through a lens of a theoretical framework that embeds second-degree price discrimination into a general equilibrium trade model. I analyze the welfare effects of the distortion induced by second-degree price discrimination, characterize optimal policy menu in a small open economy, and emphasize important differences as compared to the traditional models.

Structure of the economy

Consider the world consisting of \( J \) countries indexed by \( i \) and \( j \). Each country is populated by a mass \( L_j \) of consumers. Firms produce intermediate and final goods. There is an endogenous mass \( M_j^F \) of heterogeneous final good producers. They produce differentiated varieties of a final good by combining varieties of intermediate inputs. Final good producers differ in their productivity, denoted by \( \varphi \) (\( \varphi \) indexes final good producers and varieties they produce). The cumulative distribution function of \( \varphi \) in country \( j \) is given by \( F_{\varphi,j}(\cdot) \). There is free entry in the final good sector: each entrant pays entry costs \( f_{e,F}^j \) in terms of labor in \( j \) before they learn their productivity \( \varphi \). In case entrants find it profitable to produce, they incur fixed production costs \( f_{F}^j \) in terms of labor. The market structure on the final goods market is monopolistic competition.

Turning to the intermediate good producers, each country \( j \) has an endogenous mass \( M_j^I \) of heterogeneous intermediate good producers, indexed by productivity \( z \) with the CDF \( F_{z,j}(\cdot) \) (\( z \) indexes intermediate good producers and varieties they produce). The only input in production of intermediate varieties is labor, and production functions are linear. Similar to the final good sector, there are entry costs into the intermediate goods sector equal to the \( f_{e,I}^j \) in terms of labor. Intermediate good producers from \( i \) can sell their varieties in \( j \) if they pay a fixed cost \( f_{I}^{ij} \) in terms of their country \( i \) labor. Trade is also subject to iceberg trade costs \( \tau_{ij} \geq 1 \).

So far the model is similar to Bernard, Moxnes, and Ulltveit-Moe (2017). Deviating from the previous trade literature, I postulate that intermediate inputs can only be purchased directly from their producers. This allows intermediate good producers to potentially charge
different prices for the same good. If reselling of intermediate inputs was allowed, this would lead to arbitrage opportunities. I rule out this possibility.

I assume that intermediate good producers do not observe the individual characteristics of the final good producers, but know the distribution of their buyers’ willingness to pay. These assumptions induce intermediate good producers to construct a payment-quantity schedule to maximize their profits – a practice known as second-degree price discrimination. The assumptions of incomplete information about final good producers’ types and no reselling are key for the main theoretical result in the paper. The former assumption is plausible since in the context of international trade firms can find it costly to get full information about all of their potential buyers. The latter assumption is supported by the empirical and anecdotal evidence that firms are able to charge their buyers different prices for the same good.

I solve this model in general equilibrium in several steps:

1. Step 1. Conditional on consumer preferences and aggregate variables, I solve the problem of final good producers who maximize their profits conditional on the bundle of intermediate inputs they purchase. Note that at this point final good producers do not optimize over the bundle of intermediate inputs and thus the level of production – that bundle is taken as given. The main result from this step is a profit function of final good producers conditional on their input bundle. This profit function will be an important input in Step 2.

2. Step 2. Similar to the model discussed in Varian (1989), the problem of intermediate good producers is to design a specific payment-quantity bundle for each final good producer $\varphi$ in each of the countries. They can make this bundle compatible with the incentives of final good producers, so that final good producers of type $\varphi$ self-select to choose a payment-quantity bundle designed specifically for them. To design such a bundle, each intermediate goods producer $z$ needs to know the marginal profits that their buyers make as a function of the quantity and payment for their goods, conditional on purchases of all other inputs. At the end of step 2, I solve for the quantities of intermediate inputs purchased by each of the final good producers and a corresponding payment schedule.

3. Step 3. Using the results from Step 2, I solve for all of the aggregate variables in the multi-country model in general equilibrium. I show that the model features a gravity equation and the gains from trade formula as in Arkolakis, Costinot, and Rodríguez-Clare (2012). However, second-degree price discrimination creates a distortion. I characterize welfare in this model relative to the model with monopolistic competition on the input market and the social planner allocation in autarky.

---

1See Varian (1989) and Stole (2007) for an extensive theoretical discussion of second-degree price discrimination.
CHAPTER 2. PRICE DISCRIMINATION IN INTERNATIONAL TRADE: THEORY

4. Step 4.] Finally, I introduce taxes and solve for the menu of optimal policies in autarky and in a small open economy environment. I highlight important difference from the optimal taxation policies in traditional models.

Step 1. Solving the problem of final good producers.

Consumers. A representative consumer in country $j$ has a CES utility function given by:

$$U_j = \left( M_j^F \int_{\Omega_{\varphi,j}} q_j(\varphi) \frac{\sigma - 1}{\sigma} dF_{\varphi,j}(\varphi) \right)^{\frac{\sigma}{\sigma - 1}},$$

where $\Omega_{\varphi,j}$ is the set of final good varieties produced in country $j$ and indexed by $\varphi$, $q_j(\varphi)$ is the quantity consumed, and $\sigma > 1$ is the elasticity of substitution in consumption. Consumers in $j$ inelastically supply $L_j$ units of labor and earn wage $w_j$ per unit of labor. The budget constraint of the representative agent in country $j$ is given by:

$$L_j w_j = M_j^F \int_{\Omega_{\varphi,j}} p_j(\varphi) q_j(\varphi) dF_{\varphi,j}(\varphi),$$

where $p_j(\varphi)$ is the price of final good variety $\varphi$ produced in country $j$.

Final good producers. Consider the problem of final good producer with productivity $\varphi$ from country $j$ that purchases $x_{ij}(\varphi, z)$ units of intermediate input from producer with productivity $z$ in country $i$, and pays $T_{ij}(\varphi, z)$ in exchange. The payment and quantity are both indexed by $\varphi$. Even though intermediate good producers do not observe the types of each of their customers, they can design the payment-quantity bundles in such a way that final good producers of type $\varphi$ self-select to choose a bundle designed specifically for them. In Step 2, I study the way an intermediate good producer makes the bundles incentive-compatible. Intermediate inputs from all sources are combined by a final good producer $\varphi$ in a CES fashion into a composite input $x_{ij}^F(\varphi)$:

$$x_{ij}^F(\varphi) = \left( \sum_i M_i^I \int_{\Omega_{z,ij}} x_{ij}(\varphi, z) \frac{\epsilon - 1}{\epsilon} dF_{z,i}(z) \right)^{\frac{\epsilon}{\epsilon - 1}},$$

where $\epsilon$ is the elasticity of substitution in production of final goods with the property $\epsilon > \sigma$, and $\Omega_{z,ij}$ is the set of intermediate good producers that sell from country $i$ to final good producers in country $j$. The composite input bundle can be transformed into the final good using the following technology:

$$q_j(\varphi) = \varphi x_{ij}^F(\varphi).$$

---

2This ensures existence of interior solution.

3Even though I don’t index $\Omega_{z,ij}$ by $\varphi$ I implicitly allow final good producers to source from different sets of intermediate good producers by letting $x_{ij}(\varphi, z) = 0$ for some $z$’s.
Taking into account consumer preferences and the monopolistic competition structure of the market for final goods, I can express revenues of a final good producer \( r^F_j(\varphi) \) as:

\[
r^F_j(\varphi) = q_j(\varphi)^{\frac{\sigma-1}{\sigma}} P_j Q^1_j, \quad (2.5)
\]

\[
Q^\frac{\sigma-1}{\sigma}_j = M^F_j \int_{\Omega_\varphi,j} q_j(\varphi)^{\frac{\sigma-1}{\sigma}} dF_{\varphi,j}(\varphi), \quad (2.6)
\]

\[
P^{1-\sigma}_j = M^F_j \int_{\Omega_\varphi,j} p_j(\varphi)^{1-\sigma} dF_{\varphi,j}(\varphi), \quad (2.7)
\]

where \( P_j \) is the aggregate price index and \( Q_j \) is the aggregate real consumption in the economy \( j \), such that \( P_j Q_j = L_j w_j \).

Total expenditures on the intermediate inputs by the final good producer \( \varphi \) are given by:

\[
e^F_j(\varphi) = \sum_i M^I_i \int_{\Omega_\varphi,ij} T_{ij}(\varphi, z) dF_{z,i}(z) \quad (2.8)
\]

In addition to expenditures on intermediate inputs, final good producers incur fixed costs \( w^F_j \) and thus profits of the producer \( \varphi \) net of fixed costs and conditional on the menu of bundles \( x_{ij}(\varphi, z), T_{ij}(\varphi, z) \) purchased from different intermediate good producers are given by:

\[
\pi^F_j(\varphi) = q_j(\varphi)^{\frac{\sigma-1}{\sigma}} P_j Q^1_j - \sum_i M^I_i \int_{\Omega_\varphi,ij} T_{ij}(\varphi, z) dF_{z,i}(z) - w^F_j, \quad (2.9)
\]

\[
q_j(\varphi) = \varphi x^F_j(\varphi), \quad \text{and} \quad (2.10)
\]

\[
x^F_j(\varphi) = \left( \sum_i M^I_i \int_{\Omega_\varphi,ij} x_{ij}(\varphi, z)^{\frac{1}{\tau_{ij}}} dF_{z,i}(z) \right)^{-\frac{1}{\tau_{ij}}} \quad (2.11)
\]

**Step 2. Solving the problem of intermediate good producers.**

Let \( x_{ij}(z) \) denote the quantity of intermediate variety \( z \) produced in country \( i \) and shipped to country \( j \). Trade is subject to iceberg trade costs \( \tau_{ij} \), and so shipping \( x_{ij}(z) \) units of variety \( z \) requires

\[
l(z) = x_{ij}(z) \frac{\tau_{ij}}{z} \quad (2.12)
\]

units of labor. In addition to paying variable costs of production, intermediate good producers from country \( i \) incur fixed costs of production \( f^I_{ij} \) in terms of country \( i \) labor to serve firms in market \( j \). Thus, I can express profits of the final good producer with productivity
$z$ in country $i$ net of fixed costs as:
\[
\pi^I_i(z) = \sum_{j \in J_i(z)} \pi^I_{ij}(z),
\]
(2.13)

\[
\pi^I_{ij}(z) = M^F_j \int_{\Omega_{\varphi,j}} T_{ij}(\varphi, z) dF_{\varphi,j}(\varphi) - x_{ij}(z) \frac{w_i \tau_{ij}}{z} - w_i f^I_{ij} \quad \forall j \in J_i(z),
\]
(2.14)

\[
x_{ij}(z) = M^F_j \int_{\Omega_{\varphi,j}} x_{ij}(\varphi, z) dF_{\varphi,j}(\varphi) \quad \forall j \in J_i(z),
\]
(2.15)

where $\pi^I_{ij}(z)$ denotes profits from a specific market $j$, and
\[
J_i(z) = \left\{ j : \pi^I_{ij}(z) \geq 0 \right\}
\]
(2.16)
is the set of markets that the intermediate good producer with productivity $z$ in $i$ serves. It may not be profitable to sell to all of the available destinations due to the presence of fixed costs of production. Intermediate good producers with productivity $z$ in country $i$ choose payment-quantity bundles $\{T_{ij}(\varphi, z), x_{ij}(\varphi, z)\}$ for all possible $\varphi, j$ to maximize their profits subject to the following constraints:

1. Incentive compatibility (IC) constraints. These constraints ensure that final good producers $\varphi$ prefers to choose a bundle designed specifically for them rather than a bundle designed for some other producers $\varphi'$. In other words, final good producers’ marginal profits from selecting a bundle designed specifically for them are larger than marginal profits from any other bundle.

2. Individual rationality (IR) constraints. These constraints ensure that final good producers are better off buying a bundle designed for them rather than picking no bundle at all. In other words, final good producers’ marginal profits from selecting a bundle designed specifically for them are weakly positive.

Using the final good producers’ profit function derived in Step 1, I now characterize the IC and IR constraints. I start with deriving the marginal profit function for final good producers from sourcing a particular payment-quantity bundle. Consider a final good producer with productivity $\varphi$ in country $j$ with a composite input bundle $x^I_j(\varphi)$. Imagine that this producer is matched with an additional mass $\Delta$ of country $i$ intermediate good producers from whom the final good producer sources $x$ units of intermediate inputs in exchange for payment $T$. The new profit function of this final good producer is given by:
\[
\pi^F_j(\varphi, \Delta) = q_j(\varphi, \Delta)^{\frac{q-1}{\sigma}} P_j Q_j^{1/\sigma} - \sum_i M^I_i \int_{\Omega_{\varphi,i,j}} T_{ij}(\varphi, z) dF_{z,i}(z) - \Delta T - f^F_j w_j,
\]
(2.17)
\[
q_j(\varphi, \Delta) = \varphi x^I_j(\varphi, \Delta).
\]
(2.18)
\[
x^I_j(\varphi, \Delta) = \left( \sum_i M^I_i \int_{\Omega_{\varphi,i,j}} x_{ij}(\varphi, z) \frac{w_i}{\Psi} dF_{z,i}(z) + \Delta x^{\frac{\gamma}{\xi}} \right)^{\frac{1}{\gamma-1}}.
\]
(2.19)
Taking the derivative of $\pi_j^F(\varphi, \Delta)$ with respect to $\Delta$ and letting $\Delta \to 0$ yields:

$$\frac{\partial \pi_j^F(\varphi, \Delta)}{\partial \Delta} |_{\Delta=0} = x^\gamma \mu_j(\varphi) - T$$  \hspace{1cm} (2.20)

$$\equiv b_j(x, T; \mu_j(\varphi)), \hspace{1cm} (2.21)$$

where:

$$\gamma = \frac{\epsilon - 1}{\epsilon} \hspace{1cm} (2.22)$$

$$\mu_j(\varphi) = \frac{\sigma - 1}{\sigma} \frac{\epsilon}{\epsilon - 1} x_j^j(\varphi)^{\frac{1}{\sigma} - \frac{1}{\sigma}} P_j Q_j^1/\sigma \hspace{1cm} (2.23)$$

Therefore $b_j(x, T; \mu_j(\varphi))$ is the marginal profit of the final good producer $\varphi$ from buying $x$ units of a certain variety of intermediate inputs while paying $T$, conditional on the use of other intermediate inputs. For the sake of convenience from now on I will index final good producers by $\mu_j(\varphi)$, which reflects there willingness to pay for the intermediate inputs. I assume that there is a one-to-one mapping between $\varphi$ and $\mu_j(\varphi)$ and later verify that this assumption is correct. As a result, intermediate good producers will condition their quantity-payment schedules on the type $\mu$. Let the CDF of $\mu$ in country $j$ be given by $F_{\mu,j}(\mu)$. Given the $b(x, T; \mu)$ function the IC and IR constraints can be written in the following way:

1. **Incentive compatibility:**

$$b_j(x(\mu, z), T(\mu, z); \mu) \geq b_j(x(\mu', z), T(\mu', z); \mu) \hspace{1cm} \forall \mu' \neq \mu.$$  \hspace{1cm} (2.24)

2. **Individual rationality**

$$b_j(x(\mu, z), T(\mu, z); \mu) \geq 0.$$  \hspace{1cm} (2.25)

As a result, the problem of an intermediate good producer can be formulated as:

$$\max_{x_{ij(\mu, z), T_{ij}(\mu, z)}} \pi_i^I(z) = \sum_{j \in \mathcal{I}_i(z)} \pi_{ij}^I(z) \hspace{1cm} (2.26)$$

s.t.  \hspace{1cm} $\pi_{ij}^I(z) = M_j^F \int_{\Omega_{\mu,j}} T_{ij}(\mu, z) dF_{\mu,j}(\mu) - x_{ij}(z) \frac{w_i f_{ij}}{z} - w_i f_{ij} \hspace{1cm} \forall j \in \mathcal{I}_i(z)$  \hspace{1cm} (2.27)

$$x_{ij}(z) = M_j^F \int_{\Omega_{\mu,j}} x_{ij}(\mu, z) dF_{\mu,j}(\mu) \hspace{1cm} \forall j \in \mathcal{I}_i(z) \hspace{1cm} (2.28)$$

$$b_j(x(\mu, z), T(\mu, z); \mu) \geq b_j(x(\mu', z), T(\mu', z); \mu) \hspace{1cm} \forall \mu, \mu' \hspace{1cm} (2.29)$$

$$b_j(x(\mu, z), T(\mu, z); \mu) \geq 0 \hspace{1cm} \forall \mu, \mu' \hspace{1cm} (2.30)$$

I make the following assumptions regarding productivity distributions to solve the problem in closed form:
1. The CDF of final good producers’ productivities \( \varphi \) in country \( j \) is Pareto:

\[
F_{\varphi,j}(\varphi) \equiv 1 - \left( \frac{\varphi}{b_{\varphi,j}} \right)^{-\theta_\varphi},
\]

where \( \theta_\varphi > \sigma - 1 \).

2. The CDF of intermediate good producers’ productivities in country \( i \) is Pareto:

\[
F_{z,i}(z) \equiv 1 - \left( \frac{z}{b_{z,i}} \right)^{-\theta_z},
\]

where \( \theta_z > \epsilon - 1 \).

In the Appendix B.1 I show that the solution to the problem of intermediate good producers has the following properties:

1. Every market \( j \) is characterized by a cutoff productivity for final good producers, given by \( \varphi_{0,j} \), and a cutoff productivity for intermediate good producers from country \( i \), given by \( z_{0,ij} \), such that \( \Omega_{\varphi,j} = \{ \varphi : \varphi \geq \varphi_{0,j} \} \) and \( \Omega_{z,ji} = \{ z : z \geq z_{0,ij} \} \). There is selection into production in the final good sector and selection into exporting in the intermediate good sector.

2. All intermediate good producers who serve market \( j \) sell their goods to every final good producer in that market. It turns out that every active firm in the final good sector sources its inputs from all of intermediate good producers that serve the market. Conditional on production there is no selection into importing.\(^4\)

3. In each country \( j \), the distribution of \( \mu \) is Pareto with the shape \( \theta_\mu \), and solution to the intermediate good producer problem is given by:

\[
x_{ij}(\varphi, z) = \chi \left( \frac{z}{\mu_j} \right)^{\epsilon},
\]

\[
T_{ij}(\varphi, z) = \frac{\theta_\mu}{\theta_\mu - 1} x_{ij}(\varphi, z) \frac{w_i \tau_{ij}}{z} + \frac{1}{\epsilon - 1} \frac{\theta_\mu}{\theta_\mu - 1} x_{ij}(\varphi_{0,j}, z) \frac{w_i \tau_{ij}}{z},
\]

\(^4\)Selection into importing can be generated by relationship-specific fixed costs, as in Bernard, Moxnes, and Ulltveit-Moe (2017)
where:

$$
\chi = \left( \frac{\epsilon - 1}{\epsilon} \frac{\theta_\mu - 1}{\theta_\mu} \right) ^{\epsilon},
$$

(2.35)

$$
\theta_\mu = \frac{\theta_x \epsilon}{\sigma - 1},
$$

(2.36)

$$
\mu_j(\phi) = \left[ \frac{\sigma - 1}{\sigma} \frac{\epsilon}{\epsilon - 1} \right] ^{\sigma/\epsilon} \phi^{\sigma-1} \left( \frac{\epsilon - 1}{\epsilon} \frac{\theta_\mu - 1}{\theta_\mu} \right) ^{\sigma/\epsilon-1} \left[ \frac{\theta_z}{\theta_z - (\epsilon - 1)} \right] ^{\sigma/\epsilon-1}
$$

$$
\times \left[ \sum_i M_i \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_\mu} \left( \frac{1}{\tau_{ij} \mu_i} \right)^{\epsilon-1} \frac{z_{0,ij}^{\epsilon-1}}{\tau_{ij} \mu_i} \left( P_j Q_j^{1/\gamma} \right) ^{\sigma/\epsilon} \right].
$$

(2.37)

The payment schedule given by equation (2.34) is a two-part tariff. There is a fixed part (first term), that depends on the quantity sold to the least productive final good producer, and a variable part (second term), which is a product of marginal costs of production \( w_i \tau_{ij} \), quantity purchased, and the markup \( \frac{\theta_\mu \epsilon}{\sigma - 1} \). Unit prices, defined as \( T/x \), are decreasing in quantity \( x \), and so this pricing scheme is consistent with empirical evidence documented in the previous section.\(^5\)

To understand why unit values are decreasing in quantity, consider the case in which intermediate good producers can observe the identity \( \mu \) of their buyers that have utility \( x^\gamma \mu \). In this case, a profit-maximizing seller with marginal costs \( c \) offers quantity that makes buyer’s marginal utility equal to seller’s marginal costs, and demand the payment that is equal to the buyer’s total surplus, in other words:

$$
(x^*)^\gamma \mu = c
$$

(2.38)

$$
T^* = (x^*)^\gamma \mu
$$

(2.39)

From which it follows that optimal payment is given by \( T^* = x^* \frac{\gamma}{\mu} \). Unit prices are constant and equal to \( c/\gamma = c \). If sellers do not observe the type of the buyer, they can not offer this price-quantity schedule, as a buyer with higher willingness to pay \( \mu \) can pretend to be a buyer with lower willingness to pay, \( \mu' \), and enjoy strictly positive surplus. To induce high-demand buyers reveal their type truthfully, the seller should offer them a discount. As a result, unit prices fall with quantity.

**Step 3. Aggregation**

**Gravity and gains from trade.** To close the model, I assume that trade in intermediate inputs is balanced. Let \( X_{ij} \) stand for the total payments from the final good producers in \( j \)

\(^5\)This is not the only theoretical way to get downward-sloping relationship between prices and quantities. Other approaches can build on specific contracting choices, as in Bernard and Dhingra (2015), or variation in the bargaining power across buyers and sellers, to name a few.
to intermediate good producers in $i$. Define the trade share $\lambda_{ij}$ as a share of expenditures on intermediate inputs sourced from $i$ in total expenditures on intermediate inputs in country $j$:

$$\lambda_{ij} \equiv \frac{X_{ij}}{\sum_i X_{ij}}. \tag{2.40}$$

In the Appendix B.2, I show that the model admits a gravity equation of the form:

$$\lambda_{ij} = \frac{M_f^I \left( \frac{w_i f^I_{ij}}{b_{z,i}} \right)^{1-\frac{\theta_z}{\theta}} \left( \frac{1}{b_{zi}} \right)^{-\theta_z} (w_i \tau_{ij})^{-\theta_z}}{\sum_i M_f^I \left( \frac{w_i f^I_{ij}}{b_{z,i}} \right)^{1-\frac{\theta_z}{\theta}} \left( \frac{1}{b_{zi}} \right)^{-\theta_z} (w_i \tau_{ij})^{-\theta_z}}. \tag{2.41}$$

An equation similar to (2.41) also arises in the standard Melitz (2003) model with Pareto distribution of productivities. The trade elasticity, defined as $\eta \equiv \frac{\partial}{\partial \ln \lambda_{ij}} \ln \lambda_{ij} \ln \lambda_{jj}$ is given by $-\theta_z$. This parameter is also a sufficient statistic for the estimation of gains from trade. I define gains from trade $GT_j$ as the negative change in real wage in country $j$ when this country is moving from trade equilibrium to autarky. These gains can be computed using a simple formula (detailed derivations are provided in Appendix B.3):

$$GT_j = 1 - \lambda_{jj}^{-1/\theta_z}, \tag{2.42}$$

which is common across a variety of trade models, as shown by Arkolakis, Costinot, and Rodríguez-Clare (2012).\footnote{The main reason a model has traditional gravity equation is that total revenues of the intermediate good producer $z$ from $i$ who sells in $j$ are proportional to $z^{\epsilon-1}$, as in a standard model.}

**Distortion.** As will be shown below, second-degree price discrimination generates a distortion. I quantify the welfare effects of this distortion by comparing my model with two benchmark cases. The first benchmark is the social planner’s allocation, and the second benchmark is the model with monopolistic competition in the market for intermediate inputs.

There are two types of distortions in this model. First, there is a distortion related to price discrimination by intermediate good producers. Second, there is a distortion that stems from the fact that final good producers use labor to pay fixed and entry costs. In that case they purchase labor on a perfectly competitive market. But when they buy intermediate inputs, they pay a markup over the labor cost of production of those inputs. This distortion is also present in a model with monopolistic competition on the intermediate goods market. To separate the two distortions, I consider three different problems in autarky:

1. Social planner’s problem (SP).
2. The model in which the market structure for intermediate goods is monopolistic competition (MC).
CHAPTER 2. PRICE DISCRIMINATION IN INTERNATIONAL TRADE: THEORY

3. The model in which intermediate-good producers use second-degree price discrimination (2DPD).

All of the models have the same technology and preferences and welfare can be expressed as:

\[ Q = \psi \varphi_0 x(z_0, \varphi_0) \left[ M^F \left( \frac{\varphi_0}{b_{\varphi}} \right)^{-\theta_\varphi} \right]^{\frac{\sigma}{\sigma-1}} \left[ M^I \left( \frac{z_0}{b_z} \right)^{-\theta_z} \right]^{\frac{\epsilon}{\epsilon-1}} \]  

(2.43)

where \( \psi \) is some constant. The closed form solutions for endogenous variables in (2.43) are given in the Table 2.1 (see Appendix B.4 for detailed derivations):

I first discuss differences between endogenous variables in column one (social planner) and column two (monopolistic competition). This helps me isolate the markup distortion that is present under monopolistic competition. I then discuss the distortion that only arises in a second-degree price discrimination setting.

Comparing columns 1 and 2 in Table 2.1, one can see that there is no distortion in the production cutoffs. The share of entrants into final good sector (row 1) \( \left( \frac{\varphi_0}{b_{\varphi}} \right)^{-\theta_\varphi} \) and the share of entrants into intermediate good sector \( \left( \frac{z_0}{b_z} \right)^{-\theta_z} \) (row 2) are the same in the case of social planner and in the model with monopolistic competition in the inputs market. These shares pin down production cutoffs in both sectors, \( \varphi_0 \) and \( z_0 \), which are also the same in two models. Comparing the number of entrants and \( x(\varphi_0, z_0) \) (which is the amount of intermediate input purchased by the least productive final good producer from the least productive intermediate good producer), we see important differences:

1. The number of entrants in the final good sector \( M^F \) is larger under monopolistic competition than under first-best (social planner’s allocation).

2. The number of entrants in the intermediate good sector \( M^I \) is lower under monopolistic competition than under first-best.

3. The quantity of the smallest transaction \( x(\varphi_0, z_0) \) is lower under monopolistic competition than under first-best.

The allocation under monopolistic competition deviates from the first-best. The markup distorts final good producers’ incentives in favor of entry and away from purchasing inputs into production. Thus, the monopolistic competition setting features a larger number of final good producers, which buy less inputs into production, and a lower number of intermediate good producers, relative to the first-best allocation.

This distortion is also present in a setting where intermediate good producers are allowed price-discriminate final good producers. However, second-degree price discrimination creates an additional distortion that affects the decision to produce by firms in the final good sector. In the Appendix B.4 I show that total expenditures on intermediate inputs \( e^F(\varphi) \) by a final
good producer with productivity $\varphi$ can be written as:

$$e^F(\varphi) = \frac{\sigma - 1}{\sigma} r^F(\varphi) + \frac{1}{\epsilon - 1} \frac{\sigma - 1}{\sigma} r^F(\varphi_0),$$  \hspace{1cm} (2.44)

where $r^F(\varphi)$ denotes revenues of the final good producer $\varphi$. In the monopolistic competition setting, the second term in the sum above would be absent: final good producers would spend a $\frac{\sigma - 1}{\sigma}$ share of their revenues on intermediate inputs. In a setting with second-degree price discrimination, there is an additional term, $\frac{1}{\epsilon - 1} \frac{\sigma - 1}{\sigma} r^F(\varphi_0)$, which is related to the revenues of the smallest active final good producer $\varphi_0$ and it is constant across all active final good producers. This term acts as an additional component to the fixed costs of production. As a result, it distorts the decision to produce for firms in the final good sector. The production cutoff is higher (and so the share of entrants who stay active is lower) relative to both monopolistic competition and social planner allocations.\(^7\) Note that the production cutoff in the intermediate good sector $z_0$ is not distorted. Equation (2.34) implies that payment to the intermediate good producer has a variable and a fixed component, and the variable component depends on the product of marginal costs, quantity and markup. The markup given by $\frac{\theta_\mu}{\sigma - 1}$ is lower than in the monopolistic case.\(^8\) As a result, final good producers purchase more intermediate inputs if they decide to stay active as compared to the model featuring monopolistic competition.

The distortion generated by second-degree price discrimination in some cases mitigates the distortion generated by monopolistic competition. For example, it stimulates more entry in the intermediate good sector and less entry in the final good sector; final good producers also buy more intermediate inputs. As will be shown in the next section, in some cases welfare in the model with second-degree price discrimination is higher than in the model with monopolistic competition, but this result is not general.

### Step 4. Optimal policies

In this step, I analyze optimal policies. I start with the case of autarky, considering optimal policies under monopolistic competition and second-degree price discrimination in the inputs market. Then I characterize optimal policies in these two settings from a perspective of small open economy.

**Optimal policy in autarky.** The insights learned from this exercise will be very helpful when analyzing a policymaker’s problem in a small open economy. I restrict the set of instruments available to policymakers to ad-valorem (proportional) taxes on the purchases of intermediate inputs at at rate $t$ (I will refer to $1 + t$ as the gross tax rate), and a lump-sum production tax levied on the firms in the final good sector, $T^F$. The first instrument is enough to achieve the first-best allocation under monopolistic competition. The second instrument

\(^7\)This can be seen from the first line in the table above, and the fact that $\sigma > 1$

\(^8\)In the monopolistic competition case the markup is $\frac{\theta_\mu}{\sigma - 1}$. Since $\theta_\mu = \epsilon \frac{\theta_0}{\sigma - 1} > \epsilon$, the markups in the model with second-degree price discrimination is lower, as $\theta_\mu > (\sigma - 1)$
is necessary to replicate the first-best allocation under second-degree price discrimination.\textsuperscript{9} In a small open economy setting, I will restrict the set of available policy instruments to proportional taxes and tariffs on intermediate inputs, ruling out any lump-sum taxes, since lump-sum taxes are not a very popular policy instrument in the real world. To draw parallels with a closed economy setting I also consider a problem of a policymaker who is only allowed to impose proportional taxes on the purchases of intermediate inputs in the setting with second-degree price discrimination in autarky. In all of the models, tax proceeds are fully rebated to the consumers in a lump-sum fashion.

In Appendix B.5 I show that optimal policies in autarky are the following:

1. In the setting with monopolistic competition and proportional taxes on the purchases of intermediate inputs paid by the final good producers, the optimal policy is given by:

\[
1 + t^{MC} = \frac{\epsilon - 1}{\epsilon} < 1
\]

(2.45)

The only distortion in the economy with monopolistic competition is the markup imposed by intermediate good producers. This distortion can be mitigated with a subsidy on the purchases of intermediate inputs equal to the inverse of the markup, as in equation (2.45).

2. In a setting with second-degree price discrimination with proportional taxes on intermediate inputs and lump-sum production subsidy for final good producers, optimal policy is given by:

\[
1 + t^{2DPD^*} = \frac{\theta - 1}{\theta} < 1
\]

(2.46)

\[
T^{2DPD^*} = -\frac{\sigma - 1}{\epsilon - 1} f^F < 0
\]

(2.47)

A setting with second-degree price discrimination features a markup distortion and a production cutoff distortion. Proportional subsidy (2.46) deals with the former, while lump-sum production subsidy (2.47) deals with the latter distortion.

3. In a setting with second-degree price discrimination with only proportional taxes on intermediate inputs, optimal policy is given by:

\[
1 + t^{2DPD} = \frac{\theta - 1}{\theta} \kappa > 1 + t^{MC}, \text{ where}
\]

\[
\kappa = 1 - \frac{\sigma - 1}{\epsilon - 1} \frac{\theta - (\sigma - 1)}{\theta} \phi \kappa < \frac{\theta - 1}{\theta}
\]

(2.48) (2.49)

\textsuperscript{9}This is not the only policy menu that can achieve first-best allocation. For example, it is possible to tax labor used for entry and fixed costs in the final good sector and get the same allocation as under central planner.
If lump-sum taxes/subsidies are forbidden, a policymaker has no control over production cutoffs in the intermediate and final good sectors. The only way it can affect welfare is through the effect of proportional taxes on the mass of entrants in both sectors, which is given by:

\[ M^F = \frac{L \sigma - 1}{f^{e,F} \sigma \theta \nu} \frac{1 + t}{1 + \frac{1}{\sigma} \kappa t} \]
\[ M^I = \frac{L \sigma - 1}{f^{e,I} \sigma \theta \nu} \frac{1}{1 + \frac{1}{\sigma} \kappa t}, \]

where \( \kappa \) is given by the equation (2.49). A policymaker wants to decrease the number of entrants in the intermediate good sector since it is too high relative to the first-best allocation. Higher tax decreases the number of entrants into the intermediate goods sector and increases the number of entrants into the final good sector, since tax proceeds are rebated to consumers who purchase final goods. Parameter \( \kappa \) regulates how fast the number of entrants reacts to the changes in tax. The higher \( \kappa \) is, the higher will be the response, and so the optimal tariff is decreasing with \( \kappa \). As a result, optimal tax in (2.48) is a product of two components: i) a subsidy that counteracts the markup \( \frac{\theta_{\mu-1}}{\theta_{\mu}} < 1 \); and ii) a tax \( 1/\kappa > \frac{\theta_{\mu}}{\theta_{\mu-1}} \) that reduces the number of entrants into the intermediate good sector. The resulting gross optimal tax is always greater than \( 1 + t^{MC} \). In the limit when \( \epsilon \to \infty \), intermediate goods become perfectly substitutable and two decentralized allocations coincide. Under monopolistic competition, markup \( \frac{\epsilon}{\epsilon-1} \) converges to 1, as does the optimal tax. Under second-degree price discrimination, markup \( \frac{\theta_{\mu}}{\theta_{\mu-1}} \to 1 \), fixed payment converges to 0, and \( \kappa \to 1 \), and so optimal gross tax rate also converges to 1.

**Optimal policy in a small open economy.** I now consider the problem of a policymaker who can only levy proportional tax on the purchases of intermediate inputs in a small open economy. The world consists of 2 countries. Country 1 is a small open economy, country 2 is the rest of the world. I assume that policy developments in country 1 have no effect on aggregate variables in the rest of the world. Let \( t_{i1} \) denote a net tax rate imposed by country 1 on the intermediate inputs purchased from country \( i \).

In Appendix B.5, I show that under monopolistic competition optimal tariffs are given by:

\[ 1 + t_{11}^{MC} = \frac{\epsilon - 1}{\epsilon} < 1, \]
\[ 1 + t_{21}^{MC} = \frac{\theta_z}{\theta_z \frac{\epsilon}{\epsilon-1} - 1} = \frac{\theta_z}{\theta_z \frac{\epsilon}{\epsilon-1} - 1} (1 + t_{11}^{MC}) < 1. \]

The optimal menu of taxes implies subsidies for the purchases of domestic and imported inputs. The subsidy for domestic purchases mitigates the effect of the markup distortion. Another distortion arises in a small open economy. Exporters do not internalize the effect of their entry into country 1 market on the surplus of final good producers through the
increase in variety of intermediate inputs. Thus, a policymaker wants to subsidize imports. The nature of the second distortion was analyzed by Demidova and Rodríguez-Clare (2009) who find that the same subsidy can neutralize the consumer surplus distortion in a one-sector Melitz model.

In the Appendix B.5 I show that the optimal menu of taxes under second-degree price discrimination is given by:

\[ 1 + t_{11}^{2DPD} = \frac{\theta_{\mu} - \frac{1}{\kappa}}{\theta_{\mu}} > 1 + t_{11}^{MC} \]  
\[ 1 + t_{21}^{2DPD} = \frac{\theta_{z}}{\theta_{z} - \frac{1}{\epsilon}}(1 + t_{11}^{2DPD}) > 1 + t_{21}^{MC} \]  

Both taxes are greater than the ones in the model with monopolistic competition. Optimal tax on the purchases from domestic producers that is given by (2.54) is the same as the optimal tax under autarky when lump-sum subsidies are not allowed, as in equation (2.48). Hence the same intuition applies: a policymaker wants, on the one hand, to reduce the markup distortion, and, on the other hand, to decrease the number of intermediate good producers. The formula for the tax on imports, given by (2.55), is similar to the optimal import tariff under monopolistic competition (2.53). The difference comes from the fact that optimal tax \( t_{11}^{*} \) is higher under second-degree price discrimination then under monopolistic competition and in some cases the optimal import tariff can be greater than 1. The reason for having import tariff higher than under monopolistic competition is the incomplete passthrough that arises in the model. When a policymaker increases import tariff, importers reduce the fixed part of the payment, and so average unit values (net of tariff) fall. The incomplete passthrough gives a policymaker a way to manipulate the terms-of-trade and induces her to impose a higher import tax.

2.3 Quantification

In this section, I calibrate the main parameters of the model to perform several quantitative exercises. First, I quantify the size of the distortion induced by second-degree price discrimination in autarky. Second, I compare the effects of optimal policies under monopolistic competition and second-degree price discrimination in autarky. Finally, I quantify welfare gains from optimal policies in a small open economy emphasizing the differences between the two settings. To evaluate policy experiments considered in this section, I use the exact “hat algebra” approach as in Dekle, Eaton, and S. Kortum (2008). I can calculate changes in welfare in response to any change in taxes if I know the values of the calibrated parameters and the share of expenditures on domestic intermediate inputs. The details of the procedure are laid out in the Appendix B.6.

\[^{10}\text{At this point the optimal tariffs were derived numerically. The functional forms provided below were verified using a grid of different parameter values. Analytical derivation of optimal taxes is a work in progress.}\]
CHAPTER 2. PRICE DISCRIMINATION IN INTERNATIONAL TRADE: THEORY

Calibration

To perform these exercises I need to estimate four parameters: the elasticity of substitution in consumption $\sigma$, the elasticity of substitution in the production of final goods $\epsilon$, and the two shape parameters that govern the distribution of productivities: $\varphi$ and $z$.

According to equation (2.41), international trade flows in the model are governed by a gravity equation with trade elasticity given by $z$. I calibrate $z = 4$, which is a median trade elasticity from structural estimation of the gravity equation in Head and Mayer (2014). To calibrate the shape parameters, I use the prediction of the model on the distribution of imports and exports. The model predicts that the distribution of imports across Colombian buyers is Pareto with the shape parameter $\varphi/(\sigma - 1)$ while the distribution of exports across sellers to Colombia is Pareto with the shape parameter $z/(\epsilon - 1)$. I estimate the shape parameters using Gabaix and Ibragimov (2011) method from the following regression:

$$\ln(Rank_{ft} - 1/2) = \gamma_t - \beta_x \ln x_{ft} + \epsilon_{ft}, \hspace{1cm} (2.56)$$

where $t$ stands for year, $f$ stands for the firm (exporter or importer), $x_{jt}$ is total annual volume of export or import transactions, $Rank_{jt}$ is the rank of firm in terms of its annual export or import, $\gamma_t$ is a set of year fixed effects, and $\beta_x$ is the estimated shape parameter. I restrict the sample to the top 5% of exporters (importers) in each year. Based on the sample of importers, the estimate of the shape parameter $\varphi/(\sigma - 1)$ is equal to 1.25, while from the sample of exporters I recover the estimate of $z/(\epsilon - 1)$ equal to 1.1. Finally, I calibrate $\sigma = 4$ to yield average markups in the final good sector of 33%. This estimate is also in the range of estimates of the elasticities of substitution from Broda and Weinstein (2006).

Distortions and optimal taxes in autarky

Figure 2.1 plots the ratio of welfare under first-best allocation to the welfare in the models with monopolistic competition and second-degree price discrimination in autarky for different values of $\sigma$ and $\varphi$, keeping $\varphi/(\sigma - 1) = 1.25$.\(^{11}\) The lines can be interpreted as gains from moving from a certain laissez-faire setting to the first-best allocation. The gains are strictly declining with the value of $\sigma$ under monopolistic competition, and are non-monotone in a setting with second-degree price discrimination. The main conclusion that can be drawn is that welfare can be higher in a setting with second-degree price discrimination than under monopolistic competition. For small values of $\sigma$, distortions induced by monopolistic competition lead to bigger welfare losses relative to social planner as compared to second-degree price discrimination allocation, while the reverse is true for higher values of $\sigma$. At a benchmark value of $\sigma = 4$, second-degree price discrimination induces distortions that reduce welfare by around 5% relative to monopolistic competition.

Figure 2.2 plots the optimal taxes in monopolistic competition $t^{MC}$ (solid black line) and second-degree price discrimination $t^{2DPD}$ when lump-sum subsidies are not allowed (dashed

\(^{11}\)I use equation (2.43) and expression from Table 2.1 to calculate welfare in autarky
blue curve) for different levels of $\sigma$ keeping $\theta_{\sigma}/(\sigma - 1) = 1.25$. While optimal taxes under monopolistic competition are constant, optimal taxes under second-degree price discrimination are always higher and are increasing in $\sigma$.

On Figure 2.3 I plot the welfare gains when taxes are changed from 0 to the optimal level in a specific setting in autarky. The gains are defined as the ratio of welfare under optimal policy relative to welfare under zero tax policy and thus 1 means no change in welfare. The solid black line corresponds to monopolistic competition, the dashed blue line corresponds to second-degree price discrimination. The dotted magenta line plots welfare gains when a policymaker in the environment with second-degree price discrimination mistakenly sets taxes that are optimal under monopolistic competition. This figure illustrates several important facts about the effects of optimal policies. First, these welfare gains are higher under monopolistic competition. Second, all of the lines slope down indicating that welfare gains from optimal policies are decreasing in $\sigma$ when $\theta_{\sigma}/(\sigma - 1)$ is held constant. At the benchmark level of $\sigma = 4$, there is almost no improvement in welfare when optimal policies are adopted in the case of second-degree price discrimination. Finally, the price of “policy mistakes” is increasing in $\sigma$. For the benchmark value of $\sigma$, a policymaker who imposes optimal tax $t^{MC}$ in a setting with second-degree price discrimination reduces welfare by 0.7 percentage points.

**Optimal policies in a small open economy**

I now shift to the small open economy environment with proportional taxes on the purchases of domestic and foreign intermediate inputs and no lump-sum production subsidies. Figure 2.4 plots optimal taxes for the different levels of $\sigma$. Both domestic and import tax are higher under second-degree price discrimination and, as in autarky, the difference between optimal taxes increases in $\sigma$.

Figure 2.5 plots the ratio of welfare under optimal and zero taxes for different levels of domestic trade share. The relationship is hump shaped. For very closed and open economies, optimal policy has little impact on welfare, while the effect is largest when domestic expenditure share is around 30%-40%.

The main result that I want to highlight is the cost of 'policy mistakes'. If a policymaker adopts optimal taxes suitable for monopolistic competition environment in a setting with second-degree price discrimination, the welfare is 0.65-0.7 percentage points smaller than it would have been if policymaker imposed the correct optimal taxes. I conclude that models with monopolistic competition and second-degree price discrimination imply rather different optimal taxes, and the cost of using the wrong policy menu can be quite substantial. I conclude that the type of the market structure for intermediate inputs has important has important qualitative and quantitative implications for the optimal policies in both autarky and second-degree price discrimination settings.
2.4 Conclusions

To rationalize the findings of Chapter 1, I develop a tractable theoretical framework that embeds nonlinear pricing (second-degree price discrimination) into a standard international trade model and characterize optimal policies. The model admits a gravity equation and a standard gains from trade formula. I show that second-degree price discrimination has important policy implications. First, the welfare losses from second-degree price discrimination can be quite substantial. Non-linear pricing creates a distortion that decreases the number of producers in the final goods sector, and increases the number of entrants in the intermediate goods sector. Second, optimal taxes on the purchases of intermediate inputs are higher when firms use non-linear prices as compared to standard models. A policymaker wants to correct excess entry in the intermediate goods sector imposing a higher tax. Finally, the cost of ‘policy mistakes’ can be high. If the policymaker sets tariffs that are optimal under linear pricing, but firms use second-degree price discrimination, this will lead to significant welfare losses.
2.5 Figures

Figure 2.1: Relative Welfare in Autarky

Note: the horizontal axis represents different values of the elasticity of substitution in consumption. The solid black line represents the ratio of welfare under social planner and second-degree price discrimination allocations for different level fo $\sigma$. The dashed blue line represents the ratio of welfare under social planner and monopolistic competition allocations. The red line corresponds to the value of $\sigma = 4$. Other parameters are given by: $\theta_z = 4$, $\theta_{\varphi}/(\sigma - 1) = 1.25$, $\epsilon = 4.6$.
Note: the horizontal axis represents different values of elasticity of substitution in consumption. The solid black line represents optimal tax in autarky under monopolistic competition $t^{MC}$. The dashed blue line represents optimal tax under second-degree price discrimination when lump-sum subsidies are not feasible $t^{2DPD}$. The red line corresponds to the value of $\sigma = 4$. Other parameters are given by: $\theta_z = 4$, $\theta_p/(\sigma - 1) = 1.25$, $\epsilon = 4.6$. 
CHAPTER 2. PRICE DISCRIMINATION IN INTERNATIONAL TRADE: THEORY

Figure 2.3: Gains from changing taxes from 0 to the optimal level in autarky

Note: the horizontal axis represents different values of the elasticity of substitution in consumption. The lines represent changes in welfare when policymaker changes taxes from 0 to the optimal level in a particular setting. The solid black line represents monopolistic competition. The dashed blue line represents second-degree price discrimination. The dotted magenta line represents change in welfare in a setting with second-degree price discrimination when policymaker sets taxes at the optimal level for monopolistic competition. The red line corresponds to the value of $\sigma = 4$. Other parameters are given by: $\theta_Z = 4$, $\theta_y/(\sigma - 1) = 1.25$, $\epsilon = 4.6$. mark up = 33%
Figure 2.4: Optimal taxes in a small open economy

Note: the horizontal axis represents different values of the elasticity of substitution in consumption. The solid black lines represents optimal tax in a small open economy under monopolistic competition $t^{MC}_{11}$. The dashed blue lines represents optimal tax under second-degree price discrimination when lump-sum subsidies are not feasible $t^{2DPD}_{11}$. The red line corresponds to the value of $\sigma = 4$. Other parameters are given by: $\theta_z = 4, \theta_\phi/(\sigma - 1) = 1.25, \epsilon = 4.6$.
Figure 2.5: Gains from changing taxes from 0 to the optimal level

Note: the horizontal axis represents initial domestic expenditure share. The lines represent changes in welfare when policymaker changes taxes from 0 to the optimal level in a particular setting. The solid black line represents monopolistic competition. The dashed blue line represents second-degree price discrimination. The dotted magenta line represents change in welfare in a setting with second-degree price discrimination when policymaker sets taxes at the optimal level for monopolistic competition. The parameters are given by: $\theta_z = 4$, $\sigma = 4$, $\theta_z/(\sigma - 1) = 1.25$, $\epsilon = 4.6$
## 2.6 Tables

### Table 2.1: Allocation under social planner and decentralized equilibrium

<table>
<thead>
<tr>
<th>Variable</th>
<th>SP</th>
<th>MP</th>
<th>2DPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{\phi_0}{b_x}$</td>
<td>$\frac{f^{e,F} \theta_x}{\sigma}$</td>
<td>$\frac{f^{e,F} \theta_x}{\sigma}$</td>
<td>$\frac{f^{e,F} \theta_x}{\sigma}$</td>
</tr>
<tr>
<td>$\frac{\phi_1}{b_x}$</td>
<td>$\frac{f^{e,F} \theta_x}{\epsilon}$</td>
<td>$\frac{f^{e,F} \theta_x}{\epsilon}$</td>
<td>$\frac{f^{e,F} \theta_x}{\epsilon}$</td>
</tr>
<tr>
<td>$M^F$</td>
<td>$\frac{L^{e,F} \theta_x}{\sigma}$</td>
<td>$\frac{L^{e,F} \theta_x}{\sigma}$</td>
<td>$\frac{L^{e,F} \theta_x}{\sigma}$</td>
</tr>
<tr>
<td>$M^I$</td>
<td>$\frac{L^{e,I} \theta_x}{\sigma}$</td>
<td>$\frac{L^{e,I} \theta_x}{\sigma}$</td>
<td>$\frac{L^{e,I} \theta_x}{\sigma}$</td>
</tr>
<tr>
<td>$\frac{x^{(\phi_0,\phi_1)}}{\phi_0}$</td>
<td>$\frac{L^{e,F} \theta_x}{\sigma}$</td>
<td>$\frac{L^{e,I} \theta_x}{\sigma}$</td>
<td>$\frac{L^{e,F} \theta_x}{\sigma}$</td>
</tr>
</tbody>
</table>
Chapter 3

The Intensive Margin in Trade

This chapter is the product of joint work with Ana Margarida Fernandes, Pete Klenow, Martha Denisse Pierola, and Andres Rodriguez-Clare. Thanks to them for allowing me to input our joint work as part of this dissertation.

3.1 Introduction

Across trading partners, exports can vary along the extensive margin (number of exporting firms) and the intensive margin (average exports per firm). The classic Krugman (1980) model predicts all export variation will be on the intensive margin because all firms export to every destination. Melitz (2003) brings the extensive margin to life with fixed costs of exporting, and emphasizes the importance of selection of firms into exporting.

In this paper we document the importance of the intensive and the extensive margin empirically, exploiting a much larger set of exporting countries than in the existing literature. Our key finding is that the intensive margin is at least as important as the extensive margin in driving bilateral trade flows. We show that this finding is at odds with the Melitz model if one uses a Pareto distribution of productivities, as is customary in the quantitative trade literature. In the benchmark Melitz-Pareto model, the extensive margin is dominant. A lognormal distribution of productivities, instead, allows the Melitz model to successfully match the empirical prominence of the intensive margin. We also discuss the implications of lognormal versus Pareto productivity for how trade costs affect trade flows and welfare.

To elaborate, we use the World Bank’s Exporter Dynamics Database (hereafter EDD) to systematically examine the importance of the extensive and intensive margins. The EDD covers firm-level exports from 59 (mostly developing) countries to all destination countries.

---

in most years from 2003 to 2013. For 49 of the countries, every exporting firm’s exports to each destination in a given year can be broken down into products at the HS 6-digit level.\textsuperscript{2} We add China to the EDD set of 49 countries to arrive at 50 countries for our analysis. Having many origin and destination countries in our dataset allows us to study the role of the intensive and extensive margins while allowing for origin-year and destination-year fixed effects that control for differences in population, wages, and other country characteristics that affect firm entry into exporting and exports per firm.

We find that between 40 and 60 percent of the variation in overall exports across origin-destination pairs is accounted for by the intensive margin, with the rest accounted for by the extensive margin. This breakdown into the intensive and extensive margins is robust to using different country samples or sets of fixed effects, excluding country pairs with few exporters or tiny exporters, and looking within industries. If we place exporting firms into percentiles for each trading pair and look across pairs, the importance of the intensive margin in explaining overall exports rises steadily from around 20 percent for the smallest exporters to over 50 percent for the largest exporters.

We interpret the finding that up to 60 percent of the variation in bilateral trade flows are explained by the extensive margin as providing support for the Melitz (2003) model. But the finding that at least 40 percent of that variation is explained by the intensive margin even while allowing for origin-year and destination-year fixed effects contradicts an important special case of the Melitz model, the case with Pareto-distributed firm productivity and fixed trade costs that vary only because of separate origin and destination components. Melitz-plus-Pareto has a sharp prediction: conditional on the fixed costs of exporting, all variation in exports across trading partners should occur through the number of exporters (the extensive margin). Lower variable trade costs should stimulate sales of a given exporting firm, but draw in marginal exporting firms to the point that average exports per exporter (the intensive margin) is unchanged. This exact offset is a special property of the Pareto distribution. It is not so dependent on other aspects of the Melitz model.\textsuperscript{3}

The upshot of our EDD facts could simply be that one needs to combine Melitz with a firm productivity distribution other than Pareto. But Melitz-Pareto has become a useful and tractable benchmark in international trade. It is consistent with many firm-level facts Eaton, S. Kortum, and Kramarz (2011), generates a gravity equation Chaney (2008), and yields a simple summary statistic for the welfare gains from trade Arkolakis, Costinot, and Rodríguez-Clare (2012). We therefore explore whether it can be rescued before moving beyond Pareto.

We explore several potential explanations for the positive intensive margin elasticity (the tendency of exports per firm to rise along with overall exports) in the EDD data while

\textsuperscript{2} “Exporter behavior, country size and stage of development: Evidence from the exporter dynamics database” (2016) describe the dataset in detail.

CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

46

retaining a Melitz-Pareto core. First, we consider the possibility that fixed trade costs vary by origin-destination pair. Higher fixed trade costs raise average exports per exporter, but also lower overall exports. For the intensive margin to be increasing in overall exports, one therefore needs variable trade costs to be very negatively correlated with fixed trade costs. A corollary is that, whereas variable trade costs rise decisively with distance between trade partners, fixed trade costs would need to fall with distance between trade partners. In this explanation, however, the intensive margin elasticity would be equally important for the smallest and largest exporting firms, contrary to what we see in the data, where the importance of the intensive margin rises steadily with exporting firm size.

Second, we explore the role of multi-product firms. If the typical firm exports more products to destinations with larger overall flows, this could account for the importance of the intensive margin for exports. We find that the number of HS 6-digit products per exporting firm does indeed account for about 12 percent of the variation in overall exports, or about one-fourth of the contribution of the intensive margin to overall exports. In the context of the multi-product Melitz-Pareto model developed by Bernard, Redding and Schott (2011), however, this explanation still requires a negative correlation between firm-level fixed costs of exporting and variable trade costs, and for firm-level fixed costs to fall with the distance between trading partners. Moreover, the significant intensive margin elasticity per firm-product implies that fixed costs of exporting per product also fall with distance.

A third hypothesis we investigate is granularity — a finite number of firms. With a finite number of firms, the intensive margin (and overall exports) can be high because of favorable productivity draws from the Pareto distribution within a country. We develop an estimator for the elasticity of fixed trade costs to distance that is valid under granularity as in Eaton, S. S. Kortum, and Sotelo (2012), and continue to find that fixed trade costs must fall with distance to explain a positive intensive margin elasticity. Using simulations of finite draws from a Pareto distribution, we find that granularity generates only a modest intensive margin elasticity, and — in contrast to what we observe in the data — almost entirely in the right tail of the exporter size distribution.

After these failed attempts to rescue the Melitz-Pareto model, we depart from the comforts of that model and consider a lognormal distribution of firm productivity. Head, Mayer, and Thoenig (2014) analyze how the welfare gains from trade in the Melitz model differ with a lognormal instead of a Pareto distribution. Bas, Mayer, and Thoenig (2015) show how the trade elasticity varies with a lognormal distribution. Both papers marshal evidence from firms in France and China pointing to the empirical relevance of the lognormal distribution.4

As in Eaton, S. Kortum, and Kramarz (2011), we consider a Melitz model with firm-destination-specific demand and fixed cost shocks, but we assume that the firm productivity

4This may seem surprising in light of the evidence in Axtell (2001) for a Pareto distribution of U.S. firm sizes. Moreover, Gabaix (2009) emphasizes that a Pareto distribution emerges naturally from random growth and some extensions. But Rossi-Hansberg and Wright (2007) find thinner-than-Pareto tails of the firm and (especially) plant size distributions in the U.S. And Luttmer (2011) shows that the largest U.S. firms are far too young to emerge from random growth. A lognormal distribution, meanwhile, can arise from random growth innovations, albeit with exploding variance without mean reversion in levels.
distribution is lognormal rather than Pareto. In particular, we assume that each firm is characterized by a productivity parameter as well as an idiosyncratic demand shifter and fixed cost for each destination market, all drawn from a multivariate lognormal distribution. We allow for a non-zero covariance between the demand shifter and the fixed cost in each destination, but set all other covariances to zero. One appealing feature of the model is that it is amenable to maximum-likelihood methods. As the likelihood is potentially not concave as a function of the parameters, and since we have a large number of parameters to estimate (means, variances, covariance, and trade costs), we rely on the estimation methodology proposed by Chernozhukov and Hong (2003).

Our estimation shows that a lognormal distribution for productivity can indeed generate a sizable intensive margin elasticity. When variable trade costs fall and fixed costs are constant, the ratio of mean to minimum exports per firm increases as the productivity cutoff falls under the lognormal distribution (while being constant under Pareto).\(^5\) Shifting to lognormal productivity also changes our inference about fixed trade costs, rendering them positively correlated with variable trade costs and rising with distance. As in the data, the intensive margin elasticity rises steadily with the size percentile of exporters under a lognormal productivity distribution. We further show that lognormally distributed firm-destination-specific demand shocks and fixed cost shocks can contribute to fitting the intensive margin facts.

3.2 The Intensive Margin in the Data

The Exporter Dynamics Database

We use the Exporter Dynamics Database (EDD) described in “Exporter behavior, country size and stage of development: Evidence from the exporter dynamics database” (2016) to study the intensive and extensive margins of trade. The EDD is based on firm-level customs data covering the universe of export transactions provided by customs agencies from 59 countries (53 developing and 6 developed countries). For each country, the raw firm-level customs data contains annual export flows (in values) disaggregated by firm, destination and Harmonized System (HS) 6-digit product. Oil exports are excluded from the customs data due to lack of accurate firm-level data for many of the oil-exporting countries. For most countries total non-oil exports in the EDD are close to total non-oil exports reported in COMTRADE/WITS. More than 100 statistics from the EDD are publicly available at the origin-year, origin-product-year, origin-destination-year, or origin-product-destination-year levels. These input average exports per firm as well as the number of exporting firms.

For the descriptive analysis in this section as well as for the regression and simulation work in the sections that follow we focus on a core sample that consists of 50 countries (49

\(^5\)The result holds under other thin-tailed productivity distributions, such as bounded Pareto as in Feenstra (2014). But a bounded Pareto distribution loses all the analytical convenience of the unbounded Pareto while lacking the empirical convenience of the lognormal distribution.
from the EDD and China) for which we have the firm-level data.\textsuperscript{6} However, to use the most comprehensive sample of countries available we rely for the motivating plots below on an extended sample that inputs the 59 origin countries from the EDD plus China. Both samples cover a subset of years from 2003 and 2013 — see Table 3.1 and Table C.1 in the Appendix.

We focus on EDD statistics based on products belonging only to the manufacturing sector. Specifically, relying on a concordance across the ISIC rev. 3 classification and the HS 6-digit classification, we consider only exports of HS 6-digit products that correspond to ISIC manufacturing sub-sectors 15-37. Using these data we calculate variants of average exports per firm, number of exporting firms, and total exports at the origin-destination-year level or at the origin-product-destination-year level. The product disaggregations that we use are HS 2-digit for the extended sample and HS 2-digit, HS 4-digit, or HS 6-digit for the core sample.

**Importance of the intensive margin**

Let $X_{ij}$, $N_{ij}$ and $x_{ij} ≡ X_{ij}/N_{ij}$ denote total exports, total number of exporting firms, and average exports per firm from country $i$ to country $j$, respectively. In Figure 3.1 we plot the intensive margin ($\ln x_{ij}$) and extensive margin ($\ln N_{ij}$) vs. total exports ($\ln X_{ij}$) for the extended sample of countries. We restrict the sample to the origin-destination pairs with more than 100 exporting firms (i.e., $ij$ pairs for which $N_{ij} > 100$) to reduce noise associated with country pairs with few exporting firms.\textsuperscript{7} All variables plotted are demeaned of origin-year and destination-year fixed effects. Each dot corresponds to ($\ln x_{ij}$, $\ln X_{ij}$) (Panel A) or ($\ln N_{ij}$, $\ln X_{ij}$) (Panel B). The lines can be ignored for now.

A key statistic that we use to summarize the pattern observed in Figure 3.1 is the **intensive margin elasticity** (IME), which is the slope of the regression line in Panel A. In a given year, the IME can be obtained from an OLS regression of $\ln x_{ij}$ on $\ln X_{ij}$ with origin and destination fixed effects:

$$
\ln x_{ij} = FE_i^o + FE_j^d + \alpha \ln X_{ij} + \varepsilon_{ij}.
$$

The IME is the estimated regression coefficient

$$
\hat{\alpha} = \frac{cov(\ln \tilde{x}_{ij}, \ln \tilde{X}_{ij})}{var(\ln \tilde{X}_{ij})},
$$

where we write $\ln \tilde{z}_{ij}$ to denote variable $\ln z_{ij}$ demeaned by origin-year and destination-year fixed effects. The complement of the IME is the **extensive margin elasticity**, defined as $EME = 1 - IME$.

\textsuperscript{6} China is not inputed in the EDD due to confidentiality concerns.

\textsuperscript{7} The core sample inputs 1291 unique country pairs with $N_{ij} > 100$ while the extended sample inputs 2075 unique country pairs with $N_{ij} > 100$. 

---

\textsuperscript{6} China is not inputed in the EDD due to confidentiality concerns.

\textsuperscript{7} The core sample inputs 1291 unique country pairs with $N_{ij} > 100$ while the extended sample inputs 2075 unique country pairs with $N_{ij} > 100$. 

---
Figure 3.1 demonstrates that both the IME and the EME are positive and large. As shown in Panel A of Table 3.2, depending on the type of fixed effects inputed, the IME ranges from 0.4 to 0.46 in the core sample that we will use for the analysis in the next two sections. Our preferred estimate of the IME is 0.4 based on the inclusion of origin-year and destination-year fixed effects (as in Figure 3.1). In this estimate, the intensive margin accounts for approximately 40% of the variation in total exports across country pairs, while 60% is accounted for by the extensive margin. As the focus has so far been on accounting for the variation in bilateral trade flows while controlling for origin-year and destination-year fixed effects, it is natural to wonder how much of that variation is absorbed by the fixed effects alone. The results in Table 3.2 show that this is never more than 59 percent, implying that a large share of the variation in bilateral trade flows comes from the forces behind the estimated IME.

Robustness

The finding of a positive and large IME is robust to considering different samples. In Panel B of Table 3.2 we estimate the IME including all country pairs — even those with less than 100 exporting firms. The IME in this case reaches 0.58 when origin-year and destination-year fixed effects are inputted. In the Appendix Table C.2 we reproduce the regressions in Table 3.2 but now for the extended sample of countries. In the preferred specification with origin-year and destination-year fixed effects, the IME is 0.38 among origin-destination pairs with at least 100 exporting firms and 0.52 among all origin-destination pairs. To make sure the IME is not driven by small exporting firms, we re-estimate it after excluding firms whose annual exports fell below $1,000 in any year. The corresponding IME estimates in Table 3.3 (core sample) and Appendix Table C.4 (extended sample) change only slightly.

A separate concern is measurement error. Since total exports is the sum of firm-level exports, classical measurement error in exports per exporter \( x \) would bias the IME upward, but classical measurement error in the number of exporters \( N \) would bias the IME downward. Depending on their relative importance compared to the true IME, classical measurement error could bias the IME upward or downward. If the measurement error is serially uncorrelated, then instrumenting total exports with its leads and/or lags should yield an unbiased estimate of the IME. As shown in Appendix Table C.5, the instrumented IMEs are very close to the OLS IME, both economically and statistically.

Our results for the IME could be coming from country differences in industry composition of exports combined with industry differences in average exports per firm. In Figure 3.2

---

8To be specific, the equation estimated in this case is \( \ln x_{ijt} = FE_{it} + FE_{jt} + \alpha \ln X_{ijt} + \varepsilon_{ijt} \) using all years of available data for the country pairs inputed in the core sample.

9This percentage comes from the R-squared of an OLS regression of bilateral total exports in logs (\( \ln X_{ijt} \)) on origin-year and destination-year fixed effects.

10When we allow the IME to differ across origin countries depending on their GDP per capita, we find that IME estimates are close to or larger than 0.4 for any group of countries, as shown in Appendix Table C.3.
we plot the (demeaned) intensive and extensive margins against total exports at the origin-industry-destination-year level using HS 2-digit industries. The pattern here is similar to that in Figure 3.1. Table 3.4 shows that the IME actually increases when moving to industry-level data. At the lowest level of aggregation available (HS 6-digit), for the core sample of countries the IME is 0.51 with origin-year-industry and destination-year-industry fixed effects. The results also hold in the extended sample, for which we calculated IME disaggregated at HS2 product level. As reported in the Appendix Table C.6, this IME is also close to 0.52.\footnote{The presence of large trading firms could lead to both high exports per firm and high total exports and explain our IME estimates. While we are unable to identify large trading firms in the EDD data, we estimate the IME based on a sample including only HS 2-digit industries with low shares of firms exporting via intermediaries, as defined in Chan (2017). The results in Appendix Table C.7 show an almost unchanged IME at 0.53.}

**IME by percentiles**

A positive IME could be due to the presence of export superstars that increase both average exports per firm and total exports for some country pairs, as discussed in Freund and Pierola 2015. We study this possibility by considering separate IME regressions for each exporter size percentile. For each origin-destination-year combination we distribute the exporting firms into percentiles based on their exports. Denoting average exports per firm in percentile \( p \) as \( x_{ij}^{pct} \), we run the regressions:

\[
\ln x_{ij}^{pct} = FE_i^o + FE_d^d + \alpha^{pct} \ln X_{ij} + \epsilon_{ij}.
\]

We define the IME for each percentile as \( IME^{pct} \equiv \hat{\alpha}^{pct} \).

We plot the IME\(^{pct} \) for each percentile (with confidence intervals) in Figure 3.3 along with the horizontal line at the overall IME of 0.4.\footnote{For exporter percentiles to be well-defined we focus on country pairs for which \( N_{ij} > 100 \).} The IME is 0.5 for the highest percentile. But the positive overall IME is not coming exclusively from the export superstars: the IME\(^{pct} \) rises steadily from 0.2 at the 50th percentile to 0.3 at the 80th percentile.

**IME for multi-product firms**

We can dig deeper and study whether average exports per firm can be explained by the number of products exported per firm or by exports per product per firm. Let \( m_{ij} \) be the average number of products exported from \( i \) to \( j \) by firms exporting from \( i \) to \( j \), and let \( x_{ij}^p \equiv x_{ij}/m_{ij} \) be the average exports per product per firm exporting from \( i \) to \( j \). We define the IME at the product level as \( IME^p \equiv cov(\ln \tilde{x}_{ij}^p, \ln \tilde{X}_{ij})/var(\ln \tilde{X}_{ij}) \). Since \( x_{ij} = x_{ij}^p m_{ij} \), the IME is equal to the IME\(^p \) plus the extensive product margin elasticity,

\[
IME = IME^p + \frac{cov(\ln \tilde{m}_{ij}, \ln \tilde{X}_{ij})}{var(\ln \tilde{X}_{ij})}.
\]
Table 3.5 reports the results for the IMEp for the core sample. Most of the IME is explained by the systematic variation in average exports per product per firm, rather than in the average number of products exported by firm.

**Taking stock: the IME in the EDD**

Summarizing the results so far, we find the intensive margin elasticity to be positive and significant, both statistically and economically. This finding is robust to the inclusion of a variety of fixed effects, various samples, exclusion of small firms, and disaggregation by industry. The IME is positive and monotonically increasing across the whole distribution of exporter size. The systematic cross-country-pair variation of average exports per firm comes primarily from the behavior of average exports per product per firm.

**Correlation between intensive and extensive margin, and relation with distance**

We now move beyond the intensive margin elasticities and report additional stylized facts on the correlations between the intensive margin, the extensive margin, and distance. There is a positive and significant correlation between average exports per firm and the number of exporting firms (0.25, standard error 0.01) after taking out origin-year and destination-year effects. Table 3.6 shows how these margins vary with log distance with alternative sets of fixed effects. The elasticities are all negative and significant when controlling for origin-year and destination-year fixed effects: average exports per firm, the number of firms, average number of products exported per firm, and average exports per product per firm all decline with distance between trade partners.

**Relation to previous empirical results**

We finish this section by relating our stylized facts to those of EKK, EKS, Bernard et al. 2007 and Bernard et al. 2009. EKK use firm-level export data for a single origin (France) and show that average exports per firm increase with market size of the destination (measured as manufacturing absorption) with an elasticity of 1/3. In Figure 3.4 we plot market size (horizontal axis) against our estimated destination fixed effects (vertical axis) from a regression of average exports per firm on origin-year and destination-year fixed effects based on the extended sample and country pairs for which \( N_{ij} > 100 \). A regression line through the points in the plot implies that average exports per firm increase with destination market size with an elasticity of 0.19, a bit lower than the result in EKK.\(^{14}\)

EKK also show that firms exporting to more destinations exhibit higher sales in the domestic (French) market. Our data does not input domestic sales, but we can instead look

\(^{13}\)Bernard et al. (2009) present a similar decomposition for U.S. exports. We compare their results to ours below.

\(^{14}\)Similar findings are obtained in unreported plots where the destination fixed effects are based on the extended sample and all country pairs or based on the core sample and either country pairs for which \( N_{ij} > 100 \) or all country pairs.
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

at sales in the most popular destination market for each origin. Let $x_{ilj}$ denote average exports to destination $l$ computed across firms from $i$ that sell in markets $l$ and $j$ and let $l^*(i) \equiv \arg \max_k N_{lk}$ be the largest destination market for each origin country $i$ (e.g., the United States (U.S.) for Mexico). In Figure 3.5 we plot $\log \frac{x_{il^*(i)j}}{N_{il^*(i)}}$ (vertical axis) against $\log \frac{N_{ij}}{N_{ij^*(i)}}$ (horizontal axis) for all $i$ and $j$ for the core sample.\footnote{The EKK estimating sample inputs only firms with sales in France. To implement an approach comparable to theirs, we drop all firms from country $i$ that do not sell to $l^*(i)$, so the sample inputs only $N_{il^*(i)}$ firms for country $i$. This implies that all firms that make up $N_{ij}$ are also selling to $l^*(i)$.} It is very clear that the results derived by EKK for French firms remains valid for our data with many origin countries: firms that sell in more markets are more productive as proxied by their sales in their origin country’s most popular destination market.

EKS find that average exports per firm are very similar across four origin countries (Brazil, Denmark, France and Uruguay) for which they have customs data. They regress average exports per firm on origin and destination fixed effects and find that the origin fixed effects differ little across their four origins. Running the same regression in our dataset (but pooling across years and including year fixed effects), we find that origin fixed effects do vary significantly across countries (the coefficient of variation in the estimated origin fixed effects ranges from 0.81 to 2.56, depending on the sample used) and are higher for countries with higher GDP per capita and higher total exports.\footnote{For this purpose, we run regressions of the estimated origin fixed effects on population, GDP, GDP per capita, and total exports, jointly and separately.} Moreover, origin-year and destination-year fixed effects are not enough to capture the variation in $\ln x_{ij}$: a regression of $\ln x_{ij}$ on origin-year and destination-year fixed effects yields an R-squared of 0.65 when only country pairs with $N_{ij} > 100$ are considered and only 0.37 when all country pairs are considered.

Using firm-level export data for the U.S., Bernard et al. 2009 present a similar decomposition to the one we present above for multi-product firms, except that they cannot allow for destination fixed effects because their data is for a single origin. They find that IME$^p$ is around 0.23, which is not far from our finding of around 0.29. On the other hand, contrary to our results, Bernard et al. 2007 find that average exports per product per firm increase with distance. We believe that the difference arises from the fact that, by having data for multiple origins, we are able to control for destination fixed effects. In fact, Table 3.6 shows that regressing $\ln x_{ij}^p$ on $\ln dist_{ij}$ with only origin and year fixed effects but without destination fixed effects yields a positive and significant coefficient as in Bernard et al. 2007, whereas the coefficient becomes negative and significant when destination fixed effects are added. The same happens when regressing $\ln x_{ij}$ on $\ln dist_{ij}$.

### 3.3 The Intensive Margin in the Melitz-Pareto Model

In this section we ask how the Melitz model with Pareto distributed productivity stacks up relative to the findings of the previous section. We focus on the implications of this model for the intensive margin elasticity. We start with the simplest model, which entails a
continuum of single-product firms with a Pareto distribution for productivity as in Chaney (2008) and Arkolakis, Demidova, et al. (2008). We derive a series of properties of this model, and then explore their robustness to allowing for destination-specific demand and fixed trade cost shocks at the firm level as in EKK, for multi-product firms, and for granularity.

The Basic Melitz-Pareto Model

Theory

As this is a well-known model, we will be brief in the presentation of the main assumptions. There are many countries indexed by \(i,j\). Labor is the only factor of production available in fixed supply \(L_i\) in country \(i\) and the wage is \(w_i\). Preferences are constant elasticity of substitution (CES) with elasticity of substitution across varieties \(\sigma > 1\). Each firm produces one variety under monopolistic competition. In each country \(i\) there is a large pool of firms of measure \(N_i\) with productivity \(\varphi\) distributed Pareto with shape parameter \(\theta > \sigma - 1\) and scale parameter \(b_i\), \(\Pr (\varphi \leq \varphi_0) = G_i(\varphi_0) = 1 - (\varphi_0/b_i)^{-\theta}\). Firms from country \(i\) also incur fixed trade costs \(F_{ij}\) as well as iceberg trade costs \(\tau_{ij}\) to sell in country \(j\).

Sales in destination \(j\) by a firm from origin \(i\) with productivity \(\varphi\) are

\[
x_{ij}(\varphi) = A_j \left( \frac{w_i \tau_{ij}}{\varphi} \right)^{1-\sigma},
\]

where \(A_j \equiv P_j^{1-\sigma} w_j L_j, P_j^{1-\sigma} = \sum_i N_i \int_{\varphi \geq \varphi_{ij}^*} \left( \frac{w_i \tau_{ij}}{\varphi} \right)^{1-\sigma} dG_i(\varphi)\) is the price index in \(j\), \(\bar{\sigma} \equiv \sigma/(\sigma - 1)\) is the markup, and \(\varphi_{ij}^*\) is the productivity cutoff for exports from \(i\) to \(j\), which is determined implicitly by

\[
x_{ij}(\varphi_{ij}^*) = \sigma F_{ij}.
\]

The value of overall exports and the number of firms that export from \(i\) to \(j\) are then \(X_{ij} = N_i \int_{\varphi \geq \varphi_{ij}^*} x_{ij}(\varphi)dG_i(\varphi)\) and \(N_{ij} = N_i \int_{\varphi \geq \varphi_{ij}^*} dG_i(\varphi)\), respectively. Using again the fact that \(G_i(\varphi)\) is Pareto and assuming that \(\varphi_{ij}^* > b_i\) for all \(i,j\), we get that

\[
X_{ij} = \left( \frac{\theta}{\theta - (\sigma - 1)} \right) A_j \left( \frac{w_i \tau_{ij}}{\varphi} \right)^{1-\sigma} b_i^\theta N_i \left( \varphi_{ij}^* \right)^{\sigma-\theta-1}
\]

and

\[
N_{ij} = b_i^\theta N_i \left( \varphi_{ij}^* \right)^{-\theta}.
\]

Combining (3.38), (3.5) and (3.6), the extensive margin is

\[
N_{ij} = N_i \left( \frac{w_i}{b_i} \right)^{-\theta} \left( \frac{\sigma}{A_j} \right)^{-\theta/(\sigma - 1)} \tau_{ij}^{-\theta} F_{ij}^{-\theta/(\sigma - 1)},
\]

\(F_{ij}\) is in units of the numeraire. Since we focus on cross-section properties of the equilibrium, we do not need to specify whether the fixed trade cost entails hiring labor in the origin or the destination country.
while the intensive margin is

\[ x_{ij} \equiv \frac{X_{ij}}{N_{ij}} = \left( \frac{\theta \sigma}{\theta - (\sigma - 1)} \right) F_{ij}. \] (3.8)

We can always decompose variable and fixed trade costs as follows: \( \tau_{ij} = \tau_i^o \tau_j^d \tilde{\tau}_{ij} \) and \( F_{ij} = F_i^o F_j^d \tilde{F}_{ij} \). Taking logs in (3.7) and (3.8), and defining variables appropriately, we have

\[ \ln N_{ij} = \mu_i^{N,o} + \mu_j^{N,d} - \theta \ln \tilde{\tau}_{ij} - \bar{\theta} \ln \tilde{F}_{ij} \] (3.9)

and

\[ \ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln \tilde{F}_{ij}, \] (3.10)

where \( \bar{\theta} \equiv \frac{\theta}{\sigma - 1} \). These are the two key equations that we use to derive the results in the rest of this section.

Combining the definition of the intensive margin elasticity given in the previous section (i.e., \( \text{IME} = \frac{\text{cov}(\ln \tilde{x}_{ij}, \ln \tilde{X}_{ij})}{\text{var}(\ln \tilde{X}_{ij})} \)) with equations (3.9) and (3.10), the model implies that

\[ \text{IME} = -\left( \bar{\theta} - 1 \right) \frac{\text{var}(\ln \tilde{F}_{ij}) - \theta \text{cov}(\ln \tilde{\tau}_{ij}, \ln \tilde{F}_{ij})}{\text{var}(\bar{\theta} \ln \tilde{\tau}_{ij} - (\bar{\theta} - 1) \ln \tilde{F}_{ij})}. \] (3.11)

This result can be used to extract several implications of the model, which we present in the form of four observations in the rest of this section.

Our first observation says that if all variation in fixed trade costs comes from origin and destination fixed effects with no country-pair component, for example because \( F_{ij} \propto w_i^\gamma w_j^{1-\gamma} \) (as in Arkolakis (2010)), then the model implies that the intensive margin elasticity is zero.

**Observation 1**: If \( \text{var}(\ln \tilde{F}_{ij}) = 0 \) then IME = 0.

Since this is a key result, it is worth understanding it in more detail. Using equations (3.3) and (3.38) together with the definition of \( x_{ij} \), taking logs and differentiating w.r.t. \( \ln \tau_{ij} \) we get

\[ \frac{d \ln x_{ij}}{d \ln \tau_{ij}} = 1 - \sigma - \frac{d \ln (1 - G_i(\varphi_{ij}^*))}{d \ln \varphi_{ij}^*} \left( \frac{1 - \frac{x_{ij}(\varphi_{ij}^*)}{x_{ij}}}{\frac{x_{ij}}{x_{ij}(\varphi_{ij}^*)}} \right). \]

The first term is the direct effect on incumbent firms, while the second term captures selection. In turn, selection is the product of \(-\frac{d \ln (1 - G_i(\varphi_{ij}^*))}{d \ln \varphi_{ij}^*}\), which captures the effect of \( \tau_{ij} \) (and hence \( \varphi_{ij}^* \)) on average exports per firm through its impact on the share of firms that export, and \( \left( \frac{1 - \frac{x_{ij}(\varphi_{ij}^*)}{x_{ij}}}{\frac{x_{ij}}{x_{ij}(\varphi_{ij}^*)}} \right) \), which captures how much less firms export at the cutoff relative to the average. Obviously, if \( x_{ij} = x_{ij}(\varphi_{ij}^*) \) then there is no selection, while the effect of selection is maximized if \( x_{ij}(\varphi_{ij}^*)/x_{ij} = 0 \). With a Pareto distribution for productivity we have \(-\frac{d \ln (1 - G_i(\varphi_{ij}^*))}{d \ln \varphi_{ij}^*} = \theta \) and \( \frac{x_{ij}}{x_{ij}(\varphi_{ij}^*)} = \frac{\theta}{\theta - (\sigma - 1)} \), therefore \( \frac{d \ln x_{ij}}{d \ln \tau_{ij}} = 0 \).
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

Combined with the assumption that $\bar{\theta} > 1$, the result in equation (3.11) also implies that if the intensive margin elasticity is positive then there must be a negative correlation between the variable and fixed trade costs (ignoring origin and destination fixed costs).

**Observation 2:** If $\text{IME} > 0$ then $\text{corr}(\ln \tilde{F}_{ij}, \ln \tilde{\tau}_{ij}) < 0$.

Ignoring origin and destination fixed effects, equation (3.10) implies that
\[
\text{cov}(\ln \tilde{F}_{ij}, \ln \tilde{\text{dist}}_{ij}) = \text{cov}(\ln x_{ij}, \ln \tilde{\text{dist}}_{ij}).
\]

Thus, if average exports per firm fall with distance then fixed trade costs must also fall with distance. This is captured formally by our third observation which is related to the fixed trade costs elasticity with respect to distance.

**Observation 3:**
\[
\frac{\text{cov}(\ln x_{ij}, \ln \tilde{\text{dist}}_{ij})}{\text{var}(\ln \tilde{\text{dist}}_{ij})} = \frac{\text{cov}(\ln \tilde{F}_{ij}, \ln \tilde{\text{dist}}_{ij})}{\text{var}(\ln \tilde{\text{dist}}_{ij})}.
\]

We can go beyond the previous qualitative observations and derive the fixed and variable trade costs implied by the model so as to compute actual values for $\text{corr}(\ln \tilde{F}_{ij}, \ln \tilde{\tau}_{ij}) < 0$ and $\text{cov}(\ln \tilde{F}_{ij}, \ln \tilde{\text{dist}}_{ij})/\text{var}(\ln \tilde{\text{dist}}_{ij})$. Combining equations (3.9) and (3.10) to solve for $\ln \tilde{F}_{ij}$ and $\ln \tilde{\tau}_{ij}$ in terms of $\ln x_{ij}$ and $\ln N_{ij}$ yields
\[
\ln \tilde{F}_{ij} = \delta_{i}^{F,o} + \delta_{j}^{F,d} + \ln x_{ij}
\]
and
\[
\theta \ln \tilde{\tau}_{ij} = \delta_{i}^{\tau,o} + \delta_{j}^{\tau,d} - \bar{\theta} \ln x_{ij} - \ln N_{ij}.
\]

Model-implied values for $\ln \tilde{F}_{ij}$ are (ignoring origin and destination fixed effects) directly given by $\ln x_{ij}$, but for $\ln \tilde{\tau}_{ij}$ a value for $\bar{\theta}$ is required to go from $\ln x_{ij}$ and $\ln N_{ij}$ in the data to model-implied values for $\theta \ln \tilde{\tau}_{ij}$.

Exports of a firm in the $p^{th}$ percentile of the exporter size distribution are
\[
\sigma F_{ij} \left( \frac{\varphi^p}{\varphi^*_{ij}} \right)^{\sigma - 1},
\]
where $\varphi^p$ is such that $\text{Pr}[\varphi < \varphi^p | \varphi > \varphi^*_{ij}] = p$. Since productivity is distributed Pareto, the ratio $\varphi^p/\varphi^*_{ij}$ and thus average exports per firm in each percentile should be the same for all $ij$ pairs. This implies that the intensive margin elasticity calculated separately for each exporter size percentile is the same as the overall intensive margin elasticity.

**Observation 4:** IME$^{pct} = \text{IME}$, for all $pct$.

Data

We now use Observations 1 – 4 above to relate the simple Melitz-Pareto model to the data as described in Section 2.

Observation 1 indicates that if fixed trade costs vary by origin and destination but not across country pairs, i.e., $\text{var}(\tilde{F}_{ij}) = 0$, then the IME should be equal to zero while the EME should be equal to one. This is captured in Figures 3.1 and 3.2 by the horizontal line for the model-implied intensive margin (panel a) and the line with unit slope for the model-implied extensive margin (panel b). These implications of the model stand in sharp contrast to what
is seen in the data, both in Figures 3.1 and 3.2 and in Tables 3.2, 3.3, and 3.4, which reveal an IME of 0.4 or higher.

For the simple Melitz-Pareto model to be consistent with the data, we need to move away from $\text{var}(\tilde{F}_{ij}) = 0$. As per Observation 2, however, the positive IME seen in the data implies a negative correlation between model-implied fixed and variable trade costs. Moreover, Observation 3 combined with the result in Table 3.6 of a negative distance elasticity of average exports per firm implies that model-implied fixed trade costs fall with distance. We explore these results further by using equations (3.12) and (3.13) to compute model-implied fixed and variable trade costs. The correlation between the resulting fixed and variable trade costs is $-0.786$ (with a standard error of 0.007). Figure 3.6 plots these trade costs against distance. The Figure shows that model-implied fixed trade costs are decreasing with distance, while model-implied variable trade costs are increasing with distance. The distance elasticities corresponding to Figure 3.6 are reported in Table 3.7. For fixed trade costs this elasticity is $-0.285$ (as per Observation 3, this is equal to the distance elasticity of average exports reported in Table 3.6) while for variable trade costs the distance elasticity is 0.272, both statistically significant.

Finally, according to Observation 4, the simple Melitz-Pareto model implies that $\text{IME}_{\text{perc}} = \text{IME}$ for all perc. This theoretical prediction of a common elasticity across percentiles is captured by the horizontal line red in Figure 3.3. This is at odds with the data.

To conclude, the simple version of the Melitz-Pareto model with fixed trade costs varying only because of origin and destination fixed effects is clearly at odds with the data. One can of course allow a richer pattern of variation in fixed trade costs across country pairs to make the model perfectly consistent with the data, but then the positive IME has further puzzling implications for fixed trade costs, which should fall with distance and be very negatively correlated with variable trade costs. To the best of our knowledge, there are no models that would microfound such a strong and negative correlation between the two types of trade costs and a negative fixed trade costs elasticity with respect to distance. The data is also at odds with the implication from the Melitz-Pareto model of a constant IME across exporter size percentiles.

---

18To compute model-implied variable trade costs as in equation (3.13), values for $\theta$ and $\bar{\theta}$ are required. We set $\theta = 5$ from Head and Mayer (2014) and $\sigma = 5$ from Bas, Mayer, and Thoenig (2015), which jointly imply $\bar{\theta} = 1.25$.

19Variable trade costs must increase with distance so that total exports fall with distance, as implied by the results in Table 3.6.

20Allowing for tariffs in addition to iceberg trade costs would naturally lead to a positive correlation between model-implied variable and fixed trade costs. This is because a tariff affects trade flows both by increasing the price of the affected good, as with iceberg trade costs, and by decreasing the net profits conditional on the quantity sold, as with fixed trade costs. See the online appendix of Costinot and Rodríguez-Clare (2014), Felbermayr, Jung, and Larch (2015), and Caliendo et al. (2015).
Multi-Product Extension of Melitz-Pareto

In this section we explore whether the puzzling implications for trade costs arising from the Melitz-Pareto model can be avoided by extending the model to multi-product firms. The idea would be that average exports per firm may fall along with total exports (thereby creating a positive IME) as firms facing higher product-level fixed trade costs export fewer products (even though they export more per product). Roughly speaking, allowing for multi-product firms implies that part of the extensive margin in the basic Melitz-Pareto model now operates inside the firm and appears as an intensive margin. We will see, however, that under the Pareto assumption the effect of higher product-level fixed trade costs on the number of products exported per firm is exactly offset by higher average exports per product, so that Observation 1 in the basic model will remain valid in this extension.

Theory

We consider an extension of the Melitz-Pareto model due to Bernard, Redding and Schott (2011). Each firm can produce a differentiated variety of each of a continuum of products in the interval \([0,1]\) with productivity \(\varphi \lambda\), where \(\varphi\) is common across products and \(\lambda\) is product-specific. The firm component \(\varphi\) is drawn from a Pareto distribution \(G^f(\varphi)\) with shape parameter \(\theta_f\), while the firm-product component \(\lambda\) is drawn from a Pareto distribution \(G^p(\lambda)\) with shape parameter \(\theta_p\). To have well-defined terms given a continuum of firms, we impose \(\theta_f > \theta_p > \sigma - 1\).

To sell any products in market \(j\), firms from country \(i\) have to pay a fixed cost \(F_{ij}\), and to sell each individual product requires an additional fixed cost of \(f_{ij}\). Variable trade costs are still \(\tau_{ij}\). The cutoff \(\lambda\) for a firm from country \(i\) with productivity \(\varphi\) that wants to export to market \(j\), \(\lambda_{ij}^*(\varphi)\), is given implicitly by

\[
A_j \left( \frac{w_i \tau_{ij}}{\varphi \lambda_{ij}^*(\varphi)} \right)^{1-\sigma} = \sigma f_{ij}. \tag{3.14}
\]

We can then write the profits in market \(j\) for a firm from country \(i\) with productivity \(\varphi\) as

\[
\pi_{ij}(\varphi) \equiv \int_{\lambda_{ij}^*(\varphi)}^{\infty} \left[ \frac{\lambda}{\lambda_{ij}^*(\varphi)} \right]^{\sigma-1} - 1 \right] f_{ij} dG_p(\lambda). \tag{3.15}
\]

The cutoff productivity for firms from \(i\) to sell in \(j\) is given implicitly by \(\pi_{ij}(\varphi_{ij}^*) = F_{ij}\). As in the canonical model, the number of firms from country \(i\) that export to market \(j\) is \(N_{ij} = \left[ 1 - G^f(\varphi_{ij}^*) \right] N_i\), while the number of products sold by firms from \(i\) in \(j\) is \(M_{ij} = N_i \int_{\lambda_{ij}^*(\varphi)}^{\infty} \left[ 1 - G^p(\lambda_{ij}^*(\varphi)) \right] dG^f(\varphi)\). Combining the previous expressions, using the fact that \(G^p(\lambda)\) and \(G^f(\varphi)\) are Pareto, writing \(f_{ij} = f_i^o f_j^d \tilde{f}_{ij}\), \(F_{ij} = F_i^o F_j^d \tilde{F}_{ij}\), and \(\tau_{ij} = \tau_i^o \tau_j^d \tilde{\tau}_{ij}\), and defining variables appropriately we get

\[
\ln X_{ij} = \mu_i^{X,o} + \mu_j^{X,d} - \theta f \ln \tilde{\tau}_{ij} - \left( \frac{\theta_f}{\sigma - 1} - \frac{\theta_f}{\theta_p} \right) \ln \tilde{f}_{ij} - \left( \frac{\theta_f}{\theta_p} - 1 \right) \ln \tilde{F}_{ij}, \tag{3.16}
\]
\[ \ln x_{ij}^p \equiv \ln X_{ij} - \ln M_{ij} = \mu_i^{x,p} + \mu_j^{x,p} + \ln \tilde{f}_{ij} , \quad (3.17) \]

and

\[ \ln x_{ij} \equiv \ln X_{ij} - \ln N_{ij} = \mu_i^{x,d} + \mu_j^{x,d} + \ln \tilde{F}_{ij} , \quad (3.18) \]

It is easy to verify that if \( f_{ij} = 0 \) for all \( i, j \) then this model collapses to the canonical model with single-product firms.

Recalling our definition of the intensive margin elasticity at the firm and product level introduced in Section 2 and letting \( \bar{\theta} \equiv \theta_f / (\sigma - 1) \) and \( \chi \equiv \theta_f / \theta^p \), then from equations (3.16) to (3.18) we have

\[ \text{IME} = -\frac{(\chi - 1) \text{var}(\ln \tilde{F}_{ij}) + (\bar{\theta} - \chi) \text{cov}(\ln \tilde{f}_{ij}, \ln \tilde{F}_{ij}) + \theta \text{cov}(\ln \tilde{F}_{ij}, \ln \tilde{\tau}_{ij})}{\text{var}(\ln \tilde{X}_{ij})} \quad (3.19) \]

and

\[ \text{IME}^p = -\frac{(\bar{\theta} - \chi) \text{var}(\ln \tilde{f}_{ij}) + (\chi - 1) \text{cov}(\ln \tilde{f}_{ij}, \ln \tilde{F}_{ij}) + \theta \text{cov}(\ln \tilde{F}_{ij}, \ln \tilde{\tau}_{ij})}{\text{var}(\ln \tilde{X}_{ij})} . \quad (3.20) \]

Observation 1 in the single-product firm model remains valid in the multi-product firm model, while we now have an analogous observation for the product-level intensive margin elasticity:

**Observation 5:** If \( \text{var}(\ln \tilde{f}_{ij}) = 0 \) then IME\(^p\) = 0.

The assumption \( \theta_f > \theta^p > \sigma - 1 \) implies that \( \chi > 1 \) and \( \bar{\theta} > \chi > 1 \) and in turn this leads to the following extensions of observation 2:

**Observation 6:** If IME > 0 then either \( \text{cov}(\ln \tilde{f}_{ij}, \ln \tilde{F}_{ij}) < 0 \) or \( \text{cov}(\ln \tilde{F}_{ij}, \ln \tilde{\tau}_{ij}) < 0 \) (or both).

**Observation 7:** If IME\(^p\) > 0 then either \( \text{cov}(\ln \tilde{f}_{ij}, \ln \tilde{F}_{ij}) < 0 \) or \( \text{cov}(\ln \tilde{f}_{ij}, \ln \tilde{\tau}_{ij}) < 0 \) (or both).

Observation 3 remains valid in the multi-product firm model, and we now also have an analogous observation for product-level fixed trade costs:

**Observation 8:** \( \frac{\text{cov}(\ln \tilde{x}_{ij}, \ln \text{dist}_{ij})}{\text{var}(\ln \text{dist}_{ij})} = \frac{\text{cov}(\ln \tilde{f}_{ij}, \ln \text{dist}_{ij})}{\text{var}(\ln \text{dist}_{ij})} \).

As in the single-product case, we can use the model to back out the implied trade costs. Equation (3.18) can be used to obtain a model-implied \( \tilde{F}_{ij} \) (which would be the same as the one derived in the single-product model) while Equation (3.17) can be used to obtain a model-implied \( \tilde{f}_{ij} \), and Equation (3.16) can then be used to obtain a model-implied \( \tilde{\tau}_{ij} \).

**Data**

Since Observations 1 and 3 remain valid when the basic model is extended to allow for multi-product firms, the conclusions regarding the necessity of having fixed trade costs decrease with distance remain valid. Turning to the implications for product-level fixed trade costs,
the finding in Section 2 of a positive IME at the product level, IME\(^p\) > 0 in Table 3.5, combined with Observation 5 implies that, to be consistent with the data, the multi-product version of the Melitz-Pareto model presented above requires \( \text{var}(\ln \tilde{f}_{ij}) > 0 \). However, observations 6 and 7 imply that the two types of fixed trade costs would need to be negatively correlated, or that the covariances between those fixed trade costs and variable trade costs would have to be negative. Moreover, Observation 8 combined with the results in Table 3.6 implies that model-implied product-level fixed trade costs decrease with distance with an elasticity of \(-0.060\), as shown in the third column of Table 3.7 and illustrated in Figure 3.7. We conclude that the puzzling implications of the Melitz-Pareto model remain valid when the model is extended to allow for multi-product firms.

### Firm-Level Demand and Fixed-Cost Shocks

EKK extend the basic Melitz-Pareto model presented in Section 3.1 to allow for (log-normally distributed) firm-level destination-specific demand and fixed-cost shocks. Except for constants that capture the net effects of these shocks, our equations (3.7) and (3.8) remain valid in the EKK environment, and hence so do observations 1-3.\(^{21}\)

It is important to note, however, that if productivity is distributed Pareto then the presence of log-normally distributed demand or fixed-cost shocks would imply that equations (3.7) and (3.8) no longer hold. The critical assumption in EKK that allows their model to be consistent with our equations (3.7) and (3.8) is that, loosely speaking, they consider the limit as the scale parameter of the Pareto distribution converges to zero.\(^{22}\)

To formally establish this result, recall that to get equations (3.7) and (3.8) we assumed that \( \varphi_{ij}^* > b_i \). If instead \( \varphi_{ij}^* \leq b_i \) then \( N_{ij} = N_i \) and \( x_{ij} = \left( \frac{\theta}{\theta-(\sigma-1)} \right) A_j \left( \frac{w_{ij}}{b_i} \right)^{1-\sigma} \). In the extreme, if \( \varphi_{ij}^* \leq b_i \) holds for all \( i, j \) pairs, then we would have \( \text{IME} = 1 \) rather than \( \text{IME} = 0 \). Now think about the case with firm-specific demand and fixed-cost shocks. Specifically, assume that each firm is characterized by a productivity level \( \varphi \) as well as a demand shock \( \alpha_j \) and a fixed cost shock \( f_j \) in each destination \( j \), with \( \varphi \) drawn from a Pareto distribution (with scale parameter \( b_i \) and shape parameter \( \theta \)) and \( \alpha_j \) and \( f_j \) drawn iid from some distribution. Let \( x_{ij}(\varphi, \alpha_j) = A_j \alpha_j (\sigma \frac{w_{ij}}{\varphi} (1-\sigma)) \) and let \( \varphi_{ij}^*(\alpha_j, f_j) \) be implicitly defined by \( x_{ij}(\varphi_{ij}^*, \alpha_j) = \sigma f_j \). By the same argument we used in Section 3.1, if for all \( i, j \) and all possible \( (\alpha_j, f_j) \) we have \( \varphi_{ij}^*(\alpha_j, f_j) > b_i \), we can easily show that we still have \( \text{IME} = 0 \).\(^{23}\) However, if \( \alpha_j \) and \( f_j \) are lognormally distributed, then for \( b_i > 0 \) for all \( i \) there must be a

\(^{21}\)This can be confirmed by simple manipulation of equations (20) and (28) in EKK.

\(^{22}\)More exactly, EKK specify a function for the measure of firms with productivity above some level, with that measure going to infinity as productivity goes to zero. This is equivalent to taking a limit with the (exogenous) measure of firms going to infinity and the scale parameter of the Pareto distribution going to zero. Although equations (3.7) and (3.8) do not hold anywhere in this sequence, they do hold in the limit.

\(^{23}\)Consider the group of firms from country \( i \) that have some given draw \( \{(\alpha_j, f_j), j = 1, \ldots, n\} \). The exact same argument used in Section 3.1 can be used to show that the sample of firms obtained by combining such firms across all origins \( i \) satisfies \( \text{IME} = 0 \). One can then simply integrate across all possible draws \( \{(\alpha_j, f_j), j = 1, \ldots, n\} \) to show that \( \text{IME} = 0 \) for the whole set of firms.
positive mass of firms for which \( \varphi_{ij}^*(\alpha_j, f_j) < b_i \), and for those firms there would be a positive intensive margin elasticity. EKK essentially avoid this by taking the limit with \( b_i \to 0 \) for all \( i \).

In principle, one could use this result to argue that a Melitz model with Pareto distributed productivity but extended to allow for log-normally distributed demand and fixed-cost shocks could match the positive IME that we see in the data. However, such a model would not exhibit any of the convenient features of the canonical Melitz-Pareto model: the sales distribution is not distributed Pareto, the trade elasticity is not common across country pairs and fixed, and the gains from trade are not given by the ACR formula. Given that, our approach in this paper is to move all the way to a model where productivity as well as destination-specific demand and fixed-cost shocks are lognormally distributed. Such a model has at least the advantage that it is computationally tractable, and amenable to Maximum Likelihood Estimation, as we show in Section 4.

**Granularity**

The previous sections have considered a model with a continuum of firms. With a discrete and finite number of firms it may be possible to generate a positive covariance between the intensive margin and total exports that could in principle explain our empirical findings. We explore this possibility in this section.

**Theory**

Eaton, S. S. Kortum, and Sotelo (2012) extend the Melitz-Pareto model above to allow for granularity. Equations (3.9) and (3.10) then become

\[
\ln N_{ij} = \mu_i^{N,o} + \mu_j^{N,d} - \theta \ln \tilde{\tau}_{ij} - \bar{\theta} \ln \tilde{F}_{ij} + \xi_{ij} \tag{3.21}
\]

and

\[
\ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln \tilde{F}_{ij} + \varepsilon_{ij}, \tag{3.22}
\]

where \( \xi_{ij} \) and \( \varepsilon_{ij} \) are error terms arising from the fact that now the number of firms is discrete and random. Using the same definition for the intensive margin elasticity as in Section 3, the previous equations imply that

\[
IME = - (\bar{\theta} - 1) \frac{\text{var}(\ln \tilde{F}_{ij}) - \theta \text{cov}(\ln \tilde{\tau}_{ij}, \ln \tilde{F}_{ij}) + \text{var}(\varepsilon_{ij}) + \text{COV}}{\text{var}\left(-\theta \ln \tilde{\tau}_{ij} - (\bar{\theta} - 1) \ln \tilde{F}_{ij} + \varepsilon_{ij} + \xi_{ij}\right)}, \tag{3.23}
\]

where \( \text{COV} \equiv \text{cov}(\ln \tilde{F}_{ij} + \varepsilon_{ij}, \xi_{ij}) + \text{cov}(\ln \tilde{\tau}_{ij} + \ln \tilde{F}_{ij}, \varepsilon_{ij}) \). If \( \text{var}(\varepsilon_{ij}) \) is large relative to \(-\text{COV}\), this could explain \( IME > 0 \) even with \( \text{cov}(\ln \tilde{F}_{ij}, \ln \tilde{\tau}_{ij}) > 0 \). Thus, in theory, granularity could explain the positive intensive margin elasticity that we find in the data without relying on implausible patterns for fixed trade costs.
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

To check whether granularity is a plausible explanation for the positive IME in the data we will conduct two tests. First, we will estimate the fixed trade cost elasticity with respect to distance taking into account granularity and the possible biases it may induce. Second, we will simulate firm-level exports under granularity and the assumption of fixed trade costs that vary by origin and destination only and estimate the implied IME. We describe each of these tests in turn.

Fixed Trade Costs and Distance with Granularity

In the Melitz-Pareto model with a continuum of firms, average exports per firm can be expressed as \(x_{ij} = \kappa F_{ij}\), where \(\kappa \equiv \frac{\sigma \bar{\theta}}{\bar{\theta} - 1}\). If we relax the continuum assumption to allow for granularity, then average exports per firm can be expressed as \(x_{ij} = \kappa F_{ij} + \varepsilon_{ij}\), where \(\varepsilon_{ij}\) is an error term that arises from random realizations of productivity draws, the first moment of which is independent of any variables that determine bilateral fixed trade costs. If we further assume that \(F_{ij} = F^o_{i}F^d_{j}e^{\zeta \ln \text{dist}_{ij}} + v_{ij}/\kappa\), where \(v_{ij}\) satisfies \(\mathbb{E}(v_{ij}|\ln \text{dist}_{ij}) = 0\), we can then write

\[
    x_{ij} = \kappa^{o}_{i}F^d_{j}e^{\zeta \ln \text{dist}_{ij}} + u_{ij},
\]

where \(u_{ij} \equiv v_{ij} + \varepsilon_{ij}\) is an error term that captures both the deviation of \(F_{ij}\) from its mean as well as the granularity error term \(\varepsilon_{ij}\). Since both \(\mathbb{E}(v_{ij}|\ln \text{dist}_{ij})\) and \(\mathbb{E}(\varepsilon_{ij}|\ln \text{dist}_{ij})\) are equal to zero, it follows that \(\mathbb{E}(u_{ij}|\ln \text{dist}_{ij}) = 0\). The challenge in estimating the fixed trade costs elasticity with respect to distance, \(\zeta\), from this equation is that we cannot simply take logs to obtain a log-linear equation to be estimated by OLS, because the error term that comes from granularity is not log-additive.

To take advantage of the time dimension of our data, we extend (3.24) to allow for an origin-time and destination-time specific components in the expression of fixed trade costs,

\[
    x_{ijt} = \kappa^{o}_{it}F^d_{jt}e^{\zeta \ln \text{dist}_{ijt}} + u_{ijt},
\]

where again \(\mathbb{E}(u_{ijt}|\ln \text{dist}_{ijt}) = 0\). We estimate (3.25) using Poisson pseudo maximum likelihood method as in J. S. Silva and Tenreyro (2011) in the next subsection.

The IME under Granularity: Simulation

To assess how well granularity can explain a positive IME, we simulate exports of \(N_{ij}\) firms for each of the country pairs in the sample. We add demand shocks to allow for a less than perfect correlation between exports of different firms across different destinations. In the standard Melitz model with demand shocks, exports from \(i\) to \(j\) of a firm with productivity \(\varphi\) and destination-specific demand shock \(\alpha_{j}\) can be calculated as

\[
    x_{ij} (\varphi, \alpha_{j}) = \sigma F_{ij} \left( \frac{\alpha_{ij} \varphi}{\alpha_{ij}^{*} \varphi_{ij}^{*}} \right)^{\sigma-1},
\]

(3.26)
where $\alpha^*_ij\varphi^*_ij$ is a combination of productivity and demand shocks of the smallest exporter from $i$ selling to $j$. To estimate the IME in simulations we perform the following steps:

1. Draw $\varphi$ and $\alpha_j$ from some distribution. The number of draws is equal to $N_{ij}$, the number of exporters in the EDD dataset for each origin-destination pair in 2009. To be more precise, we draw the product $\alpha_j\varphi$ for each firm-destination pair assuming either that, as in the standard Melitz model, there are no demand shocks and hence the product $\alpha_j\varphi$ is perfectly correlated across destinations or that, at the other extreme, there is no correlation in the product $\alpha_j\varphi$ across destinations (pure demand shocks case). In both cases, we draw $\alpha_j\varphi$ from a Pareto distribution with a shape parameter to be specified below.

2. Assume that $\text{var} \left( \tilde{F}_{ij} \right) = 0$, so that $F_{ij} = F_i^o F_j^d$. This will allow us to study the IME generated by granularity by itself.

3. Use equation (3.26) to simulate the exports for each firm and to calculate average exports per firm (in total and in each percentile) for each origin-destination pair.

4. Run the IME regression 3.1 on the simulated export data, with $\ln x_{ij}$ being either the intensive margin for all firms exporting from $i$ to $j$, or for each percentile in the size distribution of exporters from $i$ to $j$.

Data

We now discuss the evidence obtained first for the fixed trade costs elasticity with respect to distance and second for the IME with simulated data.

We use equation (3.25) to estimate firm-level as well as product-level fixed trade cost elasticities with respect to distance ($\zeta$). Table 3.8 shows that both of these elasticities are negative and statistically significant, so both firm-level and product-level model-implied fixed trade costs are decreasing with distance, although with a much smaller elasticity than when not accounting for granularity (compare results of Tables 3.7 and 3.8). Hence granularity does not help to eliminate one of the puzzles emerging from the comparison between the Melitz-Pareto model and the data.

Table 3.9 reports the estimated IME using simulated data for alternative values of $\bar{\theta}$ and for either zero or perfect correlation between the product of demand and productivity shocks across destinations. We consider 4 values of $\bar{\theta}$: our estimate $\bar{\theta} = 2.4$, the value that can be inferred from standard estimates of $\theta$ and $\sigma$ in the literature (i.e., $\theta = 5$, the central estimate of the trade elasticity in Head and Mayer, 2014, and $\sigma = 5$ from Bas, Mayer, and Thoenig (2015), so $\bar{\theta} = 1.25$), as well as $\bar{\theta} = 1.75$ from Eaton, S. Kortum, and Kramarz (2011) (which they estimate using the procedure outlined in the Appendix C.1) and $\bar{\theta} = 1$ (as in Zipf’s Law).

Two broad patterns emerge from the table. First, the simulated IME decreases with $\bar{\theta}$. This is because the effect of granularity on the IME is stronger when there is more
dispersion in productivity levels. Second, the simulated IME is highest when productivity is less correlated across destinations, again because this gives granularity more room to generate a covariance between average exports per firm and total exports.

For our estimate of $\bar{\theta}$ ($\bar{\theta} = 2.4$) and with no demand shocks (so there is perfect correlation in firm-level productivity across destinations), the simulated IME of 0.001 is quite low. The highest simulated IME occurs for the case in which $\bar{\theta} = 1$ and there is no correlation between the product of demand shocks and productivity across destinations. In this case the simulated IME is 0.33, not too far from our preferred estimate based on the data of 0.4. But we think of this as an extreme case because $\bar{\theta} = 1$ is far from the estimates that come out of trade data, and because of the implausible assumption that firm-level exports are completely uncorrelated across destinations.

To explore this further, we examine the implications for the IME across percentiles. We calculate average simulated exports per firm in each percentile and use those to estimate an IME per percentile. We plot the resulting 100 IME estimates in Figure 3.8 along with the corresponding IME estimates based on the actual data. The IME based on the actual data is increasing with a spike at the top percentile. Granularity and the Pareto distribution fail to reproduce this pattern in the simulated data, since the corresponding IME is much smaller than in the data for most percentiles. The IME in the simulated data is almost zero for small percentiles and is relatively high for a small number of top percentiles. We conclude that granularity does not offer a plausible explanation for the positive estimated IME in the data.

### 3.4 The Intensive Margin in the Melitz-Lognormal Model

In this section we depart from the assumption of a common Pareto distribution of firm-level productivity and instead assume a lognormal distribution.\footnote{One could consider combining a lognormal distribution with a Pareto distribution on the right tail, as in Nigai (2017). We have used Nigai's Matlab code on our data to estimate the point of truncation (percentile) where the lognormal ends and the Pareto begins. We find that for 75% of country pairs with more than a hundred exporters the point of truncation occurs after the 99th percentile, and for the median country pair the truncation point is at the 99.9%. In light of these results, in the rest of the paper we focus on the case in which productivity is described by a fully lognormal distribution.} In the theory section we start by showing how this can lead to a positive IME in a simple Melitz model, and then propose a maximum-likelihood estimation procedure for a richer Melitz model with heterogeneous fixed costs and demand shocks. The data section presents the results from the estimation and the implications for the IME as well as for the model-implied trade costs.

We also develop a flexible methodology to calculate counterfactual changes in trade flows in response to a change in trade costs. We improve Dekle, Eaton, and S. Kortum (2008) “exact hat algebra” approach that can accommodate any distribution of productivity, de-
mand, and fixed cost shocks. Using this approach we emphasize important differences in counterfactual trade flows responses in the Melitz-Pareto and lognormal models.

Theory

A simple Melitz model with a lognormal distribution

Consider a model as that presented in Section 2 but with general CDF and PDF $G_i(\phi)$ and $g_i(\phi)$. The ratio of average to minimum exports per firm for each country pair can be written as

$$\frac{x_{ij}}{x_{ij}(\phi_{ij}^*)} = \left(\frac{\bar{\phi}_i(\phi_{ij}^*)}{\phi_{ij}^*}\right)^{\sigma^{-1}},$$

(3.27)

where $\bar{\phi}_i(\phi^*)$ is the average productivity level defined in Melitz 2003,

$$\bar{\phi}_i(\phi^*) \equiv \left(1 - \frac{1}{1 - G_i(\phi^*)} \int_{\phi^*}^{\infty} \phi^\sigma g_i(\phi) d\phi\right)^{\frac{1}{\sigma}}.$$

If firm productivity is distributed Pareto with parameter $\theta > \sigma - 1$ (as in Section 2) then

$$\frac{\bar{\phi}(\phi^*)}{\phi_{ij}^*} = \frac{\theta}{\sigma - 1}$$

for all $\phi_{ij}^*$, (see equation 3.8), so that the average to minimum ratio does not depend on selection. This property only holds with a Pareto distribution of productivity.

As argued in footnote 15 of Melitz 2003, $\bar{\phi}(\phi^*)$ is decreasing if the distribution $g_i(\phi)$ “belongs to one of several common families of distributions: lognormal, exponential, gamma, Weibul, or truncations on $(0, +\infty)$ of the normal, logistic, extreme value, or Laplace distributions. (A sufficient condition is that $g_i(\phi)/(1 - G_i(\phi))$ be increasing to infinity on $(0, +\infty)$.)” To understand the implication of this property, consider a decline in $\tau_{ij}$, so that $\phi_{ij}^*$ decreases with no effect on minimum sales (which remain at $\sigma F_{ij}$). The decline in $\tau_{ij}$ leads to an increase in exports of incumbent firms (which increases average exports per firm) and entry of low productivity firms (which decreases average exports per firm). Under Pareto these two effects exactly offset each other so there is no change in average exports per firm. If productivity is distributed in such a way that $\bar{\phi}(\phi^*)$ is decreasing then the second effect does not fully offset the first, and hence average exports per firm increase with a decline in $\tau_{ij}$. Since this also increases the number of firms that export (and hence total exports), the result is a positive IME.

We now explore the magnitude of the implied IME in the case where $g_i(\phi)$ is lognormal for every origin country $i$. Assuming that $g_i(\phi)$ is lognormal with location parameter $\mu_{\phi,i}$ and scale parameter $\sigma_{\phi}$, and letting $\Phi(\cdot)$ be the CDF of the standard normal distribution, then

$$G_i(\phi) = \Phi\left(\frac{\ln \phi - \mu_{\phi,i}}{\sigma_{\phi}}\right).$$

(3.28)
Letting \( h(x) \equiv \Phi'(x)/\Phi(x) \) be the ratio of the PDF to the CDF of the standard normal, Bas, Mayer, and Thoenig (2015) (henceforth BMT), show that

\[
\frac{\bar{\varphi}(\varphi^*_{ij})}{\varphi^*_{ij}} = \frac{h[-(\ln \varphi^*_{ij} - \mu_{\varphi,i})/\sigma_{\varphi}]}{h[-(\ln \varphi^*_{ij} - \mu_{\varphi,i})/\sigma_{\varphi} + \bar{\sigma}_{\varphi}]},
\]

(3.29)

where \( \bar{\sigma}_{\varphi} \equiv (\sigma - 1) \sigma_{\varphi} \). Combined with \( 1 - G_i(\varphi^*_{ij}) = \frac{N_{ij}}{N_i} \), we have

\[
\frac{x_{ij}}{x_{ij}(\varphi^*_{ij})} = \Omega \left( \frac{N_{ij}}{N_i} \right) = \frac{h \left( \Phi^{-1} \left( \frac{N_{ij}}{N_i} \right) \right)}{h \left( \Phi^{-1} \left( \frac{N_{ij}}{N_i} \right) + \bar{\sigma}_{\varphi} \right)}.
\]

(3.30)

Thus, the average to minimum ratio of exports per firm for country pair \( ij \) only depends on the share of total firms in \( i \) that export to \( j \), with the relationship given by the function \( \Omega(\cdot) \).

As argued by BMT, \( \Omega(\cdot) \) is an increasing function, which is consistent with the observation by Melitz 2003 above that \( \tilde{\varphi}(\varphi^*_{ij}) \varphi^*_{ij} \) is decreasing in \( \varphi^*_{ij} \) since \( N_{ij}/N_i \) is decreasing in \( \varphi^*_{ij} \).

Given values of \( \bar{\sigma}_{\varphi} \) as well as \( N_i \) for every country, we can use our data on \( N_{ij} \) to compute \( \Omega \left( \frac{N_{ij}}{N_i} \right) \) for all country pairs. Combined with \( x_{ij}(\varphi^*_{ij}) = \sigma F_{ij} \) and imposing \( F_{ij} = F^o F^d \), we can use equation (3.30) to get the model-implied average exports per firm (in logs),

\[
\ln x_{ij} = \mu^x_o + \mu^x_d + \ln \Omega \left( \frac{N_{ij}}{N_i} \right).
\]

(3.31)

In contrast to Observation 1 for the Melitz-Pareto model, under under lognormality we will have a positive IME even with \( \text{var}(\tilde{F}_{ij}) = 0 \).

We can also compute model-implied fixed and variable trade costs similarly to what we did under the assumption of Pareto-distributed productivity. First, we obtain \( \tilde{F}_{ij} \) from

\[
\ln \tilde{F}_{ij} = \delta^F_o i + \delta^F_d j + \ln x_{ij} - \ln \Omega \left( \frac{N_{ij}}{N_i} \right).
\]

(3.32)

Second, to compute \( \tilde{\tau}_{ij} \), we combine equations (3.3), (3.38), (3.28) and (3.30) to get (with appropriately defined fixed effects)

\[
(\sigma - 1) \ln \tilde{\tau}_{ij} = \delta^\tau_o i + \delta^\tau_d j - \ln x_{ij} + \ln \Omega_i \left( \frac{N_{ij}}{M_i} \right) + \bar{\sigma}_{\varphi} \Phi^{-1} \left( 1 - \frac{N_{ij}}{N_i} \right).
\]

(3.33)

Armed with estimates of \( \tilde{F}_{ij} \) and \( (\sigma - 1)\tilde{\tau}_{ij} \), we can compute their correlation and check whether \( \tilde{F}_{ij} \) increases or decreases with distance (demeaned by origin and destination fixed effects).

These empirical exercises require estimates for \( \bar{\sigma}_{\varphi} \) as well as \( N_i \) for every country. We use Bento and Restuccia (2015) (henceforth BR) data to estimate a value for \( N_i \) for all the
countries in our sample.\footnote{Using census data as well as numerous surveys and registry data, BR compiled a dataset with the number of manufacturing firms for a set of countries. Unfortunately, the sample in BR has missing observations for a number of countries in the EDD. We impute missing values projecting the log number of firms on log population. There is a tight positive relationship between log number of firms in the BR dataset and log population with an elasticity of 0.945, as reported in Table 3.10 and in Figure 3.9.} We acknowledge slippage between theory and data in that we obviously do not have a measure of the entry level $N_i$, but (at best) only for the number of existing firms, which in theory would correspond to $(1 - G_i(\varphi^*_i)) N_i$ (our approach in the next subsection avoids this problem). We use the QQ-estimation proposed by Head, Mayer, and Thoenig (2014) (henceforth HMT) to obtain estimates of $\sigma_\varphi$ and $\mu_{\varphi,i}$ for every $i$ (see Appendix C.2 for a detailed description).

Full Melitz-lognormal model

The previous section has shown that a model with a lognormal distribution of firm productivity is capable of generating a positive intensive margin elasticity conditional on fixed costs. However, the model we considered had two very stark predictions. First, fixed trade costs that are common across firms lead to the prediction that sales of the least productive exporter from $i$ to $j$ are equal to $\sigma F_{ij}$. In the data we observe many firms with very small export sales (sometimes as low as $1$) which implies unrealistic fixed trade costs. Second, as shown by Eaton, S. Kortum, and Kramarz (2011), the model implies a perfect hierarchy of destination markets (i.e., destinations can be ranked according to profitability, with all firms that sell to a destination also selling to more profitable destinations) and perfect correlation of sales across firms that sell to multiple markets from one origin. None of these predictions holds in the data.

In this section we consider a richer model with firm-specific fixed trade costs and demand shocks that vary by destination. This is similar to the setup in Eaton, S. Kortum, and Kramarz (2011). We assume that firm productivity, demand shocks (denoted by $\alpha_j$) and fixed trade costs (denoted by $f_j$) are distributed jointly lognormal, i.e., for each origin $i$:

$$\begin{bmatrix}
\ln \varphi \\
\ln \alpha_1 \\
\vdots \\
\ln \alpha_J \\
\ln f_1 \\
\vdots \\
\ln f_J
\end{bmatrix} \sim N
\begin{bmatrix}
\mu_{\varphi,i} \\
\mu_{\alpha} \\
\vdots \\
\mu_{\alpha,j} \\
\mu_{f,i} \\
\vdots \\
\mu_{f,i,J}
\end{bmatrix},
\begin{bmatrix}
\sigma^2_{\varphi,i} & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & \sigma^2_{\alpha,i} & \ldots & \sigma^2_{\alpha,i} & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \sigma^2_{f,i} & 0 & \ldots & 0 \\
0 & 0 & \ldots & \sigma^2_{f,i} & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \sigma^2_{f,i} & 0 & \ldots & 0
\end{bmatrix}.$$ (3.34)

Note that we allow mean log productivity to be origin-specific while imposing that the mean of demand shocks be the same across origin-destination pairs (however, we cannot separately identify these parameters). Mean fixed costs are allowed to vary across origin-destination pairs and can be correlated with demand shocks within destinations.
empirical estimation we will not be able to separately identify mean productivity from wages and variable trade costs – they will all be absorbed into an origin-destination fixed effect. Also, we allow the dispersion of log productivity, log demand shocks and log fixed trade costs to differ across all origins. We restrict the dispersion of log demand shocks and log fixed trade costs to be the same across destinations within a given origin.

Without risk of confusion, we change notation in this section and use \( X_i \equiv (X_{i1}, ..., X_{iJ}) \) to denote the random variable representing log sales of a firm from \( i \) in each of the \( J \) destinations, with \( x_i \equiv (x_{i1}, ..., x_{iJ}) \) being a realization of \( X_i \), and \( g_{X_i}(x_i) \) being the associated probability density function. According to the model, a firm does not export to destination \( j \) if it has a large fixed trade cost draw \( f_{ij} \) relative to its productivity and its demand shock for that destination. Let \( D_{ij} \equiv \ln \left[ A_j \left( w_i \tau_{ij} \right)^{\frac{1}{\sigma}} \right] \) and let \( Z_{ij} \equiv D_{ij} + \ln \alpha_j + (\sigma - 1) \ln \varphi \) be sales in destination \( j \) by a firm from \( i \) with productivity \( \varphi \) and demand shock \( \alpha_j \). This is a latent variable that we observe only if a firm actually exports,

\[
X_{ij} = \begin{cases} 
Z_{ij} & \text{if } \ln \sigma + \ln f_{ij} \leq Z_{ij} \\
\emptyset & \text{otherwise} 
\end{cases},
\]

with \( Z_i \equiv (Z_{i1}, ..., Z_{iJ}) \) distributed according to

\[
\begin{bmatrix} Z_{i1} \\
\vdots \\
Z_{iJ} \end{bmatrix} \sim N \left( \begin{bmatrix} d_{i1} \\
\vdots \\
 d_{iJ} \end{bmatrix} , \begin{bmatrix} \bar{\sigma}_{\varphi,i}^2 + \sigma_{\alpha,i}^2 & \cdots & \bar{\sigma}_{\varphi,i}^2 \\
\vdots & \ddots & \vdots \\
\bar{\sigma}_{\varphi,i}^2 & \cdots & \bar{\sigma}_{\varphi,i}^2 + \sigma_{\alpha,i}^2 \end{bmatrix} \right),
\]

(3.35)

where \( d_{ij} \equiv D_{ij} + \mu_{\alpha} + (\sigma - 1) \mu_{\varphi,i} \) and \( \bar{\sigma}_{\varphi,i} \equiv (\sigma - 1) \sigma_{\varphi,i} \).

Using firm-level data from the EDD and China across different origins and destinations, we can estimate the parameters in (3.35) as well as mean log fixed trade costs (up to a constant) and their dispersion using maximum likelihood methods.\(^{26}\) Appendix C.3 shows how to derive the density function \( g_{X_{i1},...,X_{iJ}}(x_{i1}, ..., x_{iJ}) \) for the case when we observe sales to \( J \) destinations. We simplify the analysis by considering only data for fifteen destinations (USA, Germany, Japan, France and the 11 largest destinations by exports value), which we label \( j = 1, ..., 15 \) for the year 2007 for each of 39 origins. We compute \( g_{X_{i1},...,X_{iJ}}(x_{i1}, ..., x_{iJ}) \) for each observation in our dataset (which is a realization of \( \{X_{i1}, ..., X_{iJ}\} \) that we observe). Since all random variables are independent across firms, we can compute the log-likelihood function as a sum of log-densities,

\[
\ln L \left( \theta_i \mid \{x_{i1}(k_i), ..., x_{iJ}(k_i)\}_{i,k_i} \right) = \sum_{k_i=1}^{N_i} \ln [g_{(X_{i1},...,X_{iJ})}(x_{i1}(k_i), ..., x_{iJ}(k_i))],
\]

(3.36)

where \( N_i \) is the number of firms from \( i \) that sell to either of the fifteen destinations we consider, and where \( k_i \) is an index for a particular observation in our dataset (for origin \( i \).

\(^{26}\)We took all of the firms from the EDD and 5% of exporters from China for computational reasons.
it takes values in 1,...,N_{ij}) and \theta_i is an origin-specific vector of parameters that we want to estimate, 
\[ \theta_i = \left\{ \{d_{ij}, \bar{\mu}_{f,ij}\}, \bar{\sigma}_{\varphi,i}, \sigma_{\alpha,i}, \sigma_{f,i}, \rho_i \right\}, \tag{3.37} \]
where \(\bar{\mu}_{f,ij} = \ln \sigma + \mu_{f,ij}\) and \(\rho = \frac{\sigma_{f,i}}{\sigma_{\alpha,i} \sigma_{f,i}}\). As the likelihood is potentially not concave in \(\theta_i\) and because there are 34 parameters per origin to estimate, we rely on the estimation methodology proposed by Chernozhukov and Hong (2003) (a detailed description of the algorithm can be found in the Appendix C.4). We use the Metropolis-Hastings MCMC algorithm to construct a chain of estimates \(\theta_i^{(n)}\) for each origin country. Chernozhukov and Hong (2003) show that \(\bar{\theta} \equiv \frac{1}{N} \sum_{n=1}^{N} \theta_i^{(n)}\) is a consistent estimator of \(\theta_i\), while the covariance matrix of \(\theta_i\) is given by the variance of \(\theta_i^{(n)}\), so we use this to construct confidence intervals for \(\theta_i\). For each origin, we run 5 different chains that start at a different random starting value \(\theta_i^{(i)}\). We then explore whether the different parameters in \(\theta_i\) converged to the same values across different chains and discuss the convergence of the chains in the Appendix C.5.

Loosely speaking, identification works as follows. First, data on export flows and the number of exporters across country pairs helps in identifying the sum of the dispersion parameters for productivity and demand shocks, \(\bar{\sigma}_{\varphi,i} + \sigma_{\alpha,i}\). Third, the extent of correlation of firm sales from a particular origin across different destinations helps in identifying \(\sigma_{\varphi,i}\) separately from \(\sigma_{\alpha,i}\): the more correlated firm sales are across destinations, the larger is \(\sigma_{\varphi,i}\) relative to \(\sigma_{\alpha,i}\). Fourth, the correlation between fixed costs and demand shocks can be inferred from the distribution of sales of small firms. Intuitively, if correlation is negative, then a firm with a bad demand shock would also likely draw a high fixed cost shock and thus will not export, hence, we would not see a lot of small firms in the data. Finally, to understand how \(\sigma_{f,i}\) is identified, imagine for simplicity that there is only one destination.

We then have
\[ g_{X_{i1}}(x_{i1}) = \frac{g_{Z_{i1}}(x_{i1}) \times \Pr \{\ln \sigma + \ln f_{i1} \leq x_{i1} | Z_{i1} = x_{i1}\}}{C} \]
where \(C \equiv \Pr \{\ln \sigma + \ln f_{i1} \leq Z_{i1}\}\) and where \(g_{Z_{i1}}()\) is the probability density function of the latent sales \(Z_{i1}\). This implies that we can get the density of \(X_{i1}\) by applying weights \(\frac{\Pr \{\ln \sigma + \ln f_{i1} \leq x_{i1} | Z_{i1} = x_{i1}\}}{C_i}\) to the density of \(Z_{i1}\). The parameter \(\sigma_{f,i}\) regulates how these weights behave with \(x_{i1}\). In the extreme case in which \(\sigma_{f,i} = 0\) then the weights are 0 for \(x_{i1} \leq \mu_{f,i}\) and \(1/C\) for \(x_{i1} > \mu_{f,i}\), while in the other extreme with \(\sigma_{f,i} = \infty\) the weights are all equal to 1. For intermediate cases the density of \(X_{i1}\) will be somewhere in the middle, with the left tail becoming fatter and the right tail becoming thinner as \(\sigma_{f,i}\) increases. This suggests that we can identify \(\sigma_{f,i}\) from the shape of the density of sales.

We will use the results of the estimation to conduct exercises similar to those in the previous sections. First, we will compute the IME for all firms and for each percentile using the estimated model. Second, after removing origin and destination fixed effects, we will compute the correlation across the estimated values of \(d_{ij}\) and \(\bar{\mu}_{f,ij}\), and between them and distance.
Data

Simple Melitz model with lognormal distribution

Appendix Table C.9 reports the QQ-estimate of $\bar{\sigma}_\varphi$. We report three sets of estimates: for the full sample, the largest 50% of firms and the largest 25% of firms for each origin-destination pair in each year. These estimates are on the high side relative to the estimate obtained by HMT, so we will use the minimum among them, $\bar{\sigma}_\varphi = 4.02$, which corresponds to the subsample with the largest 25% of firms.\footnote{See the section in the Appendix titled QQ-Estimation of $\bar{\sigma}_\varphi$ for a discussion of these estimates and their relation to the estimate in HMT.} We highlight three findings. First, a lognormal distribution allows the intensive margin elasticity to be positive even under the assumption of a continuum of firms. Second, for our estimate of the shape parameter, the implied IME is 0.28, which is close to that from the data.\footnote{Using Head, Mayer, and Thoenig (2014) estimate of $\bar{\sigma}_\varphi = 2.4$ we get IME of around 0.12} Third, most of the action comes from the right tail of the exporter size distribution, as seen in Figure 3.10.

We use equations (3.32) and (3.33) to compute the model-implied fixed and variable trade costs. The correlations between those costs and distance are reported in Table 3.11 and plotted in Figure 3.11. In contrast to our results under Pareto, now under lognormal both the model-implied variable and fixed trade costs are increasing with distance.

Overall, the model does much better in fitting the data when we assume that firm productivity is distributed lognormal than when we assume that it is distributed Pareto. However, the IME for each percentile is not a perfect match to the data, and the there is still a negative correlation between the model implied variable and fixed trade costs, although it is much closer to zero than with Pareto (-0.3 rather than -0.8). In any case, this is just a “proof of concept” that lognormally-distributed productivity can by itself improve the performance of the model relative to the data. In the next subsection we present the results obtained with the estimated full Melitz-lognormal model.

Full Melitz-lognormal model

To estimate the parameters of the full Melitz-lognormal model we use firm-level data from the EDD and China for the year 2007. This entails 39 different origins, but we had to drop 2 origins from our analysis due to convergence issues discussed in the Appendix C.5. We consider 15 destinations: USA, Germany, Japan, France and other 11 biggest destinations for each origin.

Before presenting the results of the estimation and discussing their implications for the IME, we show three figures revealing the fit of the estimated model with the data. Figure 3.12 shows a plot of the density function for standardized firm-level log sales pooled across multiple origin and destinations.

We next look at deviations from the strict hierarchy of firms sales across destinations (for each origin) in the data and in the estimated model. If there were no demand and fixed cost shocks across firms, then all firms from a given origin that export to less popular
destinations would also export to the most popular destination. The share of firms that only sell in the less popular destinations is then a measure of the extent to which this strict hierarchy predicted by the simplest model is violated. According to Figure 3.13, the share predicted by the estimated model is quite close to the one in the data.

Finally, Figure 3.14 shows the correlation in sales across the U.S. and Germany firms from a given origin that sell in both destinations. The estimated model mostly implies positive correlation driven by firm-level productivity shocks, while in the data this correlation exhibits more dispersion.

The results of the estimation for variance-covariance parameters \((\bar{\sigma}_{\varphi,i}, \sigma_{\alpha,i}, \sigma_{f,i}, \rho_i)\) are shown in Table 3.12. The median estimated values for \(\bar{\sigma}_{\varphi,i}\) and \(\sigma_{\alpha,i}\) across 37 origins are 3.19 and 2.68 respectively, while the median estimates for \(\sigma_{f,i}\) and \(\rho_i\) are 2.43 and 0.49. Even though the variance-covariance parameters were precisely estimated for each of the origins, the parameters vary quite a bit across different origins. In general, there is positive correlation between demand and fixed costs shocks, but some origins exhibit a negative correlation.

Table 3.13 and Figure 3.15 show the implications of the estimated model for the IME. We compute the IME implied by the estimated model by drawing one million firms for each origin (this implies one million latent log sales and log fixed costs for each destination), computing average sales (taking into account selection), and then multiplying average sales by \(N_{ij}\) in the data to compute total exports. We pick one million because at this point we are not interested in granularity – this is just a numerical approximation to the case with a continuum of firms. The IME implied by the model is 0.64. This is actually higher than our preferred IME estimate of 0.4 in Section 2, but the gap comes in large part from the different sample of origin-destination pairs used here. Using the same sample of 37 origins and 4 destinations for the year 2007 we estimate IME of 0.67 (with a standard error of 0.03) that is statistically indistinguishable from the one implied by our estimated lognormal model.\(^{29}\) We plot the associated IME for each percentile in Figure 3.15 – the pattern of the IME across percentiles is remarkably close to what we see in the data.

Table 3.14 shows the elasticity of variable and fixed trade costs with respect to distance (controlling for origin and destination fixed effects). Now both types of trade costs are strongly increasing in distance. Surprisingly, however, we still get a negative correlation between fixed and variable trade costs.

Overall, our estimated full lognormal-Melitz model does a very good job in fitting the EDD data and in solving the puzzles associated with the Pareto model. The lognormal model generates an IME that is close to the one we see in the EDD and implies fixed trade costs that are positively correlated with variable trade costs and distance. The implied pattern for the IME across different percentiles is also very similar to what we see in the data.

\(^{29}\)The confidence interval in Table (3.13) comes 1000 random realizations of the parameters in our Markov chains.
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

Counterfactual Analysis

In this section we study whether the counterfactual implications of the Melitz-lognormal model estimated as in the previous section differ from those of the Melitz-Pareto model estimated as described below. To conduct counterfactual analysis, we need to close the model. We do so in standard fashion by assuming that labor is the only factor of production, with wage \( w_i \) and perfectly inelastic labor supply \( L_i \) in country \( i \), by assuming that entry costs are in terms of labor, and that fixed exporting costs are in terms of labor of the exporting country. To make the model be perfectly consistent with the data, we allow for trade imbalances via exogenous international transfers, as in Dekle, Eaton, and S. Kortum 2008,

\[
\Delta_i = X_i - Y_i = \sum_l X_{li} - \sum_j X_{ij}.
\]

To derive the equilibrium conditions, it is convenient to introduce the following notation:

\[
\tilde{\varphi} = (\sigma - 1) \ln \varphi + \ln \alpha - (\sigma - 1) \mu_{\varphi,i} - \mu_{\alpha},
\]

\[
\tilde{f} = \ln f - \mu_{f,ij},
\]

\[
A_{ij} = \ln \left( \frac{(\sigma w_i r_{ij})^{1-\sigma} P_j^{\sigma-1} X_j}{\sigma w_i} \right) + (\sigma - 1) \mu_{\varphi,i} + \mu_{\alpha} - \mu_{f,ij}.
\]

Firms from \( i \) export to \( j \) if and only if \( \tilde{f} \leq A_{ij} + \tilde{\varphi} \), so the fraction of exporting firms is

\[
n_{ij} = \frac{N_{ij}}{N_i} = \int_{-\infty}^{+\infty} \int_{-\infty}^{A_{ij} + \tilde{\varphi}} g(\tilde{\varphi}, \tilde{f}) d\tilde{f} d\tilde{\varphi} \equiv h^{-1}(A_{ij}),
\]

where \( g \) is the known joint pdf of \( \tilde{f} \) and \( \tilde{\varphi} \). This can be rewritten as \( A_{ij} = h(n_{ij}) \), or

\[
h(n_{ij}) = \ln \left( \frac{(\sigma w_i r_{ij})^{1-\sigma} P_j^{\sigma-1} X_j}{\sigma w_i} \right) + (\sigma - 1) \mu_{\varphi,i} + \mu_{\alpha} - \mu_{f,ij}.
\] (3.38)

In turn, the price index is

\[
P_{1-\sigma}^j = \sum_k P_{1-\sigma}^{k,j},
\] (3.39)

with

\[
P_{1-\sigma}^{k,j} = N_i (\sigma w_i r_{ij})^{1-\sigma} e^{(\sigma-1)\mu_{\varphi,i} + \mu_{\alpha}} \int_{-\infty}^{\infty} e^{\tilde{\varphi}} \int_{-\infty}^{h(n_{ij}) + \tilde{\varphi}} g(\tilde{\varphi}, \tilde{f}) d\tilde{f} d\tilde{\varphi}.
\] (3.40)

Free entry implies that profits net of fixed costs of exporting are equal to entry costs. Profits gross of fixed costs are equal to \( \frac{1}{\sigma} \sum_j N_{ij} = \frac{1}{\sigma} \sum_j \lambda_{ij} X_j \) and total fixed costs of exporting per destination are

\[
N_i w_i e^{\mu_{f,ij}} \int_{-\infty}^{\infty} \int_{-\infty}^{h(n_{ij}) + \tilde{\varphi}} e^{\tilde{\varphi}} g(\tilde{\varphi}, \tilde{f}) d\tilde{f} d\tilde{\varphi},
\]
hence the free entry condition is

\[ F^e w_i N_i = \sum_j \left( \frac{1}{\sigma} \lambda_{ij} X_j - N_i w_i e^{\mu_{f,ij}} \int_{-\infty}^{\infty} \int_{-\infty}^{h(n_{ij})+\hat{\varphi}} e^{\tilde{f}} g(\tilde{\varphi}, \tilde{f}) d\tilde{f} d\tilde{\varphi} \right). \]

But using Equation 3.38 together with

\[ \lambda_{ij} = \frac{P_{ij}^{1-\sigma}}{P_j^{1-\sigma}} \] (3.41)

we get

\[ e^{\mu_{f,ij}} = \frac{\lambda_{ij} X_j}{e^{h(n_{ij})} N_i \sigma w_i \int_{-\infty}^{\infty} e^{\tilde{\varphi}} \int_{-\infty}^{h(n_{ij})+\hat{\varphi}} g(\tilde{\varphi}, \tilde{f}) d\tilde{f} d\tilde{\varphi}}, \]

and so we can rewrite the free entry condition as

\[ \sigma F^e w_i N_i = \sum_j \lambda_{ij} X_j \left( 1 - \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{h(n_{ij})+\hat{\varphi}} e^{\tilde{f}} g(\tilde{\varphi}, \tilde{f}) d\tilde{f} d\tilde{\varphi}}{e^{h(n_{ij})} \int_{-\infty}^{\infty} e^{\tilde{\varphi}} \int_{-\infty}^{h(n_{ij})+\hat{\varphi}} g(\tilde{\varphi}, \tilde{f}) d\tilde{f} d\tilde{\varphi}} \right). \] (3.42)

An equilibrium is defined as variables \( \{n_{ij}, \lambda_{ij}, P_{ij}\} \) and \( \{X_j, P_j, w_i\} \) such that Equations 3.38 - 3.42 are satisfied for all \( i, j \), and in addition

\[ w_i L_i = \sum_j \lambda_{ij} X_j, \] (3.43)

and

\[ X_j = w_j L_j + \Delta_j \] (3.44)

are satisfied for all \( i \).

We are interested in the effect of changes in trade costs and trade imbalances. To use the exact hat algebra approach to counterfactual analysis in Dekle, Eaton, and S. Kortum 2008, we first transform this system of equations in levels into one in hat changes, with standard hat notation \( \hat{x} = x'/x \), where we use primes to denote counterfactual values. We take \( \hat{\tau}_{ij} \) and \( \hat{\Delta}_{ij} \) as exogenous and solve for changes in endogenous variables. Since \( n_{ij} \) always appears as an argument in the known \( h() \) function, it is convenient to focus instead on \( h_{ij} \equiv h(n_{ij}). \)
We proceed in three steps: first, we consider a change in trade costs and compute the
n in the opposite holds in the Melitz-lognormal, where the trade elasticity is a function of
of the trade elasticity. This differs dramatically across the two models: while the trade
Rodríguez-Clare (2012) and, counterfactual implications depend critically on the behavior
model with those of the Melitz-Pareto model. As is well known from Arkolakis, Costinot, and

The system of equations in hat changes is then:

\[ h_{ij}(\hat{h}_{ij} - 1) = \ln \left( \frac{(\hat{w}_i\hat{\tau}_{ij})^{1-\sigma} \hat{P}_j^{1-\sigma} \hat{X}_j}{\hat{w}_i} \right) \]  \hfill (3.45)

\[ \hat{P}_j^{1-\sigma} = \sum_k \lambda_{kj} \hat{P}_k^{1-\sigma} \]  \hfill (3.46)

\[ \hat{P}_{ij}^{1-\sigma} = \hat{N}_i(\hat{w}_i\hat{\tau}_{ij})^{1-\sigma} \int_{-\infty}^{\infty} e^{\hat{\phi}} \int_{-\infty}^{\hat{h}_{ij}+\hat{\phi}} g(\hat{\phi}, \hat{f}) d\hat{f} d\hat{\phi} \]  \hfill (3.47)

\[ \hat{\lambda}_{ij} = \frac{\hat{P}_{ij}^{1-\sigma}}{\hat{P}_j^{1-\sigma}} \]  \hfill (3.48)

\[ \hat{w}_i\hat{N}_i \sum_j \lambda_{ij} X_j \left( 1 - \frac{\int_{-\infty}^{\hat{h}_{ij}+\hat{\phi}} e^{\hat{f}} g(\hat{\phi}, \hat{f}) d\hat{f} d\hat{\phi}}{e^{\hat{h}_{ij}} \int_{-\infty}^{\infty} e^{\hat{\phi}} \int_{-\infty}^{h(n_{ij})+\hat{\phi}} g(\hat{\phi}, \hat{f}) d\hat{f} d\hat{\phi}} \right) \]

\[ = \sum_j \lambda_{ij} X_j \hat{\lambda}_{ij} \hat{X}_j \left( 1 - \frac{\int_{-\infty}^{\hat{h}_{ij}+\hat{\phi}} e^{\hat{f}} g(\hat{\phi}, \hat{f}) d\hat{f} d\hat{\phi}}{e^{\hat{h}_{ij}} \int_{-\infty}^{\infty} e^{\hat{\phi}} \int_{-\infty}^{h(n_{ij})+\hat{\phi}} g(\hat{\phi}, \hat{f}) d\hat{f} d\hat{\phi}} \right) \]  \hfill (3.49)

\[ \hat{w}_i Y_i = \sum_j \lambda_{ij} X_j \hat{\lambda}_{ij} \hat{X}_j \]  \hfill (3.50)

\[ \hat{X}_j X_j = \hat{w}_j Y_j + \hat{\Delta}_j \Delta_j \]  \hfill (3.51)

Here \( \lambda_{ij}, h_{ij}, X_j, Y_i \) and \( \Delta_i \) are data, \( \hat{\tau}_{ij} \) and \( \hat{\Delta}_j \) reflect exogenous shocks, and \( \hat{h}_{ij}, \hat{\lambda}_{ij}, \hat{w}_i, \hat{N}_i, \hat{P}_j, \hat{P}_{ij} \) and \( \hat{X}_j \) are the endogenous variables that we solve for. It is important to note that, in contrast to the hat-algebra in the Melitz-Pareto model, with a lognormal distribution we also need data on \( n_{ij} \) so that we can get \( h_{ij} \).

We consider the 12 Latin American countries for which we have estimated the full lognormal model. Although we have data for \( N_{ij} \) and \( X_{ij} \) for all countries in the extended EDD, we actually need \( n_{ij} = N_{ij}/N_i \), and this requires an estimate of the total number of draws, \( N_i \), which we only have for the smaller sample of 37 countries XXX. For the counterfactual analysis we consider a world composed only of the countries in this sample, which means that all other international trade flows are implicitly assigned to domestic transactions. We also need the domestic counterpart of \( X_{ij} \) and \( N_{ij} \). Finally, we also need values of \{ \( \sigma_{\phi,i}, \sigma_{\alpha,i}, \sigma_{f,i}, \rho_i, \sigma \} \) — we use the ones estimated in the previous subsection.

We are interested in comparing the counterfactual implications of the Melitz-lognormal model with those of the Melitz-Pareto model. As is well known from Arkolakis, Costinot, and Rodríguez-Clare (2012) and, counterfactual implications depend critically on the behavior of the trade elasticity. This differs dramatically across the two models: while the trade elasticity in the Melitz-Pareto model is common across country pairs and invariant to shocks, the opposite holds in the Melitz-lognormal, where the trade elasticity is a function of \( n_{ij} \). We proceed in three steps: first, we consider a change in trade costs and compute the
counterfactual implications using the estimated Melitz-lognormal model. We consider four different variations in trade costs shocks – a 1%, a 10%, a 30%, and a 50% uniform reduction in international trade costs: \( \hat{\tau}_{ij} = \hat{\tau} \in \{0.99, 0.9, 0.75, 0.6\} \) if \( i \neq j \), while \( \hat{\tau}_{ii} = 1 \). Second, we use these results to estimate the trade elasticity in the Melitz-Pareto model by running the OLS regression

\[
\Delta \ln X_{ij} = \gamma_i^o + \gamma_j^d + \theta \Delta \ln \tau_{ij} + \zeta_{ij}.
\]

This leads to four values of the implied trade elasticity \( \hat{\theta} \) for the Pareto model. Finally, for each trade cost shock we compute the counterfactual implications in the Melitz-Pareto model with each of the four estimated trade elasticities.

We show the results of this exercise in Figures 3.16 and 3.17. We use \( \hat{W}_i^m \equiv \hat{w}_i^m/\hat{P}_i^m \) and \( \hat{X}_{ij}^m \) to denote the hat changes in welfare and trade flows for the lognormal model \( (m = LN) \) and the Pareto model \( (m = P) \). Figure 3.16 plots \( \hat{W}_i^{LN} \) (horizontal axis) versus \( \hat{W}_i^P \) (vertical axis) in response to four different trade costs shocks and for four different ways of estimating the trade elasticity in the Pareto model. It is evident that both models yield very similar results. This reveals that, for the purposes of computing the welfare effects of a trade shock, the variation in the trade elasticity across country pairs and ... the Pareto model yields a good approximation for the gains from trade liberalization as each origin has trading partners with both lower and higher trade elasticities. In addition, we estimate “average” trade elasticities to be very close across different levels of trade cost shocks, with \( \hat{\theta}_{0.99} = 4.53 \) and \( \hat{\theta}_{0.6} = 4.38 \), thus gains are not sensitive to the particular values of trade elasticity in the Pareto model.

Even though gains from trade liberalization are very close across the two models, the Pareto and the lognormal models differ in their implications for the counterfactual changes in bilateral trade flows. Figure 3.17 plots the ratio between the counterfactual changes in trade flows in the lognormal model, \( \hat{X}_{ij}^{LN} - 1 \), and in the Pareto model, \( \hat{X}_{ij}^P - 1 \), against the initial export share \( r_{ij} \equiv X_{ij}/Y_i \). Notice that the countries we use for this exercise don’t trade much with each other, since we exclude some of their biggest export markets, such as the US or Brazil. We can see that the biggest difference between the Pareto and the lognormal models occurs for the bilateral pairs that don’t trade much – those are exactly the country pairs for which trade elasticities between the two models differ most.

### 3.5 Conclusion

The canonical Melitz model of trade with Pareto-distributed firm productivities has a stark prediction: conditional on the level of the fixed costs of exporting, all variation in exports across partners should be due to the number of exporting firms (the extensive margin). There should be no variation in the intensive margin (exports per exporting firm), again conditional on fixed costs.

We use the World Bank’s Exporter Dynamics Database plus China to test this prediction. Compared to existing studies, this data allows one to look for systematic variation
in the intensive and extensive margins of trade, allowing for year, origin, and destination components of fixed trading costs.

We find that at least 40% of the variation in exports occurs along the intensive margin. That is, when exports from a given origin to a given destination are high, exports per exporting firm are responsible for at least 40% of the high exports. This finding is robust to looking at all destinations or only the largest destinations, including all firms or ignoring very small firms, including all country pairs or only ones for which more than 100 firms export, and disaggregating across industries. When we look at average exports by percentile of exporting firms (rather than the average), we find the intensive margin is more important the higher the percentile.

Although variation in fixed trade costs across country pairs can make the Melitz-Pareto model fit the intensive margin in the data, such fixed trade costs would need to be negatively correlated with variable trade costs and with distance. Moreover, variation in fixed trade costs does not reproduce the pattern of a steadily rising intensive margin by exporter percentiles. Allowing firms to export multiple products or taking into account granularity (a finite number of exporting firms) does not reverse these implications.

In contrast, moving away from a Pareto distribution and assuming that the productivity distribution is lognormal resolves the puzzles. Specifically, we estimate a Melitz model with lognormally distributed firm productivity and idiosyncratic firm-destination demand shifters and fixed costs. We estimate this model using maximum likelihood methods on the EDD firm-level data. The estimated Melitz-lognormal model is consistent with the positive intensive margin overall and with the intensive margin rising by exporting firm percentile. This specification also implies fixed trade costs that are positively correlated with variable trade costs and distance.

Whether the underlying distribution of firm productivity is Pareto or lognormal may matter for the gains from trade. First, Head, Mayer, and Thoenig (2014) argue that the static gains from trade are typically larger in a calibrated Melitz-lognormal model than under Melitz-Pareto (e.g., the equivalent of 5% vs. 2% of GDP). Second, recent models of the dynamic gains from trade have emphasized how domestic firms can learn from firms selling or producing in the domestic market. In Alvarez, Buera, and Lucas (2014), Buera and Oberfield (2015), and Perla, Tonetti, and Waugh (2015), trade liberalization boosts the level or growth rate of technology through this channel. The size of this dynamic gain should depend on whether the distribution of firm productivity is Pareto vs. lognormal, as it interacts with how trade alters the distribution of producer and seller productivity. For example, trade liberalization induces more entry of marginal exporters under Pareto than under lognormal, as seen by no change in exports per exporter under Pareto (zero intensive margin elasticity, unit extensive margin elasticity) vs. a sizable intensive margin and weaker extensive margin under lognormal.
3.6 Figures

Figure 3.1: Intensive and Extensive margins of exporting

Panel A: Average size of exporters (intensive margin) and total exports

Panel B: Number of exporters (extensive margin) and total exports

Note: the source are the statistics in the Exporter Dynamics Database for the extended sample. The x-axis represents log total exports at the origin country-destination country-year level demeaned by origin-year, and destination-year fixed effects. Only origin-destination pairs with more than 100 exporting firms considered. The dots represent the raw measures. The line is the slope predicted by the Melitz-Pareto model.
Figure 3.2: Intensive and extensive margins of exporting, by industry

**Panel a: Average size of exporters (intensive margin) and total exports**

**Panel b: Number of exporters (extensive margin) and total exports**

Note: the source are the statistics in the Exporter Dynamics Database for the extended sample. The x-axis represents log total exports at the origin country-HS 2-digit product-destination country-year level demeaned by origin-HS 2-digit-year and destination-HS 2-digit-year fixed effects. Only origin-HS 2-digit-destination triplets with more than 100 exporting firms are considered. The line is the slope predicted by the Melitz-Pareto model.
Figure 3.3: IME for each percentile, data

Note: the source is the exporter-level data used for the Exporter Dynamics Database for the core sample. The x-axis represents percentiles of the average exporter size distribution. Each dot represents the coefficient from the regression of log average exports per firm in an exporter size percentile on log total exports. The data is demeaned by origin-year and destination-year fixed effects.

Figure 3.4: Manufacturing absorption and averaged exports per firm (destination fixed effects)

Note: the source is the exporter-level data used for the Exporter Dynamics Database for the core sample. The x-axis represents the log of manufacturing absorption in each destination country measured by manufacturing gross production plus manufacturing imports minus manufacturing exports (measured in billions of USD). The y-axis represents the estimated destination fixed effects obtained from a regression of log average exports per firm on origin, destination, and year fixed effects based on the core sample considering origin-destination pairs with more than 100 exporting firms. Manufacturing gross production is calculated as manufacturing value-added from the World Development Indicators divided by 0.418 (the factor used by EKK). Manufacturing imports and exports are obtained from COMTRADE/WITS.
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

Figure 3.5: Exports to largest destination and market entry

Note: the source is the exporter-level data used for the Exporter Dynamics Database for the core sample. The x-axis represents for each country $i$ the log of the ratio of average exports per exporter to destination $j$ to average exports per exporter to $i$'s most popular destination market. The y-axis represents for each country $i$ the log of the ratio of the number of exporters to destination $j$ to the number of exporters to $i$'s most popular destination market. For the calculation of both average exports per exporter and number of exporters we focus only on firms from $i$ that sell both in $j$ and in the most popular destination.
Figure 3.6: Model-implied fixed and variable trade costs and distance

Panel a: fixed trade costs and distance

Panel b: variable trade costs and distance

Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance demeaned by origin and destination fixed effects taken from Mayer and Zignago (2011). The y-axis represents model-implied fixed or variable trade costs demeaned by origin-year and destination-year fixed effects. To calculate the model-implied fixed and variable trade costs we use $\theta = 5$ from Head and Mayer (2014) and $\sigma = 5$ from Bas, Mayer, and Thoenig (2015).\[\]
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

Figure 3.7: Fixed product-level trade costs and distance

Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance taken from Mayer and Zignago (2011). The y-axis represents model-implied fixed product-level trade costs demeaned by origin-year and destination-year fixed effects.
Figure 3.8: IME for each percentile, Pareto and granularity

Note: the source is the exporter-level data used for the Exporter Dynamics Database. The darker solid line corresponds to IME for each percentile estimated using EDD and four main destinations: France, Germany, Japan and the U.S. Dashed lines indicate 95% confidence intervals. The lighter solid line is IME for each percentile implied by the model with Pareto distribution of productivity and granularity, $\tilde{\theta} = 1$. The level of bilateral fixed trade costs was chosen to match overall IME in the data. The number of draws for each origin-destination pair is equal to the number of exporters from origin to destination in EDD as of 2009.
Figure 3.9: Number of firms and population

Note: the x-axis represents log of population taken from the World Development Indicators. The y-axis represents the number of firms as computed by Bento and Restuccia (2015).
Figure 3.10: IME for each percentile, lognormal

Note: the source is the exporter-level data used for the Exporter Dynamics Database. The darker solid line corresponds to IME for each percentile estimated using EDD and four main destinations: France, Germany, Japan and the U.S. Dashed lines indicate 95% confidence intervals. The lighter solid line is IME for each percentile implied by the model with lognormal distribution of productivity, $\bar{\sigma}_\varphi = 4.02$ (our estimate) and $\sigma = 5$ from Bas, Mayer, and Thoenig (2015). The level of bilateral fixed trade costs was chosen to match overall IME in the data. The total number of firms was imputed from Bento and Restuccia (2015).
Figure 3.11: Fixed and variable trade costs and distance, lognormal

*Panel a: fixed trade costs and distance*

*Panel b: variable trade costs and distance*

Note: source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance taken from Mayer and Zignago (2011). Only four destination countries are considered: France, Germany, Japan, and the U.S. To calculate the model-implied fixed and variable trade costs we use our estimate of $\sigma_{\phi} = 4.02$ and $\sigma = 5$ from Bas, Mayer, and Thoenig (2015), and implied number of firm from Bento and Restuccia (2015).
Figure 3.12: Full lognormal model, pdf of log sales

Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors’ calculations. The darker solid line corresponds to empirical CDF of log sales from some origin to the US. The lighter solid line corresponds to CDF implied by estimated full lognormal model.

Figure 3.13: Share of firms selling to destination X but not to destination Y

Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors’ calculations. Each point corresponds to the share of firms exporting only to less popular markets in the data (horizontal axis) and according to the estimated model (vertical axis). The figure also inputs a 45° line.
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

Figure 3.14: Correlation between log exports to top destinations

Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors’ calculations. Each point corresponds to the correlation between log exports to the top three destinations (for those firms who sell in both markets) in the data (horizontal axis) and according to the estimated model (vertical axis).

Figure 3.15: IME for each percentile, data and full lognormal model

Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors’ calculations. The x-axis represent percentiles. The dark solid line represents coefficient from the regression of log average exports in each percentile on log total exports in the data.
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

Figure 3.16: Gains from trade liberalization
Figure 3.17: Counterfactual changes in trade flows
Figure 3.18: Counterfactual changes in trade flows, $\lambda_{ij}^c > 1\%$
3.7 Tables

Table 3.1: Core Sample of EDD countries+China, years firm-level data is available

<table>
<thead>
<tr>
<th>ISO3</th>
<th>Country name</th>
<th>1st year</th>
<th>Last year</th>
<th>ISO3</th>
<th>Country name</th>
<th>1st year</th>
<th>Last year</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALB</td>
<td>Albania</td>
<td>2004</td>
<td>2012</td>
<td>KHM</td>
<td>Cambodia</td>
<td>2003</td>
<td>2009</td>
</tr>
<tr>
<td>BFA</td>
<td>Burkina Faso</td>
<td>2005</td>
<td>2012</td>
<td>LAO</td>
<td>Laos</td>
<td>2006</td>
<td>2010</td>
</tr>
<tr>
<td>BGD</td>
<td>Bangladesh</td>
<td>2005</td>
<td>2013</td>
<td>LBN</td>
<td>Lebanon</td>
<td>2008</td>
<td>2012</td>
</tr>
<tr>
<td>CHL</td>
<td>Chile</td>
<td>2003</td>
<td>2012</td>
<td>MKD</td>
<td>Macedonia</td>
<td>2003</td>
<td>2010</td>
</tr>
<tr>
<td>CHN</td>
<td>China</td>
<td>2003</td>
<td>2008</td>
<td>MMR</td>
<td>Myanmar</td>
<td>2011</td>
<td>2013</td>
</tr>
<tr>
<td>CIV</td>
<td>Cote d’Ivoire</td>
<td>2009</td>
<td>2012</td>
<td>MUS</td>
<td>Mauritius</td>
<td>2003</td>
<td>2012</td>
</tr>
<tr>
<td>CMR</td>
<td>Cameroon</td>
<td>2003</td>
<td>2013</td>
<td>MWI</td>
<td>Malawi</td>
<td>2009</td>
<td>2012</td>
</tr>
<tr>
<td>COL</td>
<td>Colombia</td>
<td>2007</td>
<td>2013</td>
<td>NIC</td>
<td>Nicaragua</td>
<td>2003</td>
<td>2013</td>
</tr>
<tr>
<td>DOM</td>
<td>Dominican Republic</td>
<td>2003</td>
<td>2013</td>
<td>PAK</td>
<td>Pakistan</td>
<td>2003</td>
<td>2010</td>
</tr>
<tr>
<td>EGY</td>
<td>Egypt</td>
<td>2006</td>
<td>2012</td>
<td>PER</td>
<td>Peru</td>
<td>2003</td>
<td>2013</td>
</tr>
<tr>
<td>ETHI</td>
<td>Ethiopia</td>
<td>2008</td>
<td>2012</td>
<td>QOS</td>
<td>Kosovo</td>
<td>2011</td>
<td>2013</td>
</tr>
<tr>
<td>GIN</td>
<td>Guinea</td>
<td>2009</td>
<td>2012</td>
<td>THA</td>
<td>Thailand</td>
<td>2012</td>
<td>2013</td>
</tr>
<tr>
<td>HRV</td>
<td>Croatia</td>
<td>2007</td>
<td>2012</td>
<td>UGA</td>
<td>Uganda*</td>
<td>2003</td>
<td>2010</td>
</tr>
<tr>
<td>IRN</td>
<td>Iran</td>
<td>2006</td>
<td>2010</td>
<td>URY</td>
<td>Uruguay</td>
<td>2003</td>
<td>2012</td>
</tr>
</tbody>
</table>

* indicates that Uganda does not have data for 2006
Table 3.2: IME regressions, core sample

<table>
<thead>
<tr>
<th>Panel a: country pairs with $N_{ij} \geq 100$</th>
<th>Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity</td>
<td>0.438*** 0.459*** 0.400***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0058] [0.0041] [0.0049]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.55 0.74 0.85</td>
</tr>
<tr>
<td>Variation in $\ln X_{ij}$ explained by FE,%</td>
<td>0.01 0.20 0.59</td>
</tr>
<tr>
<td>Observations</td>
<td>7,781 7,768 7,324</td>
</tr>
</tbody>
</table>

| Panel b: all country pairs                 |                                             |
| IM elasticity                              | 0.503*** 0.530*** 0.579***                  |
| Standard error                             | [0.0018] [0.0017] [0.0022]                  |
| $R^2$                                       | 0.77 0.81 0.85                              |
| Variation in $\ln X_{ij}$ explained by FE, % | 0.00 0.20 0.50                             |
| Observations                               | 47,129 47,129 47,037                        |

Year FE: Yes
Origin $\times$ year FE: Yes, Yes
Destination $\times$ year FE: Yes

Note: robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table 3.3: IME regression, small firms excluded, core sample

<table>
<thead>
<tr>
<th></th>
<th>Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel a: country pairs with $N_{ij} \geq 100$</strong></td>
<td></td>
</tr>
<tr>
<td>IM elasticity</td>
<td>0.437*** 0.459*** 0.398***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0058] [0.0042] [0.0050]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.54 0.74 0.85</td>
</tr>
<tr>
<td>Variation in $\ln X_{ij}$ explained by FE, %</td>
<td>0.01 0.19 0.59</td>
</tr>
<tr>
<td>Observations</td>
<td>7,698 7,684 7,234</td>
</tr>
<tr>
<td><strong>Panel b: all country pairs</strong></td>
<td></td>
</tr>
<tr>
<td>IM elasticity</td>
<td>0.497*** 0.525*** 0.573***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0013] [0.0013] [0.0015]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.77 0.81 0.84</td>
</tr>
<tr>
<td>Variation in $\ln X_{ij}$ explained by FE, %</td>
<td>0.00 0.19 0.50</td>
</tr>
<tr>
<td>Observations</td>
<td>46,925 46,925 46,832</td>
</tr>
</tbody>
</table>

Year FE         | Yes
Origin $\times$ year FE | Yes | Yes
Destination $\times$ year FE | Yes

Note: firms with annual exports lower than $1000$ excluded
Robust standard errors in brackets
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table 3.4: IME regression, disaggregated within manufacturing, core sample

<table>
<thead>
<tr>
<th>Panel a: HS 2-digit</th>
<th>Coefficient from ln $x_{ij}$ on ln $X_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity</td>
<td>0.569*** 0.510*** 0.467***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0022] [0.0017] [0.0049]</td>
</tr>
<tr>
<td>Observations</td>
<td>37,321 35,621 10,732</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel b: HS 4-digit</th>
<th>Coefficient from ln $x_{ij}$ on ln $X_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity</td>
<td>0.651*** 0.569*** 0.515***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0019] [0.0013] [0.0069]</td>
</tr>
<tr>
<td>Observations</td>
<td>62,776 58,516 4,640</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel c: HS 6-digit</th>
<th>Coefficient from ln $x_{ij}$ on ln $X_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity</td>
<td>0.664*** 0.593*** 0.508***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0020] [0.0014] [0.0094]</td>
</tr>
<tr>
<td>Observations</td>
<td>67,967 61,501 2,972</td>
</tr>
</tbody>
</table>

Note: $N_{ij} > 100$
Robust standard errors in brackets
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Product-level IME regression, core sample

<table>
<thead>
<tr>
<th>Panel a: HS 2-digit</th>
<th>Coefficient from ln $x_{ij}$ on ln $X_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity</td>
<td>0.380*** 0.397*** 0.288***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0070] [0.0054] [0.0073]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.35 0.62 0.78</td>
</tr>
<tr>
<td>Variation in ln $X_{ij}$ explained by FE,%</td>
<td>0.01 0.20 0.59</td>
</tr>
<tr>
<td>Observations</td>
<td>7781 7,768 7,324</td>
</tr>
</tbody>
</table>

| Year × HS FE | Yes |
| Origin × Year × HS FE | Yes | Yes |
| Destination × Year × HS FE | Yes |

Note: $N_{ij} > 100$
Robust standard errors in brackets
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
### CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

#### Table 3.6: Margins of trade and distance

<table>
<thead>
<tr>
<th>Elasticity with respect to distance</th>
<th>( x_{ij} )</th>
<th>( N_{ij} )</th>
<th>( x_{ij}^p )</th>
<th>( m_{ij} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard error ( [0.0150] )</td>
<td>( [0.0130] )</td>
<td>( [0.0128] )</td>
<td>( [0.0158] )</td>
<td>( [0.0146] )</td>
</tr>
<tr>
<td>( 0.123^{***} )</td>
<td>(-0.416^{***} )</td>
<td>(0.288^{***} )</td>
<td>(-0.165^{***} )</td>
<td></td>
</tr>
<tr>
<td>( -0.280^{***} )</td>
<td>(-1.010^{***} )</td>
<td>(-0.071^{***} )</td>
<td>(-0.209^{***} )</td>
<td></td>
</tr>
</tbody>
</table>

| Observations                      | 7,725          | 7,320       |

<table>
<thead>
<tr>
<th>Origin ( \times ) year FE</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Destination ( \times ) year FE</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

Note: \( N_{ij} > 100 \)

#### Table 3.7: Trade costs and distance

<table>
<thead>
<tr>
<th>( \ln \tilde{F}_{ij} )</th>
<th>( \ln \tilde{\tau}_{ij} )</th>
<th>( \ln \tilde{f}_{ij} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln \text{dist}_{ij} )</td>
<td>(-0.280^{***} )</td>
<td>(0.272^{***} )</td>
</tr>
<tr>
<td>Standard error ( [0.0140] )</td>
<td>( [0.0046] )</td>
<td>( [0.0146] )</td>
</tr>
<tr>
<td>Observations</td>
<td>7,320</td>
<td>7,320</td>
</tr>
</tbody>
</table>

Note: \( N_{ij} > 100 \)
Robust standard errors in brackets
* \( p < 0.1 \), ** \( p < 0.05 \), *** \( p < 0.01 \)

#### Table 3.8: Fixed trade costs distance elasticity and granularity

<table>
<thead>
<tr>
<th>Fixed trade costs elasticity</th>
<th>Firm level</th>
<th>Product level</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \zeta )</td>
<td>(-0.022^{***} )</td>
<td>(-0.007^{***} )</td>
</tr>
<tr>
<td>Standard error ( [0.0029] )</td>
<td>( [0.0026] )</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>7,320</td>
<td>7,320</td>
</tr>
</tbody>
</table>

Note: \( N_{ij} > 100 \)
Robust standard errors in brackets
* \( p < 0.1 \), ** \( p < 0.05 \), *** \( p < 0.01 \)
Product-level IME is calculated for the core sample
Table 3.9: IME under granularity

<table>
<thead>
<tr>
<th>$\bar{\theta}$</th>
<th>corr($\alpha_j \varphi, \alpha_k \varphi$)</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3</td>
<td>0.005</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>1.75</td>
<td>0.020</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>1.25</td>
<td>0.133</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.333</td>
<td>0.103</td>
<td></td>
</tr>
</tbody>
</table>

Note: $N_{ij} > 100$

$N_{ij}$ data as of 2007, core sample, 867 obs.

Table 3.10: Number of firms and population

<table>
<thead>
<tr>
<th>log number of firms</th>
<th>log population</th>
<th>0.945***</th>
<th>0.944***</th>
</tr>
</thead>
<tbody>
<tr>
<td>log population</td>
<td>Standard error</td>
<td>[0.0136]</td>
<td>[0.0139]</td>
</tr>
<tr>
<td>Observations</td>
<td>468</td>
<td>468</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year FE</th>
<th>Yes</th>
</tr>
</thead>
</table>

Robust standard errors in brackets
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.11: Trade costs and distance, lognormal

<table>
<thead>
<tr>
<th>log fixed costs</th>
<th>log variable costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln dist</td>
<td>0.156***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0155]</td>
</tr>
<tr>
<td>Observations</td>
<td>7738</td>
</tr>
</tbody>
</table>

$N_{ij} > 100$

Robust standard errors in brackets
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
CHAPTER 3. THE INTENSIVE MARGIN IN TRADE

Table 3.12: Estimates of dispersion, full lognormal model

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>median</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\sigma}_e$</td>
<td>3.37</td>
<td>3.19</td>
<td>0.93</td>
<td>5.82</td>
</tr>
<tr>
<td>$\sigma_\alpha$</td>
<td>2.73</td>
<td>2.68</td>
<td>1.94</td>
<td>3.64</td>
</tr>
<tr>
<td>$\sigma_f$</td>
<td>2.39</td>
<td>2.43</td>
<td>1.64</td>
<td>3.11</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.46</td>
<td>0.49</td>
<td>-0.33</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 3.13: Implied IME in full lognormal model

<table>
<thead>
<tr>
<th>IME</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>0.67 [0.61, 0.73]</td>
</tr>
<tr>
<td>Full lognormal model</td>
<td>0.64 [0.59, 0.69]</td>
</tr>
</tbody>
</table>

Table 3.14: Implied trade costs in full lognormal model

<table>
<thead>
<tr>
<th>Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr $\left( \tilde{F}<em>{ij}, \tilde{\tau}</em>{ij} \right)$</td>
<td>-0.31 [-0.16, -0.45]</td>
</tr>
</tbody>
</table>

Distance elasticity

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed costs</td>
<td>0.36</td>
<td>[0.24, 0.50]</td>
</tr>
<tr>
<td>Variable costs</td>
<td>0.34</td>
<td>[0.30, 0.37]</td>
</tr>
</tbody>
</table>
Chapter A

Appendices to Chapter 1

A.1 Data description

I construct the following variables at the importer-exporter-HS10 good-year (ietg) level:

- Year: year of transaction.
- Importer: Colombian tax number.
- Quantity: total quantity over all transactions within ietg cell.
- FOB unit values in US dollars (prices): total FOB value over all transactions divided by total quantity within ietg cell.
- HS 10 good: interaction between HS10 good and units of measurement (in some rare cases there are more than 1 units of measurement for the same HS10 good). Using first 6 digits of HS10 code I match goods to the Broad Economic Classification (BEC) codes. I classify goods in BEC categories 111, 121, 21, 22, 31, 322, 42, 53 as intermediate goods. I define manufacturing industries as all industries other than those corresponding to HS 2-digit codes 01 (live animals), 06-10 (agriculture), and 27 (fuels).
- Exporter: identified by name. Since there are no numeric identifiers for the exporters in the data, I have to rely on their names, which are sometimes reported with typos. I use a conservative approach to correcting typos, comparing the names after deleting extra spaces and punctuation marks. The transactions are considered to have the same exporter if corresponding exporters’ names match exactly after this minor correction. This way I don’t capture mistakes in letters or words, but minimize the risk of misclassifying two different exporters as being one exporter.
- Importer industry: ISIC rev 3.1 4-digit industry. Sometimes one importer reports different industries in different transactions. I define importer’s industry as the one that is reported in transactions with the largest total value of imports. Firms that belong to
2-digit industries 15-37 are classified as manufacturing firms; firms in industries 50-52 are classified as wholesale/retail trade firms. In my baseline specifications, I restrict the sample of importers to all industries other than wholesale and retail trade.

For robustness checks I use the following set of controls from Kugler and Verhoogen (2011):

- **RnD and advertising expenditures intensity.** This variable is as a share of RnD and advertising expenditures in total sales based on the U.S. US Federal Trade Commission Line of Business Survey (the measure is unavailable in the Colombian survey of manufacturers). It is constructed for a subset of 4-digit ISIC industries. The measure reflects the scope for vertical differentiation: a higher RnD and advertising intensity reflects a bigger scope for vertical differentiation.

- **Gollop-Monahan input dissimilarity index.** The larger the value of index - the more dispersion there is in input bundles across plants in a certain sector, and so potentially the more differentiated their output is. The measure is available at the level of 4-digit ISIC industries.

- **Rauch index.** Kugler and Verhoogen (2011) take Rauch (1999) "liberal" measure of the classification of industries at the SITC 4-digit level, and assign value 0 if its output is traded on organized exchanges or possess "reference prices" and assign value 1 for all other industries. Then they match SITC industries to the ISIC rev. 2 4-digit industries using concordance from OECD.

I match these variables to the HS6 goods classification. The RnD and advertising expenditures intensity measure and the Gollop-Monahan index are normalized to be mean zero with unit standard deviation across 4-digit industries. Rauch index varies between 0 (more centralized industries) to 1 (less decentralized industries).
## A.2 Appendix Tables

**Table A.1: OLS results, sample from Table 1.4**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln p$</td>
<td>$-0.209^{***}$</td>
<td>$-0.218^{***}$</td>
<td>$-0.238^{***}$</td>
</tr>
<tr>
<td>$\ln q$</td>
<td>[0.00848]</td>
<td>[0.00926]</td>
<td>[0.0120]</td>
</tr>
</tbody>
</table>

Year-Exporter-HS10 FE Yes    Yes    Yes
Year-Importer FE            No     Yes    Yes
Importer-Exporter-HS10 FE   No     No     Yes
# of clusters $i$          6,164  3,209  1,893
# of clusters $e$          7,350  5,960  2,895
Observations               118,589 105,491 70,688

Standard errors clustered at importer and exporter level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***.

**Table A.2: OLS results, sample from Table 1.5**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln p$</td>
<td>$-0.208^{***}$</td>
<td>$-0.215^{***}$</td>
<td>$-0.225^{***}$</td>
</tr>
<tr>
<td>$\ln q$</td>
<td>[0.00839]</td>
<td>[0.00922]</td>
<td>[0.0112]</td>
</tr>
</tbody>
</table>

Year-Exporter-HS10 FE Yes    Yes    Yes
Year-Importer FE            No     Yes    Yes
Importer-Exporter-HS10 FE   No     No     Yes
# of clusters $i$          6,154  3,238  1,869
# of clusters $e$          7,323  5,963  2,870
Observations               118319 105166 70491

Standard errors clustered at importer and exporter level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***.
### Table A.3: Additional OLS results in different subsamples

<table>
<thead>
<tr>
<th></th>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
<th>Column 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln ( p )</td>
<td>-0.219***</td>
<td>-0.224***</td>
<td>-0.202***</td>
<td>-0.246***</td>
<td>-0.229***</td>
<td>-0.236***</td>
</tr>
<tr>
<td>ln ( q )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.013]</td>
<td>[0.007]</td>
<td>[0.008]</td>
<td>[0.017]</td>
<td>[0.011]</td>
<td>[0.010]</td>
</tr>
<tr>
<td>Year-Exporter-HS10 FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year-Importer FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Importer-Exporter-HS10 FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td># of clusters ( i )</td>
<td>1,941</td>
<td>10,772</td>
<td>3,010</td>
<td>2,162</td>
<td>2,432</td>
<td>2,928</td>
</tr>
<tr>
<td># of clusters ( e )</td>
<td>4,251</td>
<td>13,721</td>
<td>5,499</td>
<td>3,906</td>
<td>4,457</td>
<td>5,358</td>
</tr>
<tr>
<td>Observations</td>
<td>99,059</td>
<td>645,891</td>
<td>148,015</td>
<td>80,062</td>
<td>117,922</td>
<td>144,405</td>
</tr>
</tbody>
</table>

Standard errors clustered at importer and exporter level in brackets. Firms in wholesale and retail trade and private persons excluded. Significance at 5%, 1% and 0.1% denoted by *, **, and ***. Different columns represent different samples. Relative to the baseline sample: in column 1 the sample is restricted to the importers in manufacturing industries, in column 2 wholesalers and retailers are added, in column 3 the 1% of OLS outliers is dropped, in column 4 only transactions of at least $5,000 in annual volume are considered, in column 5 the sample is restricted to goods that are intermediate inputs, in column 6 only manufacturing goods are considered. See data section of the Online Appendix for the definition of industries of firms and goods.
Chapter B

Appendices to Chapter 2

B.1 Solving the Problem of Intermediate Good Producers

The optimization problem of intermediate good producer $z$ from country $i$ is given by the following set of equations:

$$\max_{x_{ij}(\mu,z), T_{ij}(\mu,z)} \pi^I_i(z) = \sum_{j \in J_i(z)} \pi^I_{ij}(z),$$  \hspace{1cm} (B.1)

$$\text{s.t. } \pi^I_{ij}(z) = M^F_j \int_{\Omega_{\mu,j}} T_{ij}(\mu,z) dF_{\mu,j}(\mu) - x_{ij}(z) \frac{w_i T_{ij}}{z} - w_i f_{ij} \quad \forall j \in J_i(z), \quad (B.2)$$

$$x_{ij}(z) = M^F_j \int_{\Omega_{\mu,j}} x_{ij}(\mu,z) dF_{\mu,j}(\mu) \quad \forall j \in J_i(z), \quad (B.3)$$

$$b_j(x(\mu,z), T(\mu,z); \mu) \geq b_j(x(\mu',z), T(\mu',z); \mu) \quad \forall \mu, \mu', \quad (B.4)$$

$$b_j(x(\mu,z), T(\mu,z); \mu) \geq 0 \quad \forall \mu, \mu', \quad (B.5)$$

where:

$$J_i(z) = \{ j : \pi^I_{ij}(z) \geq 0 \}, \quad (B.6)$$

$$b_j(x, T; \mu_j(\varphi)) = x^\gamma \mu_j(\varphi) - T, \quad (B.7)$$

$$\gamma = \frac{\epsilon - 1}{\epsilon}, \quad (B.8)$$

$$\mu_j(\varphi) = \frac{\sigma - 1}{\sigma} \frac{\epsilon}{\epsilon - 1} \varphi \frac{\sigma - 1}{\sigma} x^I_j(\varphi)^{\frac{1}{\sigma} - \frac{1}{2}} P_j Q_j^{1/\sigma}, \quad (B.9)$$

$$x^I_j(\varphi) = \left( \sum_i M^I_i \int_{\Omega_{z,ij}} x_{ij}(\varphi, z)^{\frac{1}{\sigma} - \frac{1}{2}} dF_{z,i}(z) \right)^{\frac{1}{\frac{1}{\sigma} - \frac{1}{2}}}. \quad (B.10)$$

Assume that $\varphi$ in country $j$ is distributed Pareto with shape parameter $\theta_\varphi$, such that $\theta_\varphi > (\sigma - 1), \theta_\varphi > 1$, and scale parameter $b_{\varphi,j}$. I conjecture that $\mu$ is also distributed Pareto
with shape parameter $\theta_\mu > 1$. I solve the problem of intermediate good producers under this assumptions and later verify that this is indeed a correct conjecture. Since marginal costs of production do not depend on the volume of production, sales in one market do not affect profits in the others. Thus I solve the problem of intermediate good producer market by market, and drop subscripts $i$ and $j$.

Since $\gamma > 0$ the $b(x, T; \mu)$ function satisfies Spence-Mirrlees condition in $x, \mu$: $$\frac{\partial^2}{\partial x \partial \mu} b(x, T; \mu) > 0.$$ Using results from Varian (1989) and Stole (2007) I can now construct the optimal quantity-payment schedule as a function of $\mu$ that are implicitly given by:

$$0 = \frac{\partial b(\mu, x^*; T)}{\partial \mu} - \left( \frac{w_\tau}{z} + \frac{1 - F_\mu(\mu) \partial^2 b(\mu, x^*; T)}{F'_\mu(\mu)} \right),$$  \hspace{1cm} (B.11)

$$T^*(\mu) = b(\mu, x^*(\mu); T) - \int_{\mu_0}^{\mu} \frac{\partial b(t, x^*(t); t)}{\partial \mu} dt,$$  \hspace{1cm} (B.12)

where $\mu_0$ is the lower bound of final good producers whom intermediate good producer decides serve. Plugging in the functional forms for $b$ function and CDF of $\mu$ yields a closed-form expression for optimal quantity and payment:

$$x^*(\mu, z) = \left( \frac{\gamma z}{\tau w} \frac{\theta_\mu - 1}{\theta_\mu} \right)^{1/\gamma},$$  \hspace{1cm} (B.13)

$$T^*(\mu, z) = \gamma \mu^{1/\gamma} \left( \frac{z \gamma}{\tau w} \frac{\theta_\mu - 1}{\theta_\mu} \right)^{1/\gamma} + \left( \frac{z \gamma}{\tau w} \frac{\theta_\mu - 1}{\theta_\mu} \right)^{1/\gamma} \left[ (1 - \gamma) \mu_0^{1/\gamma} \right],$$  \hspace{1cm} (B.14)

The profits of intermediate good producer $z$ are given by:

$$\pi^I(z; \mu_0) = M^F \left( \frac{\mu_0}{b_\mu} \right)^{-\theta_\mu} \left[ \frac{\tau w}{z} \frac{\theta_\mu - 1}{\theta_\mu} - \frac{\tau w}{z} \right] \mathbb{E}[x^*(\mu)|\mu > \mu_0] + \left( \frac{\gamma z}{\tau w} \frac{\theta_\mu - 1}{\theta_\mu} \right)^{1/\gamma} \left[ (1 - \gamma) \mu_0^{1/\gamma} \right]$$  \hspace{1cm} (B.15)

$$\propto z^{\epsilon - 1} \mu_0^{-\theta_\mu + \frac{1}{1-\gamma}},$$  \hspace{1cm} (B.16)

where in the last line I used the property of Pareto distribution of $\mu$ to calculate expectation of $x^*$ and $\frac{\gamma}{1-\gamma} = \epsilon - 1$. If $\theta_\mu > \frac{1}{1-\gamma}$, then for intermediate good producers it is optimal to pick $\mu_0$ as low as possible, in other words, to sell their goods to all of the active final good producers. Profits are increasing in $z$ and since there are fixed costs of selling from $i$ to $j$ there will be a cutoff productivity of exporting $z_{0,ij}$ (assuming $b_{z,i}$ is low enough such that no market $j$ is served by all of the entrants into intermediate good sector). Plugging the
expression for $x^*(\mu(\varphi))$ into $x^I(\varphi)$, I can find the expression for $\mu_j(\varphi)$:

$$
\mu(\varphi) = \left[ \frac{\sigma - 1}{\sigma} \frac{\epsilon}{\epsilon - 1} \right]^{\sigma/\epsilon} \varphi^{\sigma - 1} \left( \frac{\epsilon - 1}{\epsilon} \frac{\theta_\mu - 1}{\theta_\mu} \right)^{\sigma/\epsilon - 1} \left[ \frac{\theta_\mu}{\theta_\mu - (\epsilon - 1)} \right]^{\sigma/\epsilon - 1}
$$

$$
\times \left[ \sum_i M_i I_i^I \left( \frac{z_{0,ij}}{b_{0,i}} \right)^{-\theta_\varphi} \left( \frac{1}{\tau_i w_i} \right)^{\epsilon - 1} z_{0,ij}^{\epsilon - 1} \right]^{\sigma/\epsilon - 1} \left( P_j Q_j^\varphi \right)^{\sigma/\epsilon}
$$

(B.17)

(B.18)

As a result, $\mu$ is indeed distributed Pareto with shape parameter $\theta_\mu = \frac{\theta_\varphi}{\sigma - 1} \epsilon > \frac{1}{1 - \gamma} = \epsilon$.

Plugging optimal quantity-payment schedules into the profit function of final good producers I can show that their profits are increasing in $\varphi$, and so there will also be a cutoff productivity $\varphi_{0,i}$, such that less productive firms will not be active on market $i$ due to the presence of fixed costs.

**B.2 Gravity**

Using equations (B.13) and (B.79) and the fact that there is no selection into importing it is possible to show that profits of intermediate good producer $z$ selling from $i$ to $j$ are given by:

$$
\pi_{ij}^I(z) = M_j^F \left( \frac{\mu_{0,j}}{b_{0,j}} \right)^{-\theta_\varphi} \mathbb{E}_\mu \left[ T_{ij}(\mu, z) - x_{ij}(\mu, z) \frac{w_i \tau_{ij}}{z} | \mu > \mu_{0,j} \right],
$$

(B.19)

$$
= M_j^F \left( \frac{\mu_{0,j}}{b_{0,j}} \right)^{-\theta_\varphi} \frac{1}{\epsilon} \mu_{0,j}^{\epsilon} \left( \frac{\epsilon - 1}{\epsilon} \frac{z_{0,ij}}{\tau_i w_i} \frac{\theta_\mu - 1}{\theta_\mu} \right)^{\epsilon - 1} \theta_\mu - 1 \frac{\theta_\mu - 1}{\theta_\mu - \epsilon}.
$$

(B.20)

Cutoff $z_{0,ij}$ is defined by zero-cutoff profit condition:

$$
w_i f_{ij}^I = M_j^F \left( \frac{\mu_{0,j}}{b_{0,j}} \right)^{-\theta_\varphi} \frac{1}{\epsilon} \mu_{0,j}^{\epsilon} \left( \frac{\epsilon - 1}{\epsilon} \frac{z_{0,ij}}{\tau_i w_i} \frac{\theta_\mu - 1}{\theta_\mu} \right)^{\epsilon - 1} \theta_\mu - 1 \frac{\theta_\mu - 1}{\theta_\mu - \epsilon}.
$$

(B.21)

For any pair of $i', i$ it follows that:

$$
\frac{z_{0,ij}}{z_{0,i'j}} = \left( \frac{w_i f_{ij}^I}{w_{i'} f_{i'j}^I} \right)^{\frac{1}{\theta_\varphi}} \frac{w_i \tau_{ij}}{w_{i'} \tau_{i'j}}.
$$

(B.22)

Total expenditures on intermediate inputs from $i$ by a final good producer $\varphi$ in $j$ are given by:

$$
M_i I_i \left( \frac{z_{0,ij}}{b_{0,i}} \right)^{-\theta_\varphi} \frac{\theta_\varphi}{\theta_\varphi - (\epsilon - 1) \frac{1}{\epsilon} \left( \frac{\epsilon - 1}{\epsilon} \frac{z_{0,ij}}{w_i \tau_{ij}} \frac{\theta_\mu - 1}{\theta_\mu} \right)^{\epsilon - 1} \left( (\epsilon - 1) \mu(\varphi)^\epsilon + \mu_{0,j}^\epsilon \right),
$$

(B.23)

(B.24)
Aggregating (B.24) over all active final good producers in \( j \), I get total expenditures in \( j \) on intermediate inputs from \( i \), denoted by \( X_{ij} \). Combining this with (B.22) for \( i' = j \) and dividing by \( \sum_i X_{ij} \), I get:

\[
\lambda_{ij} = \frac{X_{ij}}{\sum_i X_{ij}} = \frac{M_i^I (w_i f_{ij})^{1-\frac{\theta_z}{\epsilon - 1}} \left( \frac{1}{b_{\varphi,i}} \right)^{-\theta_z} (w_i \tau_{ij})^{-\theta_z}}{\sum_i M_i^I (w_i f_{ij})^{1-\frac{\theta_z}{\epsilon - 1}} \left( \frac{1}{b_{\varphi,i}} \right)^{-\theta_z} (w_i \tau_{ij})^{-\theta_z}}
\]  

(B.25)

**B.3 Gains from Trade**

Each of the final good producers in \( j \) buys intermediate inputs from all of the intermediate good producers (domestic and foreign) that serve market \( j \). Using notation from the paper, \( q_j(\varphi) \) represents the quantity of the goods produced by the final good producer with productivity \( \varphi \) in country \( j \). Since demand for final goods is CES, to sell \( q_j(\varphi) \) units, producer will charge price \( q(\varphi)^{\frac{1}{\sigma}} P_F^j Q_1^{1/\sigma} \). As a result, the price index for final goods is given by:

\[
(P_F^j)^{1-\sigma} = M_j^F \int_{\varphi_{0,j}}^{\infty} q(\varphi)^{\frac{1}{\sigma}} (PQ_1^{1/\sigma})^{1-\sigma} dF(\varphi).
\]  

(B.26)

Using the definition of \( Q_j \) and the solution to the intermediate good producers’ problem (that yields optimal quantity purchased by each of the final good producers in \( j \)), and plugging those expression in (B.26), I get the following expression:

\[
(P_F^j)^{1-\sigma} = M_j^F \varphi_{0,j}^{-1} \frac{\theta_{\varphi}}{\theta_{\varphi} - (\sigma - 1)} \left( \frac{\varphi_{0,j}}{b_{\varphi,j}} \right)^{-\theta_{\varphi}} \left( \frac{\sigma - 1}{\theta_{\mu}} \right)^{1-1} \left[ \frac{\theta_z}{\theta_z - (\epsilon - 1)} \right]^\frac{1}{\epsilon - 1}
\]

\[
\times \left[ \sum_i M_i^I \left( \frac{z_{0,ij}}{b_{0,i}} \right)^{-\theta_z} \left( \frac{1}{\tau_{ij} w_i} \right)^{\epsilon - 1} z_{0,ij}^{\frac{1}{\epsilon - 1}} \right]^\frac{1}{\epsilon - 1}.
\]  

(B.27)

Since wage in country \( j \) is given by \( w_j \) and there are no aggregate profits (due to free entry), aggregate welfare is given by:

\[
W_j = \frac{L_j w_j}{P_{F,j}^j}.
\]  

(B.28)

Without loss of generality, I normalize \( w_j = 1 \). I can decompose final good producer \( \varphi \)’s production costs into variable part \( V_j(\varphi) \) and fixed part \( F_j(\varphi) \), so that total cost of that producer is given by:

\[
T_j(\varphi) = V_j(\varphi) + F_j(\varphi),
\]  

(B.29)
where:

\[ V_j(\varphi) = \frac{\theta_z}{\theta_z - (\epsilon - 1)} M_i \sum_i \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_z} x_{ij}(\mu(\varphi), z) \frac{\tau_{ij} w_i}{z_{0,ij}} \frac{\theta_{\mu}}{\theta_{\mu} - 1}, \]  

(B.30)

\[ F_j(\varphi) = \frac{\theta_z}{\theta_z - (\epsilon - 1)} M_i \sum_i \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_z} x_{ij}(\mu(\varphi_0), z) \frac{\tau_{ij} w_i}{z_{0,ij}} \frac{\theta_{\mu}}{\theta_{\mu} - 1} \frac{1}{\epsilon - 1}. \]  

(B.31)

Due to the CES demand for final goods, final good producer \( \varphi \) in \( j \) will charge a constant markup over marginal costs. I can show that marginal costs are constant. Denoting revenues of final good producer \( j \) by \( R_j(\varphi) \), we get:

\[ V_j(\varphi) = \frac{\sigma - 1}{\sigma} R_j(\varphi) \]  

(B.32)

Using zero cutoff profit condition \( (R_j(\varphi_0) - V_j(\varphi_0) - F_j(\varphi_0) = f^F_j) \), free entry condition (expected profits net of fixed costs are zero; aggregate entry costs in the final good sector are equal to the aggregate profits net of fixed costs), and the fact that revenues are proportional to \( \varphi^{\sigma - 1} \), I can derive the following conditions for the cutoff productivity and the mass of firms:

\[ \left( \frac{\varphi_{0,j}}{b_{\varphi,j}} \right)^{-\theta_\varphi} = \frac{f^{e,F}_j \theta_\varphi - (\sigma - 1) \epsilon - \sigma}{f^F_j \theta_\varphi - (\sigma - 1) \epsilon - 1}, \]  

(B.33)

\[ M^F_j = \frac{L_j}{f^{e,F}_j \theta_\varphi}. \]  

(B.34)

Hence, \( \varphi_{0,j} \) and \( M^F_j \) are not affected by the trade share.
Using (B.22), note that:

\[
\sum_i M_i^I \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_z} \left( \frac{1}{\tau_{ij}w_i} \right)^{-\theta_0 + (\epsilon - 1)}
\]

(B.35)

\[
= \sum_i M_i^\Omega \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_z} \left( \frac{1}{w_j f_j} \right)^{-\theta_z} \left( \frac{1}{\tau_{ij} w_i} \right)^{-\theta_z + (\epsilon - 1)}
\]

(B.36)

\[
\sum_i M_i^I \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_z} \left( \frac{1}{w_j f_j} \right)^{-\theta_z} \left( \frac{1}{\tau_{ij} w_i} \right)^{-\theta_z + (\epsilon - 1)} \sum_i M_i^I \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_z} \left( \frac{1}{w_i f_i} \right)^{-\theta_z} \left( \frac{1}{\tau_{ij} w_i} \right)^{-\theta_z}
\]

(B.37)

\[
= \sum_i M_i^\Omega \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_z} \left( \frac{1}{w_j f_j} \right)^{-\theta_z} \left( \frac{1}{\tau_{ij} w_i} \right)^{-\theta_z} \sum_i M_i^I \left( \frac{z_{0,ij}}{b_{z,i}} \right)^{-\theta_z} \left( \frac{1}{w_i f_i} \right)^{-\theta_z} \left( \frac{1}{\tau_{ij} w_i} \right)^{-\theta_z}
\]

(B.38)

\[
= M_j^I \left( \frac{z_{0,j}}{b_{z,j}} \right)^{-\theta_z} \left( \frac{1}{f_j} \right)^{-\theta_z} \left( \frac{1}{\tau_{jj} w_j} \right)^{-\theta_z}
\]

(B.39)

\[
= \frac{M_j^I}{f_j} \left( \frac{z_{0,j}}{b_{z,j}} \right)^{-\theta_z}
\]

(B.40)

For intermediate good producers, I can show that revenues and costs are proportional to \(z^{\epsilon - 1}\). Using free entry condition, trade balance, and zero-cutoff profit condition, I can show that:

\[
M_j^I = \frac{L_j}{f_j} \frac{1}{\theta_z} \frac{\sigma - 1}{\sigma} \theta_\mu - 1
\]

(B.41)

\[
\left( \frac{z_{0,j}}{b_{z,j}} \right)^{-\theta_z} = \lambda_{jj} \frac{f^{\epsilon, l} \theta_z - (\epsilon - 1)}{f^{\epsilon, l} \theta_z - (\epsilon - 1)}
\]

(B.42)

Combining (B.27), (B.40), and (B.42), I can show that price index for final goods is given by:

\[
P_j^F = a_j \lambda_{jj}^{-1/\theta_z}
\]

(B.43)

where \(a_j\) is some \(j\)-specific constant. Thus, gains from trade, defined as a negative of a change in real wage (or consumption, in absence of tariffs those are the same) when a country moves trade to autarkic equilibrium are given by:

\[
GT_j = 1 - \lambda_{jj}^{-1/\theta_z}
\]

(B.44)

### B.4 Welfare in a Closed Economy

Let \(r^F(\varphi)\) and \(e^F(\varphi)\) be total revenues and expenditures on intermediate inputs of final good producer \(\varphi\). Under monopolistic competition, \(e^F(\varphi) = \frac{\sigma - 1}{\sigma} r^F(\varphi)\). Under second degree
price discrimination, there are two components of expenditures on intermediate inputs. The variable payment will still be equal to \( \frac{\sigma - 1}{\sigma} r_F(\varphi) \) because of the CES structure of the market for final good varieties (final good producers will impose a markup over marginal costs, fixed fee is not in the marginal costs). Fixed fee is a fraction of \( \frac{1}{\epsilon - 1} \) of variable payment for the least productive final good producer \( \varphi_0 \), and so:

\[
e^F(\varphi) = \frac{\sigma - 1}{\sigma} r_F(\varphi) + \frac{1}{\epsilon - 1} \frac{\sigma - 1}{\sigma} r_F(\varphi_0) \tag{B.45}
\]

**Second-degree price discrimination equilibrium**

In the previous section I derived conditions for the mass of firms and domestic productivity cutoffs in an open economy. Setting \( \lambda_{jj} = 1 \), I have the expressions for autarky. Dropping the \( j \) index, I can express labor market clearing condition as:

\[
L = f^{e,I} M^I + f^I M^I \left( \frac{z_0}{b_z} \right)^{-\theta_z} + M^F M^I \int_{z_0}^{\infty} \int_{\varphi_0}^{\infty} \frac{x(\varphi, z)}{z} dF(\varphi) dF(z) + f^{e,F} M^F + f^F M^F \left( \frac{\varphi_0}{b_\varphi} \right)^{-\theta_\varphi} . \tag{B.46}
\]

Plugging optimal values for masses and cutoffs, I can derive the expression for \( x(\varphi_0, z_0) \). The full set of equilibrium conditions is given by:

\[
\left( \frac{z_0}{b_z} \right)^{-\theta_z} = \frac{f^{e,I}}{f^I} \frac{\theta_z - (\epsilon - 1)}{\epsilon - 1} \tag{B.47}
\]

\[
\left( \frac{\varphi_0}{b_\varphi} \right)^{-\theta_\varphi} = \frac{f^{e,F}}{f^F} \frac{\theta_\varphi - (\sigma - 1) \epsilon - \sigma}{\sigma - 1} \tag{B.48}
\]

\[
M^F = \frac{L}{f^{e,F}} \frac{1}{\theta_\varphi \sigma} \tag{B.49}
\]

\[
M^I = \frac{L}{f^{e,I}} \frac{1}{\theta_z \sigma} \frac{\epsilon - 1}{\epsilon - \sigma} \tag{B.50}
\]

\[
\frac{x(\varphi_0, z_0)}{z_0} = \frac{f^F f^I}{L} \frac{1}{\sigma(\epsilon - 1)} \frac{\epsilon - 1}{\epsilon - \sigma} \tag{B.51}
\]

**Monopolistic competition**

The same logic can be used to derive equilibrium conditions for monopolistic competition, the only difference being that there will be no fixed payment part from final good producers to intermediate good producers. Two free entry conditions, two zero-cutoff profit conditions,
CHAPTER B. APPENDICES TO CHAPTER 2

and labor market clearing condition yield:

\[
\left( \frac{z_0}{b_z} \right)^{-\theta_z} = \frac{f^{e,I} \theta_z - (\epsilon - 1)}{\epsilon - 1} \tag{B.52}
\]

\[
\left( \frac{\varphi_0}{b_\varphi} \right)^{-\theta_\varphi} = \frac{f^{e,F} \theta_\varphi - (\sigma - 1)}{\sigma - 1} \tag{B.53}
\]

\[
M^F = \frac{L}{f^{e,F} \theta_\varphi} \frac{1}{\sigma} \tag{B.54}
\]

\[
M^I = \frac{L}{f^{e,I} \theta_z} \frac{1}{\epsilon} \frac{\epsilon - 1}{\sigma - 1} \tag{B.55}
\]

\[
\frac{x(\varphi_0, z_0)}{z_0} = \frac{f^F f^I}{L} \frac{1}{\sigma} (\epsilon - 1) \tag{B.56}
\]

**Social planner**

Social planner chooses how much intermediate inputs to produce by each of the intermediate good producers, how much final goods to produce by each of the final good producer, and how many firms to have in each of the sectors, subject to labor market clearing:

\[
\max_{M^F, M^I, \varphi_0, z_0, x(\varphi, z)} Q \equiv (M^F)^{\frac{\sigma}{\sigma - 1}} (M^I)^{\frac{\epsilon}{\epsilon - 1}} \left( \int_{\varphi_0}^{\infty} \varphi^{\frac{\epsilon - 1}{\sigma}} \left( \int_{z_0}^{\infty} x(\varphi, z) \frac{z}{\epsilon} dF(z) \right)^{\frac{\epsilon - 1}{\epsilon}} dF(\varphi) \right)^{\frac{\sigma}{\sigma - 1}} \tag{B.57}
\]

s.t. \( L = f^{e,I} M^I + f^{I} M^I \left( \frac{z_0}{b_z} \right)^{-\theta_z} + M^F M^I \int_{z_0}^{\infty} \int_{\varphi_0}^{\infty} \frac{x(\varphi, z)}{z} dF(\varphi) dF(z) \)

\[+ f^{e,F} M^F + f^{F} M^F \left( \frac{\varphi_0}{b_\varphi} \right)^{-\theta_\varphi} \tag{B.58}\]
FOC with respect to $M^F, M^I, \varphi_0$ are given by ($\mu$ is the Lagrange multiplier):

\[
\frac{\sigma}{\sigma - 1} \frac{Q}{M^F} = \mu \left[ f^{e,F} + f^F \left( \frac{\varphi_0}{b_{\varphi}} \right)^{-\theta_{\varphi}} + M^I \int_{z_0}^{\infty} \int_{\varphi_0}^{\infty} \frac{x(\varphi, z)}{z} dF(\varphi) dF(z) \right] \tag{B.59}
\]

\[
\frac{\epsilon}{\epsilon - 1} \frac{Q}{M^I} = \mu \left[ f^{e,I} + f^I \left( \frac{\varphi_0}{b_{\varphi}} \right)^{-\theta_{\varphi}} + M^F \int_{z_0}^{\infty} \int_{\varphi_0}^{\infty} \frac{x(\varphi, z)}{z} dF(\varphi) dF(z) \right] 
\]

\[
- (M^F) \frac{\sigma}{\sigma - 1} \left( \frac{Q}{M^F} \right)^{q^{e,F}} \left( \int_{\varphi_0}^{\infty} \left[ \varphi \left( \int_{z_0}^{\infty} [x(\varphi, z)] \frac{x}{z} dF(z) \right) \right] \frac{1}{\epsilon - 1} dF(\varphi) \right) \frac{\sigma - 1}{\sigma - 1}
\]

\times \int_{\varphi_0}^{\infty} \frac{\sigma - 1}{\epsilon - 1} \varphi_{\sigma - 1}^{\sigma - 1} \left( \int_{z_0}^{\infty} [x(\varphi, z)] \frac{x}{z} dF(z) \right) \frac{1}{\epsilon - 1} dF(\varphi) \]

\times [x(\varphi, z_0)]^{\frac{x}{z}} z_0^{-\theta_{\varphi}} dF(\varphi)
\]

\[
- \mu \left[ \theta_{z} z^{-\theta_{z}} \right] f^{I} M^I + M^F M^I \int_{\varphi_0}^{\infty} \frac{x(\varphi, z_0)}{z_0} dF(\varphi) \frac{z_0^{-\theta_{\varphi}}}{b_{\theta_{\varphi}}} dF(\varphi) \tag{B.60}
\]

FOC with respect to $z_0, x(\varphi, z)$ are given by:

\[
- (M^F) \frac{\sigma}{\sigma - 1} \left( \frac{M^I}{\sigma - 1} \right)^{q^{e,F}} \left( \int_{\varphi_0}^{\infty} \left[ \varphi \left( \int_{z_0}^{\infty} [x(\varphi, z)] \frac{x}{z} dF(z) \right) \right] \frac{1}{\epsilon - 1} dF(\varphi) \right) \frac{\sigma - 1}{\sigma - 1}
\]

\times \left[ x(\varphi, z_0) \right]^{\frac{x}{z}} z_0^{-\theta_{\varphi}} dF(\varphi)
\]

\[
\left( M^F \right) \frac{\epsilon}{\epsilon - 1} \left( \frac{M^I}{\epsilon - 1} \right)^{q^{e,F}} \left( \int_{\varphi_0}^{\infty} \left[ \varphi \left( \int_{z_0}^{\infty} [x(\varphi, z)] \frac{x}{z} dF(z) \right) \right] \frac{1}{\epsilon - 1} dF(\varphi) \right) \frac{\epsilon - 1}{\epsilon - 1}
\]

\times \left[ x(\varphi, z) \right]^{\frac{x}{z}} dF(z) dF(\varphi)
\]

\[
= \mu M^F M^I \frac{1}{z} dF(z) dF(\varphi) \tag{B.62}
\]
Rearranging the FOCs, and adding the labor market clearing condition, I get the following system of equilibrium conditions for the social planner:

\[
\left( \frac{z_0}{b_z} \right)^{-\theta_z} = \frac{f^{e,I} \theta_z - (\epsilon - 1)}{\epsilon - 1} \\
\left( \frac{\varphi_0}{b_{\varphi}} \right)^{-\theta_{\varphi}} = \frac{f^{e,F} \theta_{\varphi} - (\sigma - 1)}{\sigma - 1}
\]

\[
M^F = \frac{L}{f^{e,F} \theta_{\varphi}} \frac{1}{\sigma} \frac{1}{\epsilon - 1/\sigma}
\]

\[
M^I = \frac{L}{f^{e,I} \theta_z} \frac{1}{\epsilon} \frac{1}{\sigma - 1/\epsilon}
\]

\[
\frac{x(\varphi_0, z_0)}{z_0} = \frac{f^I f^F}{L} (\sigma \epsilon - 1)
\]

\[
(B.63) \\
(B.64) \\
(B.65) \\
(B.66) \\
(B.67) \\
(B.68)
\]

### B.5 Optimal policies

It is useful to derive the expressions for the mass of firms in different models. Let \( R_j^F \), \( C_j^F \) and \( X_{ij}^F \) be total revenues of final good producers in \( j \), total expenditures on intermediate inputs (including tariffs) in \( j \), and total revenues (net of tariffs) of intermediate good producers who sell from \( i \) to \( j \). Under monopolistic competition:

\[
C_j^F = \frac{R_j^F \sigma - 1}{\sigma} = \sum_i (1 + \tau_{ij}) X_{ij}
\]

\[
= X_j \sum_i (1 + \tau_{ij}) \lambda_{ij}
\]

Using the fact that \( R_j^F = L_j w_j + X_j \sum_i \lambda_{ij} t_{ij} \) and free entry conditions (profits in the final good sector and intermediate good sector are equal to entry costs), I can derive the mass of entrants in both sectors:

\[
M_1^F = \frac{L_1}{f^{e,F}_1} \frac{\sigma - 1}{\sigma} \frac{1 + \sum_{i=1,2} \lambda_{ii} t_{i1}}{1 + \frac{1}{\sigma} \sum_{i=1,2} \lambda_{ii} t_{i1}}
\]

\[
M_1^I = \frac{L_1}{f^{e,I}_1} \frac{\sigma - 1}{\epsilon} \frac{1}{\sigma - 1/\epsilon} \frac{1}{1 + \frac{1}{\sigma} \sum_{i=1,2} \lambda_{ii} t_{i1}}
\]

Under second-degree price discrimination, using (B.45) I can show that:

\[
C_j^F = \frac{R_j^F \sigma - 1}{\sigma} \left( 1 + \frac{1}{\epsilon - 1/\sigma} \theta_{\varphi} - (\sigma - 1) \right)
\]
and so the mass of firms is now given by:

\[
M_F^I = \frac{L_1 \sigma - 1}{f_1^e,F} \frac{1 + \frac{1}{\sigma} \lambda_{i1} t_{i1}}{\sigma \theta - 1 + \frac{1}{\sigma} \kappa \sum_{i=1,2} \lambda_{i1} t_{i1}},
\]

\[
M_I^I = \frac{L_1 \sigma - 1 \theta_\mu - 1}{f_1^e,I} \frac{1}{\theta_\mu - 1 + \frac{1}{\sigma} \kappa \sum_{i=1,2} \lambda_{i1} t_{i1}},
\]

where:

\[
\kappa = 1 - \frac{\sigma - 1}{\epsilon - 1} \frac{\sigma - 1}{\sigma}
\]

**Optimal taxes under monopolistic competition in autarky**

I allow a policymaker to impose a tax on the purchases of intermediate inputs. For every dollar paid for intermediate inputs, final good producers pay \( t \) (if \( t < 1 \) they get a subsidy). For \( \lambda_{i1} = 1 \), it turns out immediately that if \( 1 + t_{i1} = \frac{\epsilon - 1}{\epsilon} \), then (B.72) and (B.73) will coincide with social planner’s equilibrium. Cutoffs are unaffected and \( x(\varphi_0, z_0) \) will be the same as under social planner because of labor market clearing. Thus, optimal policy is given by:

\[
1 + t^{MC} = \frac{\epsilon - 1}{\epsilon}
\]

**Optimal taxes under second-degree price discrimination in autarky**

Under second-degree price discrimination, the productivity cutoff in the final good sector is different from social planner. It can be shown that providing the subsidy:

\[
T^{2DPD*} = -f \frac{\sigma - 1}{\epsilon - 1},
\]

will fix that distortion. This can be seen from (B.45) by setting

\[
r(\varphi_0) - \frac{\sigma - 1}{\sigma} r(\varphi_0) - \frac{1}{\epsilon - 1} \frac{\sigma - 1}{\sigma} r(\varphi_0) - T^{2DPD*} = f^F.
\]

Expression on the right is profit net of fixed costs. The subsidy in (B.79) ensures that \( r(\varphi_0) = \sigma f^F \), the same condition as under social planner or monopolistic competition. Writing down total revenues as a function of labor income and taxes, setting profits net of fixed costs equal to entry costs in both sectors, it can be shown that the optimal tax on intermediate inputs is:

\[
1 + t^{2DPD*} = \frac{\theta_\mu - 1}{\theta_\mu}
\]
Optimal taxes under second-degree price discrimination in autarky when lump-sum production subsidies are not allowed

Equilibrium allocation under second-degree price discrimination is given by:

\[
\left( \frac{z_0}{b_z} \right)^{-\theta_z} = \frac{f^{e,I} \theta_z - (\epsilon - 1)}{f^{I} \epsilon - 1} \quad \text{(B.82)}
\]

\[
\left( \frac{\varphi_0}{b_{\varphi}} \right)^{-\theta_{\varphi}} = \frac{f^{e,F} \theta_{\varphi} - (\sigma - 1) \epsilon - \sigma}{f^{F} \sigma - 1 \epsilon - 1} \quad \text{(B.83)}
\]

\[
M_1^F = \frac{L_1}{f_1^{e,F} \frac{\sigma - 1}{\sigma \theta_{\varphi}} 1 + \frac{1}{\sigma} \kappa t}, \quad \text{(B.84)}
\]

\[
M_1^I = \frac{L_1}{f_1^{e,I} \frac{\sigma - 1}{\theta_{\varphi}} \frac{-1}{\sigma \mu} 1 + \frac{1}{\sigma} \kappa t}, \quad \text{(B.85)}
\]

\[
\frac{x(\varphi_0, z_0)}{z_0} = \frac{f^{F} f^{I}}{L} \frac{\epsilon - 1}{\epsilon - \sigma} \quad \text{(B.86)}
\]

Expressing aggregate real consumption as:

\[
Q = \psi \varphi_0 x(z_0, \varphi_0) \left[ M^F \left( \frac{\varphi_0}{b_{\varphi}} \right)^{-\theta_{\varphi}} \right]^{\frac{\sigma}{\sigma - 1}} \left[ M^I \left( \frac{z_0}{b_z} \right)^{-\theta_z} \right]^{\frac{\epsilon}{\epsilon - 1}}, \quad \text{(B.87)}
\]

and setting the derivative of \( Q \) with respect to \( t \) to zero, yields:

\[
1 + t^{2DPD} = \frac{\theta_{\mu} - 1}{\theta_{\mu}} \frac{1}{\kappa}. \quad \text{(B.88)}
\]

It can be shown that this is a globally-optimal tax rate.

Optimal taxes under second-degree price discrimination in a small open economy

Using the results from the previous parts of this Appendix, and trade balance to derive condition for \( z_{12} \), equilibrium allocation in a small open economy under monopolistic competition
is given by:

\[
M^F_i = \frac{L_1}{f_1^{e,F}} \frac{\sigma - 1}{\sigma} \frac{1 + \frac{1}{\sigma} \sum_{i=1,2} \lambda_{ii} t_i}{1 + \frac{1}{\sigma} \sum_{i=1,2} \lambda_{ii} t_i}.
\] (B.89)

\[
M^I_i = \frac{L_1}{f_1^{e,I}} \frac{\epsilon - 1}{\epsilon} \frac{1}{1 + \frac{1}{\sigma} \sum_{i=1,2} \lambda_{ii} t_i}.
\] (B.90)

\[
\left( \begin{array}{c} \varphi_{0,1} \\ b_{\varphi,1} \end{array} \right)^{-\theta_{\varphi}} = \frac{L \theta_{\varphi} - (\sigma - 1)}{f_1^{e,F}} \frac{1}{\sigma - 1},
\] (B.91)

\[
\left( \begin{array}{c} z_{0,11} \\ b_{z,1} \end{array} \right)^{-\theta_{z}} = \frac{L \theta_{z} - (\epsilon - 1)}{f_1^{e,I}} \frac{1}{\epsilon - 1},
\] (B.92)

\[
\left( \begin{array}{c} z_{0,12} \\ b_{z,1} \end{array} \right)^{-\theta_{z}} = \frac{L \theta_{z} - (\epsilon - 1)}{f_1^{e,I}} \frac{1}{\epsilon - 1},
\] (B.93)

\[
\tilde{\lambda}_{i1} = \frac{M^I_i \left( \frac{z_{0,i1}}{b_{z,i}} \right)^{-\theta_{z}} \left( \frac{w_i t_i (1 + t_i)}{r_0 i} \right)^{1-\epsilon}}{\sum_{i=1,2} M^I_i \left( \frac{z_{0,i1}}{b_{z,i}} \right)^{-\theta_{z}} \left( \frac{w_i t_i (1 + t_i)}{r_0 i} \right)^{1-\epsilon}},
\] (B.94)

\[
\tilde{\lambda}_{i1} = \frac{\lambda_{i1} (1 + t_{i1})}{\sum_{i=1,2} \lambda_{i1} (1 + t_{i1})}.
\] (B.95)

where \(\tilde{\lambda}_{ij}\) is denotes aggregate expenditures including tariffs on inputs from \(i\). I can construct optimal taxes by replicating social-planners allocation from a perspective of a small open economy, as in Demidova and Rodríguez-Clare (2009).

Social planner’s problem in a small open economy under monopolistic competition is
given by:

$$\max_Q \phi_0 \left\{ M^F \tilde{\varphi} \left( \frac{\varphi_0}{b_\varphi} \right)^{-\theta_\varphi} \right\} \frac{\sigma}{\epsilon} \left( M^I a \frac{z_0^{\epsilon-1} \tilde{\theta}_z}{b_z} \right)^{-\theta_z} \left( z_0 \frac{z_0}{b_z} \right)^{-\theta_z} + M^I a_m \frac{z_m^{\epsilon-1} \tilde{\theta}_z}{b_{z,m}} \left( z_m \frac{z_m}{b_{z,m}} \right)^{-\theta_z} $$

(B.96)

s.t. 

$$L = f^{e.I} M^I + f^I M^I \left( \frac{z_0}{b_z} \right)^{-\theta_z} + M^F M^I a \varphi_0^{\sigma-1} \frac{z_0^{\epsilon-1} \tilde{\theta}_z \varphi_0}{b_z} \left( \frac{\varphi_0}{b_\varphi} \right)^{-\theta_\varphi} + M^F M^I \left( \frac{z_e}{b_z} \right)^{-\theta_z} \left( \frac{\varphi_0}{b_\varphi} \right)^{-\theta_\varphi} + M^I a_e \frac{z_e^{\epsilon-1} \tilde{\theta}_z}{b_z} \left( \frac{z_e}{b_z} \right)^{-\theta_z}$$

$$M^I M^F a_m \frac{\epsilon}{\epsilon - 1} w_m \tau z_m^{\epsilon-1} \varphi_0^{\sigma-1} \frac{z_0^{\epsilon-1} \tilde{\theta}_z \varphi_0}{b_z} \left( \frac{z_m}{b_z} \right)^{-\theta_z} \left( \frac{\varphi_0}{b_\varphi} \right)^{-\theta_\varphi} = M^I a_e \frac{z_e^{\epsilon-1} A^{1/\epsilon}}{\epsilon - 1} \tilde{\theta}_z \left( \frac{z_e}{b_z} \right)^{-\theta_z}$$

(B.97)

$$\epsilon f^I_e = a_e \frac{\epsilon-1}{z_e^{\epsilon-1} A^{1/\epsilon}}$$

(B.98)

$$\epsilon f^I_m = M^F \varphi_0^{\sigma-1} \frac{z_0^{\epsilon-1} \tilde{\theta}_z \varphi_0}{b_z} a_m \frac{\epsilon}{\epsilon - 1} w_m \tau$$

(B.99)

$$X = \left[ M^F, M^I, \varphi_0, z_0, z_m, z_e, a, a_m, a_e \right]$$

(B.100)

where maximization problem is subject to trade balance and zero-cutoff profit conditions, \(a, a^e, a^m\) are smallest quantities produced by domestic intermediate good producer for sales at home, by domestic intermediate good producer for sales abroad, by foreign intermediate good producer for sales at home. Taking first order-conditions, rearranging them, I can show that optimal taxes under monopolistic competition that make two allocations coincide are given by:

$$1 + t_{11}^{MC} = \frac{\epsilon - 1}{\epsilon} < 1,$$

(B.101)

$$1 + t_{21}^{MC} = \frac{\theta_z}{\theta_z^{\epsilon-1} - 1} = \frac{\theta_z}{\theta_z^{\epsilon+\epsilon^{*11}} - 1} < 1.$$  

(B.102)
Optimal policy under second-degree price discrimination in a small open economy when lump-sum taxes are not allowed.

Using the results from the previous sections of this Appendix, I can express the system of equations that determine equilibrium allocation under second-degree price discrimination as:

\[
\begin{align*}
M_{Fri1}^F &= \frac{L_1}{f_1^{e,F}} \frac{\sigma - 1}{\sigma \theta \varphi} \left( 1 + \sum_{i=1,2} \lambda_{i1} t_{i1} \right), \\
M_{Fri1}^I &= \frac{L_1}{f_1^{e,I}} \frac{\sigma - 1}{\theta \mu} \left( 1 + \frac{1}{\theta \kappa} \sum_{i=1,2} \lambda_{i1} t_{i1} \right), \\
(z_{0,11})^{-\theta_z} &= \frac{\lambda_{11}}{f_1^{e,I}} \frac{L \theta_z - (\epsilon - 1)}{\epsilon - 1}, \\
(z_{0,12})^{-\theta_z} &= \frac{\lambda_{21}}{f_1^{e,I}} \frac{L \theta_z - (\epsilon - 1)}{\epsilon - 1}, \\
\tilde{\lambda}_{i1} &= \frac{\frac{M_i}{f_1^{e,I}} \left( \frac{z_{0,i1}}{b_{z,i}} \right)^{-\theta_z} \left( \frac{w_{i1} \tau_{i1}(1+t_{i1})}{z_{0,i1}} \right)^{1-\epsilon}}{\sum_{i=1,2} \frac{M_i}{f_1^{e,I}} \left( \frac{z_{0,i1}}{b_{z,i}} \right)^{-\theta_z} \left( \frac{w_{i1} \tau_{i1}(1+t_{i1})}{z_{0,i1}} \right)^{1-\epsilon}}, \\
\tilde{\lambda}_{i1} &= \frac{\lambda_{i1}(1+t_{i1})}{\sum_{i=1,2} \lambda_{i1}(1+t_{i1})},
\end{align*}
\]

Using grid search over feasible values of parameters, I show that welfare is maximized for:

\[
1 + t_{11}^{2DPD} = \frac{\theta \mu - 1}{\theta \mu} \frac{1}{\kappa},
\]

\[
1 + t_{21}^{2DPD} = \frac{\theta_z}{\theta_z - \frac{1}{\epsilon}}(1 + t_{11}^{2DPD})
\]

**B.6 Hat algebra**

Using the approach of Dekle, Eaton, and S. Kortum (2008), I can calculate changes in equilibrium variables using the 'hat algebra', by expressing the system of equilibrium equations for second-degree price discrimination in terms of changes of variables, denote by \( \hat{y} = y' / y \) where \( y' \) is a new value in equilibrium. I can compute changes in all variables in response to the new level of taxes \( t'_{ij} \). For this I need to know the values of parameters \( \sigma, \epsilon, \theta_z, \theta \varphi \) and initial domestic trade share \( \lambda_{11} \). In the paper, I calibrate the values of the parameters and calculate changes in welfare for a range of \( \lambda_{11} \).
Chapter C

Appendices to Chapter 3

C.1 Estimation of $\bar{\theta}$

An estimate of $\bar{\theta}$ is required to compute model-implied $\ln \tilde{F}_{ij}$ and $\ln \tilde{\tau}_{ij}$ as functions of $\ln x_{ij}$, $\ln N_{ij}$, and estimated fixed effects. We follow Eaton, S. Kortum, and Kramarz (2011) and derive the following expression

$$\frac{x_{ilj}}{x_{il}} = \left( \frac{N_{ij}}{N_{il}} \right)^{-1/\bar{\theta}}$$

where $x_{ilj}$ are average exports per firm for firms from $i$ that sell in market $l$ but restricted to those firms that sell in markets $l$ and $j$. EKK have information on domestic sales for each firm, so they use $l = i$. We do not have such information, so we use $l^*(i) \equiv \arg \max_k N_{ik}$, that is, the largest destination market for each origin country $i$ (e.g., the United States for Mexico). Letting

$$z_{ij} \equiv \frac{x_{il^*(i)j}}{x_{il^*(i)|l^*(i)}}$$

and

$$m_{ij} \equiv \frac{N_{ij}}{N_{il^*(i)}}$$

then we have

$$\ln z_{ij} = -\frac{1}{\bar{\theta}} \ln N_{ij}.$$ 

This suggests an OLS regression to recover an estimate for $\bar{\theta}$.

Eaton, S. Kortum, and Kramarz (2011) estimate this regression for French firm-level data (including information on sales in France) and obtain a coefficient of $-0.57$, which implies $\bar{\theta} = 1.75$. In their case, they keep in their estimating sample only firms with positive sales in France, so the variables $x_{FF|j}$ and $N_{Fj}$ are calculated based on the same set of firms. To implement an approach comparable to theirs, we drop all firms from country $i$ that do not sell to $l^*(i)$, so the sample includes only $N_{il^*(i)}$ firms for country $i$. This implies that all firms that make up $N_{ij}$ are also selling to $l^*(i)$. Figure 3.5 in the paper reproduces Figure 3 from
Eaton, S. Kortum, and Kramarz (2011) by plotting the variables in equation (C.4). The slope in the graph is equal to 1/\(\hat{\theta}\), and the corresponding estimated values are reported in Table C.8. Based on all observations in the core sample of countries and using no weighting, the estimated \(\hat{\theta}\) is over 19. But in Figure 3.5 for small values of \(m_{ij}\), which correspond to small values of \(N_{ij}\), there is a lot of dispersion in \(z_{ij}\). To minimize the effect of that noise we weight observations by \(\sqrt{N_{ij}}\) and this lowers the estimate of \(\bar{\theta}\) to 4.8. Finally, when we drop all observations with \(N_{ij} < 100\) (remember that here \(N_{ij}\) measures the number of firms from country \(i\) that sell to country \(j\) and also to \(l^* (i)\)) we obtain \(\bar{\theta} = 2.3\), which is still higher than in Eaton, S. Kortum, and Kramarz (2011). We will use this estimate in our simulations of the intensive margin elasticity.

C.2 QQ-Estimation of \(\bar{\sigma}_\varphi\)

Exports from country \(i\) to country \(j\) of a firm with productivity \(\varphi\) in the model with CES preferences and monopolistic competition is given by

\[x_{ij} (\varphi) = \sigma F_{ij} \left( \varphi/\varphi^*_{ij} \right)^{\sigma-1}.\]

Since \(\ln \varphi \sim N(\mu_{\varphi,i}, \sigma_{\varphi})\) then \(\ln x_{ij}(\varphi) \sim \text{Ntrunc}(\mu_{\varphi,ij}, \sigma_{\varphi}; \ln (\sigma F_{ij}))\), where \(\sigma_{\varphi} = \sigma \varphi (\sigma - 1)\), \(\mu_{\varphi,ij} = \mu_{\varphi,i} (\sigma - 1) + \ln (\sigma F_{ij}) + (1 - \sigma) \ln (\varphi^*_{ij})\), and the truncation point is \(\ln (\sigma F_{ij})\).

As in HMT, we estimate \(\bar{\sigma}_\varphi\) using a quantile-quantile regression, which minimizes the distance between the theoretical and empirical quantiles of log exports. Empirical quantiles are given by:

\[Q_{ij,n}^E = \ln x_{ij,n}\]

where \(n\) is the rank of the firm among exporters from \(i\) to \(j\). We calculate theoretical quantiles of exports from \(i\) to \(j\) as

\[Q_{ij,n}^T = \mu_{\varphi,ij} + \sigma_{\varphi} \Phi^{-1} \left( \hat{\Phi}_{ij,n} \right),\]

where \(\hat{\Phi}_{ij,n} = \frac{N_i - (n-1)}{N_i}\) is the empirical CDF and \(N_i\) is the imputed number of firms from the BR data. Following HMT we adjust the empirical CDF so that \(\hat{\Phi}_{ij,n} = \frac{N_i - (n-1) - 0.3}{N_i + 0.4}\) since otherwise we would get \(\Phi^{-1} \left( \hat{\Phi}_{ij,1} \right) = \infty\) when \(n = 1\). The QQ-estimator of \(\bar{\sigma}_\varphi\) is the coefficient \(\beta\) obtained from the regression

\[\ln x_{ij,n} = \alpha_{ij} + \beta \Phi^{-1} \left( \hat{\Phi}_{ij,n} \right) + \varepsilon_{ij,n}.\]

Table C.9 reports the QQ-estimate of \(\bar{\sigma}_\varphi\). We report three sets of estimates: for the full sample, the largest 50% of firms and the largest 25% of firms for each origin-destination pair in each year. According to the model, the estimates of the slope should not change when we consider different sub-samples, but this is not the case in Table C.9. This comes from a not very surprising empirical failure of the simple Melitz-lognormal model outlined in the first part of the previous section: whereas this model implies that the sales distribution for any
country pair should be distributed as a truncated lognormal (with the truncation at sales of $\sigma_{F_{ij}}$), no such truncation exists in the data (i.e., we observe exporters with very small sales).

A related issue is that our estimates for either of the sub-samples are significantly larger than the HMT estimate of 2.4. The difference comes from the fact that HMT assume that the sales distribution for any $ij$ pair is lognormal, whereas we stick close to the simple model and assume that it is a truncated lognormal, and then use data for $N_{ij}$ and our estimated values $N_i$ to derive implicit truncation points. These truncation points tend to be on the right tail of the distribution, since $N_{ij}/N_i$ tends to be quite low, hence the small $\bar{\sigma}_\sigma$ estimated by HMT would not be able to match the observed dispersion in the sales of exporters. In general, the higher the $N_i$ one takes as an input in the QQ regression, the higher the estimate of the shape parameter one obtains.

In private correspondence, the authors of HMT pointed out that their approach would be consistent with the Melitz-lognormal model if one allows for heterogeneous fixed costs and lets the variance of these costs go to infinity, whereas our approach would be right if the variance goes to zero. This is part of our motivation in allowing for heterogeneous fixed costs and then in using MLE to estimate the full Melitz-lognormal model.

C.3 Quasi-Bayesian Estimation for the full Melitz-lognormal model

The likelihood function is a product of density functions of individual firms that sell or do not sell to multiple destinations. In this section we will use the notation from Section 4 of the paper. Let $\bar{\varphi}_i \equiv (\sigma - 1)[\ln \varphi - \mu_\varphi, i]$ be a random variable that denotes deviations from mean productivity for country $i$. Individual firm density of export sales $(x_{i1}, ..., x_{iJ})$ can be written as:

$$f_{X_{i1}, ..., x_{ij}, X_{ij}}(x_{i1}, ..., x_{iJ}) = \int_\omega f_{X_{i1}, ..., X_{ij} \mid \bar{\varphi}_i}(x_{i1}, ..., x_{iJ} \mid \omega)f_{\bar{\varphi}_i}(\omega)d\omega$$  \hspace{1cm} (C.8)

$$= \int_\omega \prod_j f_{X_{ij} \mid \bar{\varphi}_i}(x_{ij} \mid \omega)f_{\bar{\varphi}_i}(\omega)d\omega$$  \hspace{1cm} (C.9)

where the second equality comes from the fact that conditional on productivity, sales are independent across markets (as well as the probability of selling to those markets). We now need to characterize $g_{X_{ij} \mid \bar{\varphi}_i}(g_{ij} \mid \omega)$ to calculate the likelihood function. In general we have:

$$f_{X_{ij} \mid \bar{\varphi}_i}(x_{ij} \mid \omega) = [f_{Z_{ij} \mid \bar{\varphi}_i}(x_{ij} \mid \omega) \Pr\{Z_{ij} \geq \ln \sigma + \ln f_{ij} \mid \bar{\varphi}_i = \omega, Z_{ij} = x_{ij}\}]^{1(x_{ij} \neq \emptyset)} \times$$

$$\times [\Pr\{\ln \sigma + \ln f_{ij} \geq Z_{ij} \mid \bar{\varphi}_i = \omega\}]^{1(x_{ij} = \emptyset)}$$  \hspace{1cm} (C.10)
The term on the first line of C.10 corresponds to the density function for the cases when we observe exports, while the second line corresponds to the mass at the point $x_{ij} = \emptyset$.

For the case when sales are not zero $X_{ij} = Z_{ij}$ and

$$Z_{ij} | [\bar{\varphi}_i = \omega] = \omega + d_{ij} + \ln \alpha - \mu_\alpha \quad (C.11)$$

$$Z_{ij} | [\bar{\varphi}_i = \omega] \sim N(d_{ij} + \omega, \sigma_{\alpha,i}^2) \quad (C.12)$$

In addition

$$\Pr[Z_{ij} \geq \ln \sigma + \ln f_{ij} | \bar{\varphi}_i = \omega, Z_{ij} = x_{ij}] = \Pr[\ln \sigma + \ln f_{ij} - x_{ij} | \ln \alpha - \mu_\alpha = x_{ij} - d_{ij} - \omega]$$

$$\ln \sigma + \ln f_{ij} | [\ln \alpha - \mu_\alpha = x_{ij} - d_{ij} - \omega] \sim N(\mu_1, \sigma^2_{1,i}) \quad (C.13)$$

$$\mu_1 = \bar{\mu}_{f,ij} + \frac{\sigma_{\alpha f,i}}{\sigma_{\alpha,i}}(x_{ij} - d_{ij} - \omega)$$

$$\sigma^2_{1,i} = \sigma^2_{f,i} (1 - \rho^2_i) \quad (C.14)$$

Finally:

$$\Pr[Z_{ij} \leq \ln \sigma + \ln f_{ij} | \bar{\varphi}_i = \omega] = \Pr[- \ln \sigma - \ln f_{ij} + (\ln \alpha - \mu_\alpha) + d_{ij} \leq -\omega] \quad (C.15)$$

$$- \ln \sigma - \ln f_{ij} + (\ln \alpha - \mu_\alpha) + d_{ij} \sim N(- \bar{\mu}_{f,ij} + d_{ij}, \sigma^2_{2,i}) \quad (C.16)$$

$$\sigma^2_{2,i} = \sigma^2_{f,i} + \sigma^2_{\alpha,i} - 2\sigma_{\alpha f,i}$$

Let $\phi$ and $\Phi$ denote PDF and CDF of standard normal. Plugging functional forms into C.10 we can get the object of interest:

$$f_{X_{i1}, \ldots, X_{iJ}}(x_{1i}, \ldots, x_{iJ}) =$$

$$\int \prod \left\{ \left[ \frac{1}{\sigma_{\alpha,i}} \phi \left( \frac{x_{ij} - d_{ij} - \omega}{\sigma_{\alpha,i}} \right) \right] \Phi \left( \frac{x_{ij} - \left[ \bar{\mu}_{f,ij} + \frac{\sigma_{\alpha f,i}}{\sigma_{\alpha,i}}(x_{ij} - d_{ij} - \omega) \right]}{\sqrt{\sigma^2_{f,i} (1 - \rho^2_i)}} \right) I(x_{ij} \neq \emptyset) \right\}$$

$$\times \left\{ \Phi \left( \frac{-\omega + \bar{\mu}_{f,ij} - d_{ij}}{\sqrt{\sigma^2_{f,i} + \sigma^2_{\alpha,i} - 2\sigma_{\alpha f,i}}} \right) I(x_{ij} = \emptyset) \right\} \frac{1}{\sigma_{\bar{\varphi},i}} \phi \left( \frac{\omega}{\sigma_{\bar{\varphi},i}} \right) d\omega \quad (C.17)$$

$$g_{X_{i1}, \ldots, X_{iJ}}(x_{1i}, \ldots, x_{iJ}) = f_{X_{i1}, \ldots, X_{iJ} \cap \text{Is an exporter}}(x_{1i}, \ldots, x_{iJ} \cap \text{Is an exporter}) \quad (C.18)$$

However, since we only have a truncated sample of $X_{ij}$'s (as we don’t observe sales of firms that do not export), we need to normalize the density by the inverse of probability that a firm is selling to at least one destination, and so we are interested in the object:

$$g_{X_{i1}, \ldots, X_{iJ}}(x_{1i}, \ldots, x_{iJ}) = f_{X_{i1}, \ldots, X_{iJ} \cap \text{Is an exporter}}(x_{1i}, \ldots, x_{iJ} \cap \text{Is an exporter}) \quad (C.19)$$
The likelihood function is a product of density functions (C.20). Parameters to estimate are

\[ \theta_i = \{d_{ij}, \bar{\mu}_{f,i,j}, \bar{\sigma}_{\varphi,i}, \bar{\sigma}_{\alpha,i}, \bar{\sigma}_{f,i}, \rho_i \} \]

To compute density (C.18), which is in the numerator of (C.20). We can think of (C.18) in the following general form:

\[
f_{X_{i1},...,X_{ij}}(x_{i1},...,x_{ij}) = \int_{\omega} G(\omega) \phi \left( \frac{\omega}{\sigma_{\varphi,i}} \right) d\omega
\]

where \( G(\omega) \) is a known function of \( \omega \). Using change of variables \( \tilde{\omega} = \frac{\omega}{\sqrt{2}\sigma_{\varphi,i}} \) and \( d\omega = \sqrt{2}\sigma_{\varphi,i} d\tilde{\omega} \) we can write:

\[
f_{X_{i1},...,X_{ij}}(x_{i1},...,x_{ij}) = \int_{\tilde{\omega}} G(\sqrt{2}\sigma_{\varphi,i} \tilde{\omega}) \frac{\sigma_{\varphi,i}}{\sqrt{\pi}} \exp \left( -\left[ \frac{\tilde{\omega}}{\sqrt{2}\sigma_{\varphi,i}} \right]^2 \right) d\tilde{\omega}
\]

We can speed up the process to calculate object in (C.22) by applying a Gauss-Hermite quadrature. In general:

\[
\int_x g(x) \exp(-x^2) dx \approx \sum_i (g(x_i)w_i)
\]

we calculate 33 values of the variable \( x_i \) as well as weights \( w_i \) using the Gauss-Hermite method.

### C.4 MCMC algorithm

Since the likelihood function in the equation 3.36 in the paper is highly nonlinear there may exist multiple local maxima. We thus estimate a vector of parameters \( \theta \) (defined in equation (C.20)).
3.37 in the paper) using the methodology developed in Chernozhukov and Hong (2003). This procedure not only gives us point estimates, but also yields confidence intervals for the estimated parameters, intensive marginal elasticity implied by the model, and elasticity of trade costs with respect to distance. We implement Chernozhukov and Hong (2003) procedure using Metropolis-Hastings Monte-Carlo Markov chain algorithm. This algorithm yields a chain of parameter draws \( \{\theta_i^{(n)}\}_{n=1}^N \) for each origin \( i \) such that \( \bar{\theta}_i \equiv \frac{1}{N} \sum_n \theta_i^{(n)} \) is a consistent estimate of \( \theta_i \). Moreover, using the values of the parameters in the chain we can construct confidence intervals for some functions \( f(\theta_i) \). The chain of parameters \( \{\theta_i^{(n)}\}_{n=1}^N \) for each origin is constructed in the following way:

- **Step 1.** Randomly choose a starting guess \( \theta_i^{(0)} \).
- **Step 2.** Draw a candidate vector of parameters for the chain’s \( n + 1 \) value as \( \tilde{\theta}_i^{(n+1)} = \theta_i^{(n)} + \psi^{(n)} \), where \( \theta_i^{(n)} \) is the value \( n \) in the chains and \( \psi^{(n)} \) is a vector of \( iid \) shocks taken from the mean-zero normal distribution with some variance. At each step all but 1 elements of \( \psi^{(n)} \) are zero. In other words, we only add an \( iid \) shock to one parameter at each step of the chain. Since the vector \( \theta_i \) has 34 elements for each \( i \), we try a new value for each parameter every 34 steps.
- **Step 3.** Calculate \( \theta_i^{(n+1)} \) in the following way:

\[
\theta_i^{(n+1)} = \begin{cases} 
\tilde{\theta}_i^{(n+1)} & \text{with probability } \min \left[ 1; L(\tilde{\theta}_i^{(n+1)}) - L(\theta_i^{(n)}) \right] \\
\theta_i^{(n)} & \text{otherwise},
\end{cases}
\]

where \( L(\theta) \) is define in the equation 3.36 in the paper.

We repeat the procedure until we have at least 34 million draws (1 million draws per parameter) in the chain after we discard the first 100,000 iterations (known as “burn-in period”). For each origin we construct 5 different chains with different starting guesses to check that our estimates are robust with respect to the starting values (we discuss convergence of the chains in the Online Appendix C.5).

Having estimated the chains, we take 1,000 random draws from the chains for each origin with replacement. We use those draws to calculate point estimates (averages) as well as 95% confidence intervals. Using those draws we simulate the model 1,000 times and run IME regressions the way we run them in the data. Finally, run 1,000 regressions of the estimates \( d_{ij} \) and \( \mu_{f,ij} \) on log distance with origin and destination fixed effects for the 4 destinations (USA, France, Germany, and Japan) and interpret the results as elasticity of variable and fixed trade costs with respect to distance.
C.5 Convergence of the Monte Carlo Markov chains

As mentioned before, we ran our estimation algorithm 5 times per origin country which gave us 5 different chains that started at different random initial guesses. This lets us compare underlying distributions of parameter estimates across different chains. We checked the convergence of the chains using the following criteria:

- comparing the means of parameter estimates in the first and second half of the chain;
- comparing parameter means across different chains for the same origin.

If the means calculated for the two parts of the chains are statistically indistinguishable, we conclude that the chains converged. If the means across different chains are statistically indistinguishable, we conclude that multiple chains converged to the same region. It turns out that in some cases the chains didn’t converge.\(^1\) In those cases we disregarded the chains that didn’t converge. Out of remaining chains, we randomly choose one chain per origin to conduct our numerical analysis. Figures C.1-C.34 plot the means for the first and second half of each chains, as well as an overall mean with the range between 2.5\(^{th}\) and 97.5\(^{th}\) percentiles of the draws in those chains that converged. As can be seen, means and their confidence intervals are virtually indistinguishable for all countries and all parameters.

\(^1\)It happens, for example, when some of the values for \(\sigma_{\varphi,i}, \sigma_{\alpha,i}\) exploded
Figure C.1: Summary plots for QBE chains: $d_{i1}$

- mean: full chain
- $\star$ mean: 1st half
- $\blacklozenge$ mean: 2nd half
- 2.5 - 97.5 percentiles range
Figure C.2: Summary plots for QBE chains: $d_{i2}$
Figure C.3: Summary plots for QBE chains: $d_{i3}$
Figure C.4: Summary plots for QBE chains: $d_{i4}$
Figure C.5: Summary plots for QBE chains: $d_{i5}$

- Mean: full chain
- Red: Mean: 1st half
- Blue: Mean: 2nd half
- Line: 2.5 – 97.5 percentiles range
Figure C.6: Summary plots for QBE chains: $d_{a6}$

<table>
<thead>
<tr>
<th>country 33</th>
<th>country 35</th>
<th>country 36</th>
<th>country 37</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>country 27</td>
<td>country 28</td>
<td>country 29</td>
<td>country 30</td>
</tr>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>country 30</td>
<td>country 31</td>
<td>country 32</td>
<td></td>
</tr>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>country 14</td>
<td>country 15</td>
<td>country 16</td>
<td>country 17</td>
</tr>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>country 18</td>
<td>country 19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>country 8</td>
<td>country 9</td>
<td>country 10</td>
<td>country 11</td>
</tr>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>country 12</td>
<td>country 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>country 1</td>
<td>country 2</td>
<td>country 3</td>
<td>country 4</td>
</tr>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>country 5</td>
<td>country 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
</tbody>
</table>

$\times$ mean: full chain
\* mean: 1st half
\* mean: 2nd half
--- 2.5 – 97.5 percentiles range
Figure C.7: Summary plots for QBE chains: $d_{i7}$
Figure C.8: Summary plots for QBE chains: $d_{i8}$

- Mean: full chain
- Mean: 1st half
- Mean: 2nd half
- 2.5 – 97.5 percentiles range
Figure C.9: Summary plots for QBE chains: $d_{ij}$
Figure C.10: Summary plots for QBE chains: $d_{i10}$
Figure C.11: Summary plots for QBE chains: $d_{i11}$
CHAPTER C. APPENDICES TO CHAPTER 3

Figure C.12: Summary plots for QBE chains: $d_{112}$

<table>
<thead>
<tr>
<th>country 33</th>
<th>country 35</th>
<th>country 36</th>
<th>country 37</th>
</tr>
</thead>
<tbody>
<tr>
<td>country 27</td>
<td>country 28</td>
<td>country 29</td>
<td>country 30</td>
</tr>
<tr>
<td>country 20</td>
<td>country 22</td>
<td>country 23</td>
<td>country 24</td>
</tr>
<tr>
<td>country 14</td>
<td>country 15</td>
<td>country 16</td>
<td>country 17</td>
</tr>
<tr>
<td>country 8</td>
<td>country 9</td>
<td>country 10</td>
<td>country 11</td>
</tr>
<tr>
<td>country 1</td>
<td>country 2</td>
<td>country 3</td>
<td>country 4</td>
</tr>
</tbody>
</table>

$d_{112}$

- mean: full chain
- mean: 1st half
- mean: 2nd half
- 2.5 – 97.5 percentiles range
Figure C.13: Summary plots for QBE chains: $d_{13}$
Figure C.14: Summary plots for QBE chains: $d_{14}$

- country 33
- country 35
- country 36
- country 37
- country 27
- country 29
- country 30
- country 31
- country 32
- country 20
- country 22
- country 23
- country 24
- country 25
- country 26
- country 14
- country 15
- country 16
- country 17
- country 18
- country 19
- country 8
- country 9
- country 10
- country 11
- country 12
- country 13
- country 1
- country 2
- country 3
- country 4
- country 5
- country 6

$d_{14}$

- mean: full chain
- mean: 1st half
- mean: 2nd half
- $2.5 - 97.5$ percentiles range
Figure C.15: Summary plots for QBE chains: $d_{i15}$

- Mean: full chain
- Mean: 1st half
- Mean: 2nd half
- 2.5 - 97.5 percentiles range
Figure C.16: Summary plots for QBE chains: $\mu_{f,i}$
Figure C.17: Summary plots for QBE chains: $\mu_{f,i2}$
Figure C.18: Summary plots for QBE chains: $\mu_{f,3}$
Figure C.19: Summary plots for QBE chains: $\mu_{f,i\delta}$
Figure C.20: Summary plots for QBE chains: $\mu_{f,5}$
Figure C.21: Summary plots for QBE chains: $\mu_{f,6}$
Figure C.22: Summary plots for QBE chains: $\mu_{f,i}$

- **Mean:** full chain
- **Red Star:** mean: 1st half
- **Blue Star:** mean: 2nd half
- **Black Line:** 2.5 - 97.5 percentiles range
Figure C.23: Summary plots for QBE chains: $\mu_{f,8}$

<table>
<thead>
<tr>
<th>country 33</th>
<th>country 35</th>
<th>country 36</th>
<th>country 37</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
</tr>
<tr>
<td>chain 4</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
</tr>
<tr>
<td>chain 3</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
</tr>
<tr>
<td>chain 2</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>country 27</th>
<th>country 28</th>
<th>country 29</th>
<th>country 30</th>
<th>country 31</th>
<th>country 32</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 4</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 3</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 2</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>country 20</th>
<th>country 22</th>
<th>country 23</th>
<th>country 24</th>
<th>country 25</th>
<th>country 26</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 4</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 3</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 2</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>country 14</th>
<th>country 15</th>
<th>country 16</th>
<th>country 17</th>
<th>country 18</th>
<th>country 19</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 4</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 3</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 2</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>country 8</th>
<th>country 9</th>
<th>country 10</th>
<th>country 11</th>
<th>country 12</th>
<th>country 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 4</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 3</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 2</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>country 1</th>
<th>country 2</th>
<th>country 3</th>
<th>country 4</th>
<th>country 5</th>
<th>country 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>chain 5</td>
<td>chain 4</td>
<td>chain 3</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 4</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 3</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 2</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
<tr>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
<td>chain 1</td>
<td>chain 5</td>
<td>chain 2</td>
</tr>
</tbody>
</table>

$\mu_{f,8}$

- `mean: full chain`
- `mean: 1st half`
- `mean: 2nd half`
- `2.5 - 97.5 percentiles range`
Figure C.24: Summary plots for QBE chains: $\mu_{f,9}$
Figure C.25: Summary plots for QBE chains: $\mu_{f,i10}$
Figure C.26: Summary plots for QBE chains: $\mu_{f,i11}$
Figure C.27: Summary plots for QBE chains: $\mu_{f_{i12}}$
Figure C.28: Summary plots for QBE chains: $\mu_{f,13}$
Figure C.29: Summary plots for QBE chains: $\mu_{f,14}$
Figure C.30: Summary plots for QBE chains: $\mu_{f,15}$
Figure C.31: Summary plots for QBE chains: $\bar{\sigma}_{\phi,i}$

- Country names: country 33, country 35, country 36, country 37, country 27, country 28, country 29, country 30, country 31, country 32, country 20, country 22, country 23, country 24, country 25, country 26, country 14, country 15, country 16, country 17, country 18, country 19, country 8, country 9, country 10, country 11, country 12, country 13, country 1, country 2, country 3, country 4, country 5, country 6.

Legend:
- *: mean: full chain
- : mean: 1st half
- : mean: 2nd half
- 2.5 - 97.5 percentiles range
Figure C.32: Summary plots for QBE chains: $\sigma_{\alpha,i}$

<table>
<thead>
<tr>
<th>chain</th>
<th>country 33</th>
<th>country 35</th>
<th>country 36</th>
<th>country 37</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>chain</th>
<th>country 27</th>
<th>country 28</th>
<th>country 29</th>
<th>country 30</th>
<th>country 31</th>
<th>country 32</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>chain</th>
<th>country 20</th>
<th>country 22</th>
<th>country 23</th>
<th>country 24</th>
<th>country 25</th>
<th>country 26</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>chain</th>
<th>country 14</th>
<th>country 15</th>
<th>country 16</th>
<th>country 17</th>
<th>country 18</th>
<th>country 19</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>chain</th>
<th>country 8</th>
<th>country 9</th>
<th>country 10</th>
<th>country 11</th>
<th>country 12</th>
<th>country 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>chain</th>
<th>country 1</th>
<th>country 2</th>
<th>country 3</th>
<th>country 4</th>
<th>country 5</th>
<th>country 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $\sigma_{\alpha,i}$
- mean: full chain
- mean: 1st half
- mean: 2nd half
- 2.5 - 97.5 percentiles range
Figure C.33: Summary plots for QBE chains: $\sigma_{f,i}$
Figure C.34: Summary plots for QBE chains: $\rho_i$
## C.7 Tables

Table C.1: Additional EDD countries and years in the Extended Sample (only EDD statistics available)

<table>
<thead>
<tr>
<th>ISO3</th>
<th>Country name</th>
<th>1st year</th>
<th>Last year</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRA</td>
<td>Brazil</td>
<td>2003</td>
<td>2013</td>
</tr>
<tr>
<td>DK</td>
<td>Denmark</td>
<td>2003</td>
<td>2012</td>
</tr>
<tr>
<td>ESP</td>
<td>Spain</td>
<td>2005</td>
<td>2013</td>
</tr>
<tr>
<td>EST</td>
<td>Estonia</td>
<td>2003</td>
<td>2011</td>
</tr>
<tr>
<td>KWT</td>
<td>Kuwait</td>
<td>2009</td>
<td>2010</td>
</tr>
<tr>
<td>LKA</td>
<td>Sri Lanka</td>
<td>2013</td>
<td>2013</td>
</tr>
<tr>
<td>NOR</td>
<td>Norway</td>
<td>2003</td>
<td>2013</td>
</tr>
<tr>
<td>PRT</td>
<td>Portugal</td>
<td>2003</td>
<td>2012</td>
</tr>
<tr>
<td>SWZ</td>
<td>Swaziland</td>
<td>2012</td>
<td>2012</td>
</tr>
<tr>
<td>TUR</td>
<td>Turkey</td>
<td>2003</td>
<td>2013</td>
</tr>
</tbody>
</table>
Table C.2: IME regressions, extended sample

<table>
<thead>
<tr>
<th></th>
<th>Coefficient from ln $x_{ij}$ on ln $X_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel a: country pairs with $N_{ij} \geq 100$</td>
<td></td>
</tr>
<tr>
<td>IM elasticity</td>
<td>0.437*** 0.450*** 0.381***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0037] [0.0028] [0.0034]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.55 0.75 0.83</td>
</tr>
<tr>
<td>Variation in ln $X_{ij}$ explained by FE,%</td>
<td>0.01 0.14 0.55</td>
</tr>
<tr>
<td>Observations</td>
<td>14,318 14,300 13,964</td>
</tr>
<tr>
<td>Panel b: all country pairs</td>
<td></td>
</tr>
<tr>
<td>IM elasticity</td>
<td>0.434*** 0.477*** 0.516***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0016] [0.0016] [0.0016]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.69 0.77 0.80</td>
</tr>
<tr>
<td>Variation in ln $X_{ij}$ explained by FE, %</td>
<td>0.00 0.22 0.56</td>
</tr>
<tr>
<td>Observations</td>
<td>52,775 52,775 52,658</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
</tr>
<tr>
<td>Origin $\times$ year FE</td>
<td>Yes       Yes</td>
</tr>
<tr>
<td>Destination $\times$ year FE</td>
<td>Yes       Yes</td>
</tr>
</tbody>
</table>

Note: robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table C.3: IME regression by income group

<table>
<thead>
<tr>
<th></th>
<th>Panel a: core sample</th>
<th>Panel a: extended sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity</td>
<td>0.567*** 0.504*** 0.471***</td>
<td>0.682*** 0.626*** 0.550***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0032] [0.0028] [0.0085]</td>
<td>[0.0027] [0.0025] [0.0046]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.82 0.89 0.95</td>
<td>0.81 0.88 0.94</td>
</tr>
<tr>
<td>Observations</td>
<td>14,893 14,326 4,108</td>
<td>24,227 23,519 12,537</td>
</tr>
</tbody>
</table>

HS2 × year FE Yes
Origin × HS2 × year FE Yes Yes
Destination × HS2 × year FE Yes

Note: HS2 industries with large share of intermediaries (food, apparel, textiles, metals, machinery, wood, and miscellaneous) are dropped

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
CHAPTER C. APPENDICES TO CHAPTER 3

Table C.4: IME regression, small firms excluded, extended sample

| Panel a: country pairs with $N_{ij} \geq 100$ |  |  |
| IM elasticity | Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$ |  |  |
|   | 0.437*** | 0.450*** | 0.379*** |
| Standard error | [0.0037] | [0.0027] | [0.0034] |
| $R^2$ | 0.55 | 0.75 | 0.83 |
| Variation in $\ln X_{ij}$ explained by FE, % | 0.01 | 0.13 | 0.55 |
| Observations | 14,216 | 14,196 | 13,858 |

| Panel b: all country pairs |  |  |
| IM elasticity | 0.431*** | 0.475*** | 0.512*** |
| Standard error | [0.0015] | [0.0015] | [0.0016] |
| $R^2$ | 0.69 | 0.77 | 0.80 |
| Variation in $\ln X_{ij}$ explained by FE, % | 0.00 | 0.21 | 0.56 |
| Observations | 52,593 | 52,593 | 52,447 |

Note: robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5: IME regressions, IV

| Panel a: country pairs with $N_{ij} \geq 100$ |  |  |  |
| IV lag | IV lead | IV lag and lead |
| IM elasticity | 0.392*** | 0.392*** | 0.399*** |
| Standard error | [0.0056] | [0.0057] | [0.0061] |
| Observations | 6,372 | 6,181 | 5,224 |

| Panel b: all country pairs |  |  |  |
| IV lag | IV lead | IV lag and lead |
| IM elasticity | 0.476*** | 0.479*** | 0.468*** |
| Standard error | [0.0018] | [0.0022] | [0.0024] |
| Observations | 36,065 | 36,065 | 28,672 |

| Year FE | Yes |
| Origin × year FE | Yes | Yes |
| Destination × year FE | Yes | Yes |

Note: robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table C.6: IME regression, disaggregated by HS2 product, extended sample

<table>
<thead>
<tr>
<th>Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$</th>
<th>IM elasticity</th>
<th>Standard error</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.646***</td>
<td>[0.0020]</td>
<td>58,609</td>
</tr>
<tr>
<td></td>
<td>0.598***</td>
<td>[0.0016]</td>
<td>56,560</td>
</tr>
<tr>
<td></td>
<td>0.518***</td>
<td>[0.0029]</td>
<td>29,906</td>
</tr>
<tr>
<td>Year $\times$ HS FE</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Origin $\times$ Year $\times$ HS FE</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Destination $\times$ Year $\times$ HS FE</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: $N_{ij} > 100$
Robust standard errors in brackets
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.7: IME regression, dropping industries with large share of intermediaries

<table>
<thead>
<tr>
<th>Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$</th>
<th>Panel a: core sample</th>
<th>Panel a: extended sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IM elasticity</td>
<td>Standard error</td>
</tr>
<tr>
<td></td>
<td>0.567***</td>
<td>[0.0032]</td>
</tr>
<tr>
<td></td>
<td>0.504***</td>
<td>[0.0028]</td>
</tr>
<tr>
<td></td>
<td>0.471***</td>
<td>[0.0085]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IH elasticity</td>
<td>Standard error</td>
</tr>
<tr>
<td></td>
<td>0.682***</td>
<td>[0.0027]</td>
</tr>
<tr>
<td></td>
<td>0.626***</td>
<td>[0.0025]</td>
</tr>
<tr>
<td></td>
<td>0.550***</td>
<td>[0.0046]</td>
</tr>
<tr>
<td>Year $\times$ HS FE</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Origin $\times$ HS $\times$ year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Destination $\times$ HS $\times$ year FE</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

Note: HS2 industries with large share of intermediaries (food, apparel, textiles, metals, machinery, wood, and miscellaneous) are dropped
Robust standard errors in brackets
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Table C.8: Estimates of $\bar{\theta}$

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\theta}$</th>
<th>s. e.</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>All observations, no weights</td>
<td>18.61***</td>
<td>[0.787]</td>
<td>39,712</td>
</tr>
<tr>
<td>Weights $\sqrt{N_{ij}}$</td>
<td>4.481***</td>
<td>[0.0360]</td>
<td>39,712</td>
</tr>
<tr>
<td>Dropping $N_{ij} &lt; 100$</td>
<td>2.657***</td>
<td>[0.0175]</td>
<td>7,781</td>
</tr>
<tr>
<td>Dropping $M_{ij} &lt; 100$</td>
<td>2.360***</td>
<td>[0.0147]</td>
<td>5,267</td>
</tr>
</tbody>
</table>

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

$N_{ij}$ denotes the number of exporters from $i$ to $j$

$M_{ij}$ denotes the number of exporters from $i$ to $j$ that also export to $i$’s largest destination

Table C.9: QQ estimates of $\bar{\sigma}_\varphi$

<table>
<thead>
<tr>
<th></th>
<th>All firms</th>
<th>Top 50%</th>
<th>Top 25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\sigma}_\varphi$</td>
<td>6.829***</td>
<td>4.676***</td>
<td>4.020***</td>
</tr>
<tr>
<td></td>
<td>[0.0010]</td>
<td>[0.0006]</td>
<td>[0.0008]</td>
</tr>
<tr>
<td>Observations</td>
<td>11,902,823</td>
<td>5,917,685</td>
<td>2,949,514</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.81</td>
<td>0.93</td>
<td>0.94</td>
</tr>
<tr>
<td>Bilateral FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Bibliography


Arkolakis, Costas, Natalia Ramondo, et al. (2014). “Innovation and Production in the Global Economy”. In:


Bas, Maria, Thierry Mayer, and Mathias Thoenig (2015). “From Micro to Macro: Demand, Supply, and Heterogeneity in the Trade Elasticity”. In:

Bastos, Paulo, Joana Silva, and Eric Verhoogen (2016). “Export Destinations and Input Prices”.


Brambilla Irene, Daniel Lederman and Guido Porto (2016). “Exporting Firms and the Demand for Skilled Tasks”.
Chan, Jackie ML (2017). “Financial Frictions and Trade Intermediation: Theory and Evidence”. In:
Costinot, Arnaud, Andrés Rodríguez-Clare, and Iván Werning (2016). “Micro to macro: Optimal trade policy with firm heterogeneity”.


Feenstra, Robert C (2014). “Restoring the product variety and pro-competitive gains from trade with heterogeneous firms and bounded productivity”. In:


