Essays on Firms in Developing Countries

Matthieu Teachout

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

ABSTRACT

Essays on Firms in Developing Countries

Matthieu Teachout

In this dissertation, I study the behavior and the factors that impact the performance of firms in developing countries. Chapter 1 and 3 investigate the determinants of patterns of trade in Myanmar, a country which over the past decade has been undergoing an extraordinary transition, from military control and diplomatic isolation to political and economic liberalization. Chapter 2 studies how firms upgrade the quality of their output to increase sales abroad.

Specifically, in Chapter 1, I investigate the hypothesis that if matching frictions in international trade are important, a seller's ability to connect with buyers could explain a substantial part of exporters heterogeneity in size. I do so in Myanmar's bean export market. Despite beans having all the attributes of a commodity, there is significant transaction price dispersion across both exporters and foreign buyers. Empirical patterns are consistent with foreign buyers facing search costs to find exporters. I estimate a model of search and auctions, where foreign buyers first search for a set of exporters, and then run a competitive bidding process between exporters within that set. In the model, exporters are described by two parameters: a visibility parameter that impacts their likelihood of being found by foreign buyers and a cost parameter that drives the level of their price quotes and thus their market share with each foreign buyer. Visibility explains an important part of the firm size distribution. On the buyer side, searching for an additional exporter has an estimated cost of about $2,000. Moving to a centralized market would lead to a five percent decrease in transaction prices.

Chapter 2 looks at the relationship between firms' output quality and their organizational structure. Using data on the production and transaction chain that makes up Peruvian fishmeal manufacturing, we establish three results. First, firms integrate

existing suppliers when the quality premium rises for exogenous reasons. Second, suppliers change their behavior to better maintain input quality when vertically integrated. Third, firms produce a higher share of high-quality output when supplier availability constraints shift them into using integrated suppliers. Overall, our results indicate that quality upgrading is an important motive for integrating suppliers facing a quantity-quality trade-off, as classical theories of the firm predict.

Chapter 3 quantifies the impact of import license liberalization in Myanmar's unique political economy environment. By contrast to previous literature on the issue, we find that liberalization did not lead to substantial entry in the sectors populated by firms connected to the party in power. We document two facts that rationalize these findings. First, connected firms tend to import products subject to important economies of scale, which provide opportunities for rent-seeking and act as a "natural" barrier to entry for small firms. Second, we show that a subset of the products liberalized de jure were not liberalized de facto. Products not liberalized de facto are more likely to be sectors where connected firms are present and where economies of scale are less important. This last result suggests that institutional arrangements were made to protect connected firms in the sectors where they faced higher potential competition.

Contents

List of Tables v

List of Figures vii

Acknowledgements viii

Chapter 1 Buyer Search and the Determinants of Exporter Size: Evidence from the Bean Sector in Myanmar 1

1.1 Introduction 2

1.2 Background 7

1.2.1 The bean export sector 7

1.2.2 Data 10

1.3 Descriptive results 11

1.3.1 Price dispersion 11

1.3.2 Evidence of search 12

1.3.3 Evidence of auctions 13

1.3.4 Buyer-exporter relationships 14

1.4 Model 15

1.4.1 Search 15

1.4.2 Auctions 17

1.5 Estimation 18

1.5.1 Estimation of buyers sets $N_{j,s}$ 18
1.5.2 Estimation of visibility γ_i 19
1.5.3 Estimation of marginal costs $mc_{i,s}$ 20
1.5.4 Recovering search costs c_j 22
1.6 Results 22
1.7 Counterfactuals 25
1.7.1 The sources of exporter heterogeneity 25
1.7.2 Lower search costs and centralized market 26
1.8 Conclusion 28

Chapter 2 Vertical Integration, Supplier Behavior, and Quality Upgrading among Exporters 41

2.1 Introduction 42
2.2 Background on Peru's Fishmeal Manufacturing Sector 52
2.2.1 Sector profile 52
2.2.2 Product differentiation and quality 53
2.2.3 Organizational structure 54
2.2.4 Contracting and supplier incentives 56
2.3 Theoretical Framework 57
2.3.1 Description 57
2.3.2 Model details 59
2.3.3 Discussion 61
2.4 Data, Variables, and the Relationship of Interest 63
2.4.1 Data 63
2.4.2 Variables of interest 65
2.4.3 Relationship of interest 67
2.5 The Quality Premium and Organizational Structure 69

2.5.1 Estimating how the quality price premium affects vertical integration 69
2.5.2 Interpretation 74
2.6 Firms' Organizational Structure and Supplier Behavior 76
2.6.1 Estimating how vertical integration affects suppliers' quality-enhancing actions 76
2.6.2 Interpretation 79
2.7 Vertical Integration and Output Quality 80
2.7.1 Estimating how vertical integration affects output quality 81
2.7.2 Interpretation 84
2.8 Conclusion 85

Chapter 3 Import Liberalization and Political Connections: Evidence from Myanmar 97
3.1 Introduction 98
3.2 Background and data 102
3.2.1 Historical Background 102
3.2.2 Measures of connectedness 105
3.3 Import license liberalization: implementation and impact 108
3.3.1 *De facto* vs. *de Jure* liberalization 108
3.3.2 The effect of import liberalization 109
3.4 Connected firms and economies of scale 112
3.4.1 Patterns of entry and presence of connected firms 112
3.4.2 Measuring economies of scales 114
3.4.3 Connected firms, economies of scale and *de facto* liberalization . . . 116
3.5 Conclusion 118

Bibliography 128

Appendix A Appendix to Chapter 1 **141**

A.1 Estimation procedure in details 141

A.2 Additional Figures 143

Appendix B Appendix to Chapter 2 **147**

B.1 Additional Tables and Figures 147

B.2 Dynamic Theoretical Framework and Relational Contracts 159

B.3 How Demand for Output Quality Affects Vertical Integration 171

B.3.1 Empirical strategy 171

B.3.2 Results 173

Appendix C Appendix to Chapter 3 **176**

C.1 Additional Tables and Figures 176

List of Tables

1.1 Descriptive statistics . . . 29
1.2 Evidence of search . . . 30
1.3 Evidence of auctions . . . 31
1.4 Exporter-buyer matches and divorces . . . 32
1.5 Estimated parameters . . . 33
1.6 Analysis of estimated $mc_{i,s}$ . . . 34
1.7 Search costs . . . 35

2.1 Summary statistics . . . 88
2.2 Summary statistics on Integration . . . 89
2.3 Output quality and vertically integrated suppliers . . . 90
2.4 Vertical integration and the Quality price premium . . . 91
2.5 Supplier behavior and vertical integration . . . 92
2.6 Output quality and share of inputs from vertically integrated suppliers . 93

3.1 Connected and Not Connected Importers . . . 120
3.2 Import license liberalization and connected sectors . . . 121
3.3 Decomposition of the effect of import license liberalization . . . 122
3.4 Connected sectors and entry . . . 123
3.5 *De facto* liberalization . . . 124

B.1 Main importers of Peruvian fishmeal and Average Quality Imported . . . 148
B.2 Vertical integration and the Quality price premium – First Stage . . . 149
B.3 Vertical integration and the Quality price premium – Robustness Checks . 150

B.4 Supplier characteristics 151
B.5 Output Quality and Share of Inputs from Vertically Integrated Suppliers - Robustness checks using a major firm in the Peruvian fishmeal industry. 152
B.6 Output Quality and Share of Inputs from Vertically Integrated Suppliers - First Stage 153
B.7 Countries' output quality and vertical integration in export manufacturing 154
B.8 Supplier Behavior and relational outsourcing 168
B.9 Output quality and share of inputs from vertically integrated suppliers and suppliers under a relational outsourcing contract 169
B.10 Supplier Behavior, Vertical Integration and Output Quality 170
B.11 Vertically Integrated Share of Inputs and Output Quality - Instrumenting with Firm-specific Demand Shocks 175

C.1 HS-Chapter Import Share by Connectedness Status 176
C.2 Robustness check - Import license liberalization and connected sectors . . 177
C.3 Shipment size and Number of transactions 178

List of Figures

1.1 Distribution of price residuals and Exporter and Buyer FEs 36
1.2 Simulated densities 37
1.3 Correlation between mc_i and survey data 38
1.4 Correlation between γ_i and survey data 39
1.5 Counterfactual distributions 40

2.1 Interactions between manufacturers and integrated and independent suppliers 94
2.2 Average Fishmeal Price and Quality Premium in Peru 95
2.3 Countries' output quality and vertical integration in export manufacturing 96

3.1 Distribution of the presence of connected firms across sectors 125
3.2 Share of products liberalized by HS chapter 126
3.3 Economies of scale 127

A.1 Quantity bins fixed effects 143
A.2 Prices over time within a buyer-exporter relationship 144
A.3 Distribution of order sizes for small and large buyers 145
A.4 Distribution of profits 146

B.1 Average output quality and firm size 155
B.2 Density of inventories 156
B.3 Evolution of the Vertically Integrated share of inputs Industry-wide . . . 157
B.4 Fishmeal prices in Peru and other countries 158

Acknowledgements

I would like to thank my dissertation committee, fellow PhD students, faculty and staff at Columbia University for their support during the PhD program.

I am particularly grateful to Jonas Hjort and Amit Khandelwal for their guidance. Jonas offered my a position as a Research Assistant which led to what is now Chapter 2 in this thesis. Amit introduced me to the International Growth Centre (IGC) in the second year of the program and together we wrote a Policy Note for the government of Myanmar. This initial work led me to conduct three research projects in Myanmar, of which two are chapters of this thesis. I was also lucky to work with Rocco Macchiavello, from the London School of Economics, on two of my projects in Myanmar. Working with Amit, Jonas and Rocco was undoubtedly a very enriching experience, and so not just for the PhD program. I also benefited from conversations and comments from other faculty at Columbia University. In particular, I would like to thank Reka Juhasz, Suresh Naidu, Andrea Prat, Tobias Salz, Joseph Stiglitz and Eric Verhoogen.

I am also indebted to development and trade colloquium participants for their helpful comments and suggestions. Support from classmates is very precious while during research. Golvine de Rochambeau, Sandesh Dhungana, Christopher Hansman, Nandita Krishnaswamy, Sun Kyoung Lee, Anh Nguyen, Valentin Somma and Danyan Zha contributed to making the PhD program a friendly place.

This thesis would not have been possible without backing from the IGC, which funded the projects behind Chapter 1 and 3. I am very appreciative of the support of Ian Porter, Tim Dobermann, Cormac Mangan and Siddhartha Basu who were or are working for the IGC office in Myanmar. I am also grateful to Sut Doi Aung for his work running the surveys for Chapter 1 of this thesis.

Finally, I would like to thank my wife, Violaine Pierre, who was always there, through the ups and downs of the program, and whose encouragements have been invaluable.

Chapter 1

Buyer Search and the Determinants of Exporter Size: Evidence from the Bean Sector in Myanmar

1.1 Introduction

Most export markets around the world are decentralized. Foreign buyers need to search for exporters to import goods from a country. This costly search process benefits sellers by allowing them to charge higher prices. Moreover, if buyers can only afford to look for a subset of available suppliers, a firm's ability to be found easily may be an important source of differences in exporter performance.

In this chapter, I test the hypothesis that a firm's capacity to capture foreign demand — which I term its *visibility* — contributes to its place in the distribution of exporter size. This conjecture is inspired by recent developments in the trade literature where firm-level demand is quantitatively important in explaining a firm's success (Hottman et al., 2016; Redding and Weinstein, 2018). Differences in firm level demand are typically thought to be driven by vertical or horizontal differentiation. However, conditional on the attributes of the goods produced or exported, firms may be able to capture a higher share of total demand if they are better at building and maintaining business relationships with potential buyers. In differentiated good markets, it is impossible to distinguish whether higher firm level demand is due to consumers' taste for a firm's product or to the marketing capacity of the firm. Thus, testing my hypothesis requires a setting in which firms do not differ in the set or attributes of products they sell.

Myanmar's bean export market is an ideal setting in which to consider the role of visibility for at least four reasons. First, the market is decentralized, there is no central location where buyers can meet exporters and ask for price quotes for a specific shipment. In order to import beans, buyers first need to identify local exporters. Second, beans are vertically differentiated but the quality of the good is determined upstream and is measurable. Exporters are simply intermediaries that purchase the beans from farmers or local traders and have very limited influence over the quality of their output product. The capacity needed to process and prepare orders is the same across all types of beans so all exporters have the same product scope. Third, there is significant variation in exporter size in this sector[1]. Fourth, there are a large number of buyers and exporters in the market and high turnover, meaning buyers must regularly search for new exporters and exporters constantly need to build new relationships with buyers.

I begin by documenting significant residual price dispersion using transaction level customs data. While measurement error undoubtedly accounts for a large portion of the variation in residual prices, the significant price differences across exporters and buyers cannot entirely be explained by *iid* draws of reporting errors. Several empirical patterns indicate that search frictions play an important role in generating this dispersion. First, entrant foreign buyers pay higher prices than established buyers for the same type of bean purchased in the same week. This difference fades out as buyers accumulate more experience. Second — as predicted by a model of sourcing through auction — buyers who purchase beans from a wider set of exporters pay lower prices in a given season. In a given week, for the same type of bean, exporters charge higher prices to buyers who have built fewer relationships with exporters.

Based on qualitative interviews with buyers and exporters in this sector and the empirical findings presented above, I model the market as follows. Before the season starts, buyers search for and choose a set of exporters for the season. Searching for one more exporter is costly but decreases the average price paid for orders. The reason is that for each shipment, buyers run a competitive bidding process between all the exporters they have built a relationship with, pre-season. A larger set of sellers to source from thus gives lower prices in expectation. On the other side of the market, exporters are characterized by two parameters: a *visibility* parameter that determines their likelihood of being found by foreign buyers pre-season and a *marginal cost* parameter that impacts the level of their price quotes to buyers and hence their capacity to win the bidding process. Visibility captures any marketing capacity or any other characteristics that buyers might value in exporters that are unrelated to the product itself. Empirical analysis reveals that an exporter's likelihood to be sampled by buyers is not correlated with any

[1] In fact, the distribution is very similar to the distribution of exporters in differentiated good sectors e.g. garments

observable characteristic. The cost parameter is more standard and reflects productivity or any supply-side advantage exporters may have in this market.

I estimate the model in two stages. First, the visibility parameters are obtained for each exporter from the new relationships that are built over time in my sample. I then estimate exporters' marginal costs, using methods inspired by recent developments in the empirical auction literature. An exporter's cost parameter impacts his likelihood of winning a given auction with a buyer as well as the average transaction price with that buyer. Finally, I recover search costs from the trade-off buyers face. In equilibrium, the marginal cost of searching should equal the price benefits of running auctions with one more exporter. Given the estimated distribution of exporter costs, I can compute the marginal benefit of sourcing from one additional seller and hence buyers' search costs.

Results are as follows. Firms vary in both their visibility and cost parameters but the two are orthogonal. Marginal cost parameters are time persistent and are correlated across bean types. I use survey data collected from exporters to look for predictors of visibility and costs. Exporters who purchase directly from farmers (rather than solely from local traders) or trade in the domestic market (as opposed to trading solely in the export market) have lower marginal costs. Exporters with higher visibility typically have made investments consistent with trying to get more attention from buyers: they are more likely to be found online and are more likely to have an office in downtown Yangon, Myanmar's economic capital. The median search cost of finding one exporter is $2,588 for entrant buyers and $1,725 for continuing buyers, in line with the burden faced by buyers who have to initiate a relationship for the first time: contacting an exporter, meeting, and then conducting due diligence.

With my estimates in hand, I can perform several counterfactual analyses to quantify the sources of exporter heterogeneity. I simulate the model under two scenarios. If exporters had equal marginal cost parameters, the standard deviation of the logarithm of firm size would be 13 percent lower. On the other hand, if all exporters had the same visibility, the same number would fall by 30 percent, highlighting the importance of relationship building in explaining exporter performance. The majority of the variation though is random from the exporters' perspective. It comes from the distribution of quantity purchased by the buyers an exporter connects with.

An important consideration is the external validity of these results. The magnitude of search costs is quite high, but as interviews with buyers revealed, a large fraction of these costs can be attributed to vetting and auditing exporters before doing business. As such, the poor contracting environment in Myanmar is likely responsible for the search process [2] buyers have to go through and so search could be less important in other exporting countries. However, in differentiated good markets, the auditing process buyers need to go through is likely to be even more costly as it typically includes discovering the precise characteristics of the goods a firm can produce[3]. Similarly, in differentiated products, the marketing capacity of firms is likely to play an even more important role in shifting firm-level demand.

A natural question emerging from this chapter is why a decentralized market subsits in this sector. Another counterfactual exercise offers insight. If buyers were able to connect and get price quotes from all available exporters, prices would be on average 5.3 percent lower. A centralized market would be equivalent to a surplus transfer from exporters to buyers. As the number of competitors on any given auction increases, exporters submit lower price bids and thus charge lower markups. Low cost and low visibility exporters would benefit from moving to a centralized market but model estimates suggest that the majority of exporters would make lower profits under a common exchange market. Implementing such a platform could be initiated from the buyer side but high turnover and limited coordination are likely to make such a plan infeasible.

This chapter relates to several distinct literatures. First, it highlights the importance of firm-level demand in explaining exporter performance. Hottman et al. (2016) and Redding and Weinstein (2018) similarly show the importance of demand idiosyncrasies

[2] Kranton and Swamy (2008) also study contracting problems in export markets where the goods are likewise supplied through an auction mechanism.

[3] See e.g. Startz (2016) for an empirical investigation and Anderson and Renault (1999) or Kuksov (2004) for models of buyer search with heterogenous products.

in explaining firm sales and trade patterns, but in their setting, variation in demand is attributed to horizontal or vertical differentiation in the products firms sell. While this chapter investigates the importance of a seller's *visibility* in the static distribution of sales, recent papers have shown that exogenous shocks to firm-level demand can have long-lasting effects both on productivity (Atkin et al., 2017b) and the likelihood to capture a higher share of demand in the future (Ferraz et al., 2015).

Second, this paper contributes to the literature on search in decentralized markets. Work in this area was led by Stigler (1961), who explores the pricing behavior of firms when consumers need to search for the lowest price in decentralized markets[4]. The goal of the empirical literature in this area is to back out buyer search costs from price and quantity data (see Hortacsu and Syverson (2004) on the mutual fund industry) or from price data alone (Hong and Shum, 2006; Giulietti et al., 2014).

In the trade literature, Allen (2014) shows how information frictions generates price dispersion in commodities across local markets whereas this paper explores how buyer search can create price dispersion across exporters within a given market. This study also relates to the literature on the matching process between exporters and buyers using customs data (Blum et al., 2010, 2011; Benguria, 2015; Kamal and Sundaram, 2016; Yoichi et al., 2016; Bernard et al., 2018; Cajal-Grossi et al., 2019). Eslava et al. (2015), using customs data from Colombia, estimate the search costs of identifying potential clients and the costs of maintaining business relationships with existing buyers[5].

Finally, this paper uses recent developments in the auction literature to identify sellers' costs from observed transaction prices and quantities (Salz, 2018; Cuesta and Sepulveda, 2018). The choice of using an auction model not only maps into the mechanism that buyers actually follow in the sector studied, but also allows to overcome the prob-

[4]While this paper assumes that search costs are constant, Stiglitz (1987) shows that when search costs are increasing with the number of previous searches, more firms in the market does not necassirily lead to more competition and thus lower prices

[5]The estimated search costs are convex in buyer arrival hazards, from $1,405 per year for an expected yield of one potential client every two years but rising to $51,471 for an expected yield of one potential client per year. In this paper, buyers search for exporters and search cost estimates are much lower.

lem that the full range of price quotes a buyer receives, which is derived from the cost distribution, is unobservable. Most papers in this field use a result from Athey and Haile (2002) to identify non-parametrically the distribution of costs from transaction prices and the identity of the auction winner. In this chapter, I follow the same intuition but I impose structure on the distribution of sellers' costs over time. This allows me to identify a cost parameter for *each* firm rather than the *distribution* of costs across all exporters.

The remainder of the chapter is organized as follows. Section 1.2 provides relevant facts about the sector and describes the data. Section 1.3 establishes empirical facts that are consistent with a model of search and auction. Section 1.4 describes the model, and Section 1.5 the estimation. Section 1.6 describes the results and Section 1.87 counterfactual analyses.

1.2 Background

1.2.1 The bean export sector

Several types of beans are produced in Myanmar. This chapter focuses on the three main bean types that dominate the export market: black matpe (or urad, 53 percent), toor whole (or pigeon peas, 21 percent) and mung beans (or green grams, 17 percent). Exports in other types of beans (black eyed beans, kidney beans, bamboo beans...) account for less than 9 percent of the total quantities exported. Black matpe and mung beans are also consumed locally but in very small quantities compared to the amounts exported. Production is seasonal but exports are usually more spread out during the year as the beans can be stored up to 18 months with a sales peak from April to September.

In total, beans account for approximately seven percent of the country's total exports and about 10 percent of the non mineral resources exports. India is by far the main destination of Myanmar beans exports (71 percent) as it is also the main consumer of that ingredient in the world. The Mumbai commodity market is the world reference price for these beans.

Exporters are all intermediaries, none of them produce the beans themselves. They buy the good from domestic traders or farmers, and store them in their warehouse, usually in the suburbs of Yangon, Myanmar's economic capital. There, they clean the beans (remove dirt), repackage, store and ship the goods to the port of Yangon. Some exporters are also involved in other activities. Survey data collected with a sample of exporters revealed that approximately 63 percent of them are also involved in domestic trading activities and 70 percent in import activities.

Foreign buyers are generally intermediaries who conduct search and ship the beans to their final destination on behalf of a final buyer (most of the time based in India). The final buyer sends an order for a certain number of containers of a bean type at a price (generally around the current Mumbai market price) and the foreign buyer conducts search in Myanmar to find the lowest possible price and hence maximize profits on the order. The foreign buyer or intermediary takes ownership of the beans at the port of Yangon and is in charge of shipping the good to the destination country. These buyers usually operate from Singapore but some of them also have a local office in Yangon. Most of them also trade other commodities in all of South East Asia.

Interviews with these foreign buyers provided insight into how they operate. Buyers build relationships with exporters before the season starts. Initiating a relationship typically involves contacting the exporter, meeting with him several times, conducting an audit of the exporter's past records and visiting the warehouse. This vetting process has to be done for two reasons. First, once the season starts, buyers must be able to sign contracts with exporters and ship the goods in a few days. They would not have time to conduct search and the vetting process for every new order. Second, once the exporter and the buyer agree on a price, they sign a contract and the exporter usually has a few days to prepare and ship the order. Yet, half of the order payment is typically done at the time when the contract is signed and the other half upon delivery of the order at the port. The value of these shipments being relatively large (around $100,000), buyers need to make sure that exporters have built some form of reputation and would not run off with half the payment upon signing the contract. All buyers interviewed for this project referred to such events having happened in the past. As such, Myanmar's poor contracting environment contributes significantly to buyers' search costs. Interviews with buyers and exporters revealed that even when a relationship has been going on for several years, both parties meet at least once a year to maintain the relationship. Over the course of the season, upon receiving a request from a final buyer, foreign buyers call each of the exporters they have built a relationship with to ask them for a price quote for the quantity and the type of beans they need. Buyers then sign a contract with the exporter who offers the lowest price. When asked why they do not buy beans from the same exporter for every order, all interviewed buyers responded that an exporter who offers the lowest price on a given week might not offer the lowest price the following week, hinting at variation in exporters' marginal costs over the course of a season. This flucuation in exporters' costs is likely to be due to inventory management. As the local price fluctuates, exporters who bought their stock at different points in time face different marginal costs when contacted by buyers.

Beans can be of two qualities labeled as Fair Average Quality (FAQ) or Superior Quality (SQ). These quality labels refer to the size of the beans which is determined by the seeds used by farmers or the conditions under which the beans are grown. Critically, exporters have no influence over the quality of the beans. As quality can be perfectly measured, it is specified in the contract and buyers typically send an inspector to check the quality of the shipment in the exporter's warehouse before it is sent to the port. More importantly for this paper, the quality of beans is reported in the customs data and can be accounted for when measuring price dispersion. Another dimension in which quality could matter is the amount of dirt mixed with the beans which can hardly be measured and hence cannot be specified in a contract. The quality inspectors sent by the buyers examine the beans and if they are not cleaned enough, ask that they be processed in the warehouse once more before being shipped. Buyers interviewed for this project declared that such a case happens less than 5 percent of the time and there is no significant difference across exporters in how likely the quality inspector would ask the exporter to reprocess the goods. Thus, quality is unlikely to be a determinant of vertical differentiation across exporters.

Buyers interviewed for this project listed trust and how easy it is to do business with an exporter as the most important determinants that make them choose some suppliers over others. Importantly, they did not list any product characteristics such as vertical or horizontal differentiation to be important.

1.2.2 Data

This chapter uses Myanmar export customs data which was provided by the Ministry of Commerce there. It is available at the transaction level between April 2011 and March 2017. The data provides the name of the exporter, the type of beans, the quality of the beans (see above), the quantity and value of the shipment, the date at which the shipment is cleared by customs, the name of the foreign buyer and the destination to where the beans will be shipped. The data contains around 38,000 transactions.

Table 2.1 provides summary statistics. There are on average of 228 distinct buyers and 168 exporters in the market on a given season. Buyers make on average 24 orders per season. 41 percent of buyers on a given season are entrant and 23 percent of exporters are new in the market, so turnover on both sides of the market is high, reflecting the need for buyers to constantly keep searching for exporters. On average, buyers purchase from 3.4 distinct exporters on a given season.

In addition to this data, I collected in the first semester of 2018, surveys from firms sampled from the top 100 exporters of beans in the last year available in the customs data. The survey gathered information about how exporters build relationships with buyers as well as how they source and process the beans (capacity, location of warehouse...). Unfortunately, reaching out to these exporters was extremely difficult and only 35 exporters were surveyed in the end. However, these interviews provided insight into how the market operates to build the model presented below. The few variables collected also allow to check for correlations between the model estimates and observable characteristics of exporters.

1.3 Descriptive results

1.3.1 Price dispersion

I first document residual price dispersion which is inconsistent with a model of perfect competition across exporters. Over the sample period, export prices fluctuate a lot due to variations in demand in destination countries and supply conditions in Myanmar and other producer countries. Thus, I analyze price dispersion from price residuals rather than just prices. I construct price residuals from the following regression:

$$p_{ijpqw} = \alpha_{pqw} + \delta_k \cdot \mathbb{1}(qty = k) + \widetilde{p}_{ijpqw} \tag{1}$$

where i in an exporter, j a foreign buyer, p a bean type, q a given quality, w is the week during which the transaction happens, p_{ijpqw} is the transaction unit price and δ_k are quantity bins fixed effects. Price residuals $\widetilde{p}_{ijpqw}$ are devations from product-quality-week fixed effects. Across the sample, there are an average of 29 transactions for each bean-quality pairs on a given week so the fixed effect should reflect the trend in the average price. I also include quantity bin fixed effects as there are consistent price differences across transactions of different sizes reflecting the fixed costs (clearing customs, transport) associated to delivering the order to the port (see Appendix Figure A.1). For the rest of the chapter, I work exclusively with the constructed price residuals and refer to these residuals when using the term transaction price. In particular, residual prices can be negative.

The distribution of residual prices is presented in Panel a of Figure 1.1. The standard deviation is $73 with the average transaction price around $800, reflecting significant price dispersion within a week. Of course, a large fraction of this variation is due to measurement error in the customs data. Quantities and values from which the transaction unit price is computed are likely to be misreported by customs agents. Nonetheless, I can rule out that all of this variation in price residuals is entirely due to measurement error. If within a week, transaction prices were equal but only differed in the data by an *iid* error term, the distribution of error term should be equal to the distribution of the price residuals plotted in panel a of Figure 1.1. Thus, I can construct a 95 percent confidence interval for exporter (or buyer) price fixed effects from the empirical distribution of price residuals and compare where the actual price fixed effect stands compared to the confidence interval. I plot the confidence interval and the residual price fixed effect for the largest exporter-season pairs in Panel b of Figure 1.1 and for the largest buyer-season pairs in Panel c. The price fixed effects are outside the 95 percent confidence intervals for more than 5 percent of the cases and for some exporter-season pairs, the price residual fixed effects are really far from the boudaries of the confidence interval. In other words, there is too much dispersion in exporter-season and buyer-season price residual fixed effects for the variation in price residuals to be entirely explained by *idd* error terms. The measurement error in the customs data could be firm specific if e.g. some exporters consistently over-report quantities or values to customs. Another potential explanation could be that there are important switching costs which allow exporters who have been in a relationship with buyers for a long time to charge higher prices. If this hypothesis were true, the average price an exporter charges to a given buyer should increase with the length of the relationship. Appendix Figure A.2 rules out the presence of switching costs as within a relationship, prices are constant over time. Moreover, the next set of results suggest that part of the price residual dispersion can be attributed to buyers facing costly search to find and build a relationship with exporters.

1.3.2 Evidence of search

In Panel A of Table 1.2, I report evidence that new buyers, for whom search costs should be relatively higher, pay higher prices. Column 1 shows that entrant buyers in their first season pay on average $13 (1.6 percent) higher prices than other buyers. But as column 2 reveals, this difference fades over time. For every additional year of experience, entrant buyers pay on average $2 (0.3 percent) lower prices.

From the list of buyers in the customs data, I collected information about them. Column 3 of Panel A of Table 1.2 shows that buyers who have an office in Myanmar, and thus are likely to face lower search costs than buyers operating from abroad, pay on average $15 (1.8 percent) lower prices.

The auction mechanism which buyers go through for every order as described in section 1.2 suggests that buyers who purchase beans from a wider set of exporters should pay lower prices. The distribution of the minimum bid shifts to the left as more exporters are included in the competitive bidding process. Panel B of Table 1.2 explores this hypothesis. Column 1 shows in a cross-section that foreign buyers who purchase from one more exporter over the course of a season pay on average $1.9 lower prices. Column 2 studies the same correlation but within buyers over time. When buyers add one more exporter to the set of firms they purchase from, transaction prices decrease by an average of $1.4. Finally, column 3 explores how prices vary within the set of sales an exporter does within a week in a particular type of good. For each additional exporter included in the supply set of a given buyer, exporters charge on average $2.5 higher prices. Even if the coefficient is not significant in column 3, these results clearly show that buyers who purchase from a wider set of exporters get lower prices. This is consistent with the assumption that buyers face a trade-off when searching for more exporters. One the one hand, they need to pay a cost to search for more sellers. On the other hand, including more exporters in their buying set allows them to source the beans at lower prices on average over the course of the season. This relationship is key to estimating search costs in the structural exercise that follows.

1.3.3 Evidence of auctions

Customs data shows evidence of the auction mechanism described by buyers which they use to find the lowest price available for each order. A classic auction model predicts that exporters who consistently offer lower prices to buyers should get a higher market share. As the price data is particularly noisy, I do not observe a clear monotonic relationship between exporter price fixed effects and market shares but the two are clearly negatively correlated.

Column 1 of Table 1.3 shows that exporters with low price fixed effects have a higher market share in the sector as a whole. Within each buyer, the negative correlation is even stronger. If buyers were including all exporters available in their auctions, the slope of the coefficients in column 1 should be similar to the one in column 2 and 3. Thus, results in this table are consistent with buyers using auctions to source the beans and that these auctions do not include all available exporters in the market.

1.3.4 Buyer-exporter relationships

Decomposing total quantity sold for each exporter into the number of buyers he sells to and the mean quantity sold to each one of these buyers reveals that about 40 percent of the variation in the distribution of exporter size can be attributed to the number of buyers they sell to. This highlights the importance of building connections with buyers for explaining an exporter's success.

At this stage, an important question is whether exporters who can offer lower prices to buyers can signal this before the season starts and are thus more likely to build new relationships with buyers. In Panel A of Table 1.4, I find that exporters' average prices in the current or previous season do not correlate with the number of new matches they make with buyers in the current season.

Looking at exporter-buyer divorces, Panel B of Table 1.4 shows that the length of the relationship is a factor that significantly impacts the rate of survival of exporters with buyers. However, and perhaps surprisingly, an exporter's market share or average transaction price with a given buyer does not affect its survival rate in the next season. This confirms what buyers reported during interviews. They value other characteristics than average prices when choosing which exporters they want to do business with.

Overall, these results suggest that the matching process between exporters and buyers is not driven by exporter characteristics that are observable in the customs data. Nonetheless, as the number of buyers a given exporter sells to is important for explaining firm size, I assume in the model section below that exporters vary in their probability to be sampled by buyers.

1.4 Model

1.4.1 Search

For each season s, there are a total of M_s foreign buyers indexed by j. Among those M_s buyers are $\widetilde{M}_s$ new buyers (entrant) and $(M_s - \widetilde{M}_s)$ continuing buyers from season $s-1$. Each buyer j needs to make $S_{j,s}$ orders to local exporters over the course of the season, each of size q [6].

Buyer j can only source orders from local exporters which he has searched and built a relationship with. I note $\Omega_{j,s}$ this set, $N_{j,s}$ the number of exporters included in that set and c_j buyer j's cost of searching one additional exporter. The number and identities of exporters in $\Omega_{j,s}$ is fixed for the duration of the season. In line with the sourcing procedures described by foreign buyers during interviews, I assume that for each order t, buyer j runs a first-price sealed auction between all exporters included in his set $\Omega_{j,s}$.

I first consider entrant buyers. They start with an empty set and must decide how many exporters to include in their set before the season starts. Each new buyer minimizes total costs of the orders he will make over the course of the season and the cost associated to the initial search of the $N_{j,s}$ exporters to include in his set $\Omega_{j,s}$. His objective

[6] I abstract from the distribution of order size. In reality, orders vary in size but 60% of orders are of exactly 5 containers and there is no systematic difference in the top of the distribution of order size across buyers. See Appendix Figure A.3.

function can be expressed as:

$$\min_{N_{j,s}} \qquad q \cdot S_j \cdot \mathbb{E}\left[p_t \middle| N_{j,s}\right] + c_j \cdot N_{j,s} \tag{2.1}$$

where $\mathbb{E}\left[p_t \middle| N_{j,s}\right]$ is the expected unit price obtained for each order t from running an auction with $N_{j,s}$ exporters. It is the average lowest bid submitted by $N_{j,s}$ exporters (see the next sub-section).

For a continuing buyer, I assume that among the $N_{j',s-1}$ exporters buyer j' had in his set $\Omega_{j',s-1}$ in the previous season, only $\widetilde{N}_{j',s-1} < N_{j',s-1}$ relationships can be carried forward into season s. Following the empirical analysis in section 1.3, I assume that this number is exogenous: some exporters leave the market and some relationships don't survive. In particular, the probability that a relationship ends on a given season is a function of the length of the relationship but is the same across all exporters and all buyers. I assume that the cost of maintaining relationships with existing exporters is $\mathbb{c}_{j'} < c_{j'}$ so that buyer j' continues working with all exporters for which the relationship is not exogenously broken. Continuing buyer j' must then decide the number of new exporters $\breve{N}_{j',s}$ to include in his set $\Omega_{j',s}$ for season s. Similarly to entrant buyers, buyer j''s objective function is given by:

$$\min_{\breve{N}_{j',s}} \qquad q \cdot S_{j'} \cdot \mathbb{E}\left[p_t \middle| \widetilde{N}_{j',s-1} + \breve{N}_{j',s}\right] + c_{j'} \cdot \breve{N}_{j',s} + \mathbb{c}_{j'} \cdot \widetilde{N}_{j',s-1} \tag{2.2}$$

On the other side of the market are a total of N exporters indexed by i. Exporter i is characterized by his *visibility* parameter γ_i, i's probability of being sampled in a search for one exporter among all N exporters. If all exporters have the same visibility, then $\gamma_i = \frac{1}{N}$. The probability $\mathbb{P}_{i,n}$ that i is sampled in a search for $n > 1$ exporters is a complex function of n and all the γ_i's since search is conducted without replacement. But for $n << N$, an approximation can be derived from a model of search with replacement, $\mathbb{P}_{i,n} \approx \gamma_i n$.

1.4.2 Auctions

On a given season s, exporter i is also described by his average *marginal cost* $mc_{i,s}$ which reflects i's productivity in cleaning the beans and preparing the order[7] as well as i's ability to buy beans in the local market at a low price. Across all exporters, $mc_{i,s}$ is an *iid* draw from distribution $\mathcal{F}(.)$. I assume that i's marginal costs are not fixed over time within a season. As explained in section 1.2, exporters' optimal bids are not constant, even relative to the average market price, which explains why foreign buyers need to run a new auction for every order. For each order t, i faces marginal cost $mc_{i,t}$, an *iid* draw from a continuous distribution $mc_{i,t} \sim \tilde{\mathcal{F}}(.|mc_{i,s})$. Finally, I note $\mathcal{G}(.)$ the overall (across all exporters) distribution of marginal costs in the market at any point in time.

Section 1.3 above showed that exporters offer lower prices to buyers with larger sets $\Omega_{j,s}$. Assuming that exporters know the number of exporters $N_{j,s}$ included in $\Omega_{j,s}$ but do not know the identity nor the average marginal cost parameter of the other exporters included in $\Omega_{j,s}$, the optimal bidding function is symmetric across all exporters competing to sell an order to the same buyer j [8]. I note $\beta_{j,s}(.)$ the optimal bidding strategy, which is the solution to the following maximization problem:

$$\max_{b} = (b - mc_t) \cdot \left[1 - \mathcal{G}\left(\beta_{j,s}^{-1}(b)\right)\right]^{N_{j,s}-1} \tag{3}$$

where $\left[1 - \mathcal{G}\left(\beta_{j,s}^{-1}(b)\right)\right]^{N_{j,s}-1}$ is the exporter perceived probability that his bid b will be lower than all other exporters submitting a bid for the same order. The optimal bidding strategy for a first-price sealed auction is a well-know function (see Krishna, 2009) and is given by:

[7] Preparing the order includes putting the beans into bags, loading them on trucks and delivering the orders to the port

[8] Moving away from this assumption would be significantly complicate the model. A closed form expression for the bidding strategies in first-price asymmetric auction may not be available. See Chapter 4 in Krishna (2009).

$$\beta_{j,s}(mc) = \mathbb{E}[Y_{j,s}|Y_{j,s} > mc] \tag{4}$$

where $Y_{j,s}$ is the first order statistic, i.e. the minimum, of $N_{j,s} - 1$ random draws from the cdf $\mathcal{G}$. In particular, $\beta_{j,s}$ is strictly increasing in its argument mc.

1.5 Estimation

This section describes how the set of parameters in the model $\{N_{j,s}, c_j, \gamma_i, mc_{i,s}\}$ are estimated from the data. I estimate marginal costs parameters for each type of bean separately but abstract from bean type notations in the estimation procedure described below.

Most empirical estimations of models of search and auctions estimate the overall distribution of search costs or marginal costs. I estimate a search cost for each foreign buyer and a visibility and a marginal cost parameter for each exporter. I do so to study the factors that influence visibility and costs from survey data and identify winners and loosers in counterfactual settings. However, estimating a cost parameter for each exporter poses a number of challenges as explained below.

1.5.1 Estimation of buyers sets $N_{j,s}$

I assume that $N_{j,s}$ is the number of exporters who do at least one transaction with buyer j during season s, I note this number $\widehat{N_{j,s}}$. In reality, it is possible that $N_{j,s} > \widehat{N_{j,s}}$ but that only a subset of the exporters in the set end up making at least one transaction with buyer j over the course of the season. In this case, the search costs estimated below would be an upper bound to the true search costs and the average marginal costs estimated below would be lower than their actual value. Nonetheless, I am more interested in the relative value of marginal costs across exporters than their absolute level. Moreover, exporters who are included in a buyer's set but do not end up making a transaction with that buyer on a given season are likely to be high marginal cost exporters. Thus, the estimation procedure described below is likely to underestimate the visibility parameter γ_i for high marginal cost exporters.

The auction model described above is not defined for $N_{j,s} = 1$ and a large proportion of buyers only buy from one exporter over the course of a season. If $N_{j,s} = 1$, exporter's bids would always be the same, at the buyer's reservation price. However, as long as exporters expect $N_{j,s}$ to be greater than one, the auction model is defined.

I do not find any statistical difference in prices between buyers who buy from one exporter and buyers who buy from two exporters over the course of the season. Thus, I assume that buyers who end up buying from only one exporter have actually two exporters in their set. In the data, while there are numerous buyers who only have one exporter in their set, the majority of these buyers tend to make only one order on a given season and so these buyers only account for a small share of total transactions. However, I estimate the auction model and the marginal cost parameters only for buyers with more than one exporter (see Section 1.5.3).

1.5.2 Estimation of visibility γ_i

In the data, in any buyer's set, the number of exporters is much smaller than the total number of exporters (see Table 2.1). Thus, as highlighted in the model section, I can make the assumption that the sampling of each exporter is done with replacement. I gather all buyer season pairs $r = \{j, s\}$ where a buyer samples $\widehat{N_r}$ new exporters (either entrant buyers or new relationships for continuing buyers as described in the model section above). I note $\widehat{\Omega}_r$ the realized set of sampled exporters. An estimator for γ_i is then simply given by:

$$\widehat{\gamma}_i = \sum_r \frac{\mathbb{1}\{i \in \widehat{\Omega}_r\}}{\widehat{N_r}}$$

1.5.3 Estimation of marginal costs $mc_{i,s}$

For the estimation of the marginal cost parameters, I assume that $\tilde{\mathcal{F}}(.|mc_{i,s})$ follows a normal distribution with mean $mc_{i,s}$ and variance v^2. Similarly, $\mathcal{F}$ is the cumulative distribution function of a normal with mean μ and variance σ_m^2 so that $\mathcal{G}$ is also normal with mean μ and variance $\sigma^2 = \sigma_m^2 + v^2$. Thus, the set of parameters to estimate here are $\{mc_{i,s}, \sigma_m, v\}$.

In the data, only the transaction prices are observed. I note $\mathcal{H}_{ijs}(.)$ the distribution of optimal bids submitted by i in response to a request by buyer j during season s. Given that the optimal bidding strategy $\beta_{j,s}(.)$ is strictly increasing in its argument, if $X \sim \mathcal{F}(.|mc_{i,s})$, then $Y = \beta_{j,s}(X) \sim \mathcal{H}_{ijs}(.)$. Submitting a bid b, i's probability of winning the auction, i.e. submitting the lowest bid, is the probability that all exporters in buyer j's set submit higher bids that i:

$$\mathbb{P}(i \text{ wins } j\text{'s auction } t) = \prod_{k\in\Omega_j, k\neq i} \Big[1 - \mathcal{H}_{kjs}(b)\Big]$$

The challenge of the estimation lies in translating the parameters that define $\mathcal{F}(.|mc_{i,s})$ which we want to identify into the parameters that define $\mathcal{H}_{ijs}(.)$. In Appendix A.1, using results from Chen and Tyler (1999) and Benaroya et al. (2005), I show that $\mathcal{H}_{ijs}$ can be very well approximated by a normal distribution with mean $\mu_{ijs} = m(\mu, \sigma, N_{j,s}, mc_{i,s})$ and standard deviation $\sigma_{ijs} = s(\mu, \sigma, N_{j,s}, mc_{i,s}, v)$.

Finally, as noted in section 1.3, errors in the transaction price reported in the data are likely to be very large. To account for this, I assume that the observed price in the data $\widetilde{p}_{ijt}$ is given by $\widetilde{p}_{ijt} = p_{ijt} + \epsilon_{ijt}$ with $\epsilon_{ijt} \sim \mathcal{N}(0, \sigma_e)$. Then, the log-likelihood contribution of a contract between exporter i and buyer j at an observed price $\widetilde{p}_{ijt}$ is:

$$L(\widetilde{p}_{ijt}, \theta_s) = \sum_{\substack{k\in\Omega_j \\ k\neq i}} \log\Big[1 - \widetilde{\mathcal{H}}_{kjs}(\widetilde{p}_{ijt})\Big] + \log\Big[\widetilde{h}_{ijs}(\widetilde{p}_{ijt})\Big] \tag{5}$$

where $\widetilde{\mathcal{H}}_{ijs}$ is the cdf of a normal with mean μ_{ijs} and variance $\sigma_{ijs}^2 + \sigma_e^2$ and $\theta_s =$

$\{mc_{i,s}, v, \mu, \sigma, \sigma_e\}$.

Two things prevent the simultaneous estimation of all parameters in θ using equation 4. First, it is not possible to differentiate v from σ_e as both parameters increase the variance of the observed price $\widetilde{p}_{ijt}$ but do not impact its level. Second, a constraint imposed by the model is that $\mu = \frac{1}{N}\sum_{i=1}^{N} mc_i$ and $\sigma^2 = \sigma_m^2 + v^2$ and an unconstrained estimation of equation 5 might not lead to these equalities holding with the estimated parameters.

To solve this issue, I estimate the parameters of the model in two steps:

1. I first estimate μ, σ and σ_e by simulating the model to match three moments in the data: the mean and standard deviation of the price residuals $\widetilde{p}_{ijt}$, and the coefficient in column 1 Panel B of Table 1.2 which drives how average prices change with the number of exporters included in a buyer's set. To do so, I assume that visibility and marginal costs are orthogonal, so that the expected marginal cost parameter of any sampled exporter is just μ. I check the validity of this assumption ex-post by checking the correlation between the estimated marginal costs and visbility paramaters. The key parameter to estimate at this stage is σ. Intuitively, higher variation in marginal costs not only shifts the distribution of the first-order statistics to the left but also increases the gap between the distribution of the minimum between N and N+1 draws. A higher σ thus increases the coefficient obtained in a similar regression as in column 1 Panel B of Table 1.2.

2. I follow the literature on auctions to estimate the $mc_{i,s}$ and v by maximizing the log-likelihood in equation 4, taking the estimated $\hat{\mu}, \hat{\sigma}$ and $\hat{\sigma_e}$ from step 1 [9]. Note that estimation through maximum likelihood is similar to estimating the $mc_{i,s}$ with observed transaction prices and within buyer market shares as the function in equation 5 measures both the likelihood of observing a certain price bid level for exporter i as well as the probability that i wins a given auction with that bid.

[9] For that, I first take a guess at v and estimate the $mc_{i,s}$ and then reestimate the $mc_{i,s}$ with a new guess of $v = \sqrt{\sigma^2 - SD(mc_{i,s})^2}$ until the estimated $mc_{i,s}$ and v converge.

1.5.4 Recovering search costs c_j

Having estimated the distribution of marginal costs, I can then recover search costs for each entrant buyer j from the first order condition of the maximization problem presented in equation 2.1:

$$\widehat{c}_j = -\hat{q} \cdot \hat{S}_j \cdot \frac{\partial \widehat{\mathbb{E}}\left[p_t | N_{j,s}\right]}{\partial N_{j,s}}$$

where $\frac{\partial \widehat{\mathbb{E}}[p_t | N_{j,s}]}{\partial N_{j,s}}$ can be approximated by the derivative of the optimal bidding strategy taken at the average marginal cost: $\beta(\hat{\mu}|N_{j,s}+1) - \beta(\hat{\mu}|N_{j,s})$.

Similarly, for continuing buyers, searching for $\breve{N}_{j',s}$ new exporters, searched costs can be recovered from the first order condition of the maximization problem presented in equation 2.2:

$$\widehat{c}_{j'} = -\hat{q} \cdot \hat{S}_{j'} \cdot \frac{\partial \widehat{\mathbb{E}}\left[p_t \middle| \widetilde{N}_{j',s-1} + \breve{N}_{j',s}\right]}{\partial \breve{N}_{j',s}}$$

Note that search costs can only be estimated for continuing buyers with $\breve{N}_{j',s} > 1$, that is buyer-season pairs that build at least one new relationship on a given season. In practice, only four continuing buyer-season pairs in the data do not build any new relationship with exporters on a new season.

1.6 Results

The estimated parameters are presented in Table 1.5 for each type of bean separately and for a specification with all beans together. The estimates are similar across bean types. Average marginal costs are between \$1 and \$7. Recall that the model is estimated with the transaction price residuals which have mean zero, so the level of marginal costs is expressed relative to the weekly price averages. In other words, the average marginal costs are between \$1 and \$7 higher than the average weekly transaction price. This is not surprising as variations in marginal costs make the expected value of the lowest bid lower than the average marginal cost paramater. The variation in marginal costs across exporters σ_m is about half the variation in marginal costs within exporters but across transactions v. As expected, the variance of the error term is very large and so measurement error accounts for the majority of the variation in the observed residual transaction prices $\widetilde{p}_{ijt}$. Figure 1.2 shows how the model moves from the distribution of marginal costs to the distribution of optimal bids and transaction prices. Panel a presents the distribution of marginal costs and optimal bids for N=2 and N=8, Panel b the distribution of transaction prices and Panel c the simulated and the actual distribution of residual transaction prices.

In Table 1.6, I study various correlations with the estimated exporter-level average marginal cost parameters $mc_{i,s}$. Panel A shows that the estimated marginal costs for each type of beans correlate with the exporter-season price fixed effects. This is not surprising as higher marginal costs imply higher price bids and hence higher transaction prices in the model described above. Columns 1 and 3 of Panel B show that if exporters have a high marginal cost in one bean type, they are likely to have high marginal costs in other bean types as well. Columns 2 and 4 look at the same correlations but within exporters and across time: if an exporter faces an increase in marginal costs from one season to the next in one bean type, marginal costs in other beans types are also likely to increase. Panel C looks at serial correlation of marginal costs. An exporter who has faced low marginal costs in the previous season is likely to have low marginal costs in the current season. However, this serial correlation is not perfect which might explain why buyers do not seem to end relationships with exporters based on the average transaction price they got from this exporter in the previous season. Overall, these results suggest that marginal costs are driven by an exporter supply ability. Finally, Panel D shows that the estimated marginal cost paramaters are orthogonal to the estimated visibility paramaters, confirming the validity of the assumption made in section 1.5.3. This result suggests that an exporter's efficiency to supply the goods at a low cost is unrelated to the exporter's ability to make business connections with foreign buyers.

I then investigate how the estimated marginal costs correlate with some supply-side variables from the survey data. Obviously, as the number of exporters in the sample is very small, catching a significant correlation with any of the survey variable is challenging. However, two variables stand out. Figure 1.3 shows that exporters who buy some of their beans directly from farmers (as opposed to traders) and exporters who also trade in the domestic market (as opposed to only selling on export markets) have lower marginal costs (significant at the 10 percent level). None of the self-reported estimates on the cost of cleaning the beans or the cost of shipping the goods from the warehouse to the port are significant, suggesting that exporters compete more on their ability to find cheap beans than in their ability to process and ship them.

Looking at the factors that impact an exporter's visibility, I do not find any of the variables collected in the survey (number of trips made aborad, number of people in the firm who speak different languages, number of years of experience...) to be significantly correlated with an exporter's visibility. However, looking at the variables collected to build the survey sample frame, I find that having an office in downtown Yangon and having a website are correlated with an exporter's estimated visibility (see Figure 1.4). Of course, these correlations could simply be the result of these exporters already dealing with a lot of buyer relationships and making a location choice and setting up a website as a result of that. But these correlations suggest that being seen by buyers seems to be more important in making connections with them than any outreach capacity such as traveling abroad as measured in the survey.

Search costs are reported in Table 1.7 for entrant and continuing buyers separately. The median search cost is about $2,600 for entrant buyers and $1,700 for continuing buyers. These costs are somewhat aligned with the process buyers need to go through to find buyers as described in section 1.2. Search costs are on average $1370 lower for buyers who have an office in Myanmar and hence do not have to travel or use local intermediaries to search and connect with local exporters.

Finally, the model and estimation allow to estimate markups for each exporter. Conditional on winning an auction process with a buyer, the markup of exporter i that faces marginal cost $mc_{i,t}$ is given by: $\beta_{j,s}(mc_{i,t}) - mc_{i,t}$. However, for a given transaction, $mc_{i,t}$ is not observed but knowing all the exporters that are in buyer j's set and their marginal cost parameters, one can estimate the distribution of $mc_{i,t}$ conditional on i winning (having the lowest bid) against all other exporters included in j's set. Thus, the expected markup is given by: $\mathbb{E}\left[\beta_{j,s}(mc_{i,t}) - mc_{i,t} \middle| i \text{ wins bid with buyer } j\right]$. Abstracting from fixed costs, I can then estimate total profits for an exporter by aggregating expected transaction markups. The distribution of profits is given in Appendix Figure A.4.

1.7 Counterfactuals

1.7.1 The sources of exporter heterogeneity

I first explore the sources of exporter heterogeneity. Three factors impact the size of exporters: i) the marginal cost parameter which drives exporters' market share with each of the buyers they are connected to, ii) the visibility parameter which drives the number of buyers they do business with and iii) the quantity bought by each of the buyers who exporters are connected to. From the perspective of exporters, iii) is random so I only quantify how i) and ii) contribute to exporter heterogeneity. Unfortunately, the model does not lead to a simple formula to decompose the contribution of each factor in firms' total sales. Thus, I perform two counterfactual analysis to quantify the importance of each parameter in explaning exporter heterogeneity: one where I set all marginal cost parameters to be equal across exporters and one where I set visibility to be the same across all firms.

If all exporters had the same average marginal cost parameter, the likelihood of winning an auction for an order with a given buyer would be the same across exporters. So two exporters included in the sets of the same buyers would sell the same amount over the course of the season. I simulate the distribution of exporters under the hypothesis that they all have the same marginal cost paramater in Panel a of Figure 1.5. I then compare that distribution to the actual distribution of exporters. Under the equal marginal cost assumption, the likelihood of observing very small exporters who only do a couple of shipments per year because their high average marginal cost prevents them from winning a significant number of bids is lower. Similarly, observing exporters who have high market shares with each of the buyers they serve is also less likely. The standard deviation of the distribution of log(orders) is about 13 percent lower with equal marginal cost parameters than the distribution in the data.

The exporter size distribution under the equal visibility assumption is presented in Panel b of Figure 1.5. Under this scenario, the variation in exporter size would be even smaller. The difference between the two distributions suggest that many of the small (big) exporters in the current distribution are firms with low (high) visibility and hence are able to sell to very few (many) buyers. The standard deviation of log(orders) would be about 30 percent lower if all exporters had the same visibility parameter.

Overall, these simulations suggest that visibility, or an exporter's ability to make connections with foreign buyers is much more important in explaining an exporter performance than his average marginal cost, i.e. his supply-side productivity.

1.7.2 Lower search costs and centralized market

I first consider the case where buyers' search costs would be twice lower. This would correspond to a case where it would be easier for them to search for exporters with for example a list of all exporters with their contact information available online. Under this hypothesis, buyers would on average be purchasing from 5.3 exporters on a given season, compared to 3.4 in the data. In turn, trasaction prices would be on average $6.2 lower. As the average price is about $800/ton, this is roughly equivalent to 0.7 percent lower prices.

A more interesting counterfactual is the centralized market case where search costs are brought to zero. Implementing such a setting would require a platform where buyers could post bid requests to all exporters at the same time with the platform having a market clearing agent to avoid any payment risk for which buyers need to vet exporters for. Under this case, not only would search cost vanish but visibility would be the same across all exporters.

Simulations of such a scenario reveal that transaction prices would be on average $42 lower (or 5.3 percent lower). Having all exporters included in the auction process significantly reduces the expected lowest bid. This exercise also provides insight into why such a decentralized market does not exist. As the drop in prices reflect, moving to a centralized market is equivalent to an important transfer of surplus from exporters to buyers. The logistics needed to move to a centralized market would definitely need to come from the exporter side but this analysis reveals that exporters would not have incentives to do so. While moving to a decentralized market would create winners and loosers among Myanmar exporters, the loosers clearly outweight the winners. First, under a centralized market, about 6 percent of exporters would be driven out of the market, not making any transaction with foreign buyers over the course of the season. Given the high number of firms bidding for each order, the likelihood that high marginal cost exporters have the lowest bid for a given order is close to zero. Second, simulating profits shows that 55 percent of exporters would make lower profits under a common exchange market than under the current setting and for 60 percent of the exporters who would make higher profits, the additional profits would be less than $8,000.

However, moving to a centralized market is likely to increase demand from foreign buyers which could benefit farmers upstream. Yet, the setting does not allow to estimate the foreign buyer's price elasticity so the exact impact of such a policy cannot be quantified. Moreover, lower transaction prices for foreign buyers might not pass through to lower prices for the final buyers in the destination country which are the ones determining the quantities bought from Myanmar. Finally, in recent years, India has imposed quotas on some bean varieties imported so the quantities exported in the future are likely to be determined by such policies rather than transaction prices in Myanmar.

1.8 Conclusion

This chapter studies the importance of an exporter's *visibility* to foreign buyers in explaining its performance. I argue that even for homogeneous goods, conditional on their costs, firms differ in the level of demand they capture. The search process that buyers need to go through in a decentralized market generates more sales opportunities for exporters who can easily be found. This allows high cost exporters to subsist in the market by reducing the degree of competition they are exposed to.

Methodogically, this study builds on recent developments in the search and auction literature to identify a cost and visibility parameter for each exporter and a search cost for each foreign buyer. The auction model allows to identify sellers' costs from transaction prices and the identity of the seller who wins the contract. Because of the time variability in exporters' marginal costs, running a competitive bidding process with one more supplier allows buyers to source the good at lower prices on average. However, searching for additional exporters is costly, and this trade-off provides a way to identify search costs.

The estimates reveal that the magnitude of buyers' search costs is important and that *visibility* explains about 30 percent of the exporter size distribution. While this paper focuses on trade from a single country but looks at the allocation of purchases across exporters, a parallel can be drawn into sourcing decisions from multiple countries. In a world where buyers can source goods from different origins but need to pay a fixed cost for searching in each country separately, a country's visibility and marketing capacity could be an important factor in explaining world trade patterns. As such, this paper also provides a rationale for export promotion at the country level.

Tables and Figures

Table 1.1: Descriptive statistics

Foreign buyers	
Average number of buyers per season	228
Average number of buyers making more than one order per season	145
Average number of orders made per buyer per season	24
Entrant share of total buyers	41%
Buys Black Matpe (Urad)	0.84
Buys Toor Whole (Pigeon peas)	0.39
Buys Mung Beans (Green grams)	0.66
Exporters	
Average number of exporters per season	168
Entrant share of total exporters	23%
Buyer-exporter relationships	
Average number of exporters per buyer	3.4

Notes: This table presents summary statistics from the customs data. Many buyers are small and only make one purchase per season. Buyers may buy one, two or three different types of beans. There is high turnover of buyers and exporters.

Table 1.2: Evidence of search

Panel A: Entry of foreign buyer

Dep. var:	Price residual - $\tilde{P}_{ijt}$ (in $)		
	(1)	(2)	(3)
Entrant foreign buyer	12.958^{***}		
	(1.709)		
Foreign buyer experience		-2.410^{*}	
		(1.305)	
Foreign buyer has office in Myanmar			-14.583^{***}
			(0.826)
Foreign Buyer FEs	No	Yes	No

Panel B: Number of exporters in foreign buyer's set

Dep. var:	Price residual - $\tilde{P}_{ijt}$ (in $)		
	(1)	(2)	(3)
Number of exporters in foreign buyer's set	-1.901^{***}	-1.359^{***}	-2.467
	(0.295)	(0.499)	(1.813)
Foreign Buyer FEs	No	Yes	No
Exporter X Product X Week FEs	No	No	Yes

Notes: This table shows price patterns that are consistent with buyers paying search costs to find exporters. The unit of observation is a transaction. The dependent variable is $\tilde{P}_{ijt}$ the price residuals obtained from Equation 1. In Panel A, *Entrant foreign buyer* is a dummy equal to one if the buyer was not purchasing in the prior season. *Foreign buyer experience* is the number of season since a given buyer entered the market. *Foreign buyer has office in Myanmar* is a dummy equal to one if the buyer is listed in the yellow pages as having an office in Yangon (as opposed to operating solely from abroad). In Panel B, *Number of exporters in foreign buyer's set* is the number of distinct exporters a buyer purchases from on a given season. Standard errors clustered at the exporter level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.3: Evidence of auctions

Dep. var:	Overall market share	Within buyer market share	Within buyer-product market share
	(1)	(2)	(3)
$\tilde{P}_{ijt}$ - Exporter X Season FEs	-0.002^{***}	-0.024^{***}	-0.015^{**}
	(0.000)	(0.006)	(0.007)
Season FEs	Yes	No	No
Foreign Buyer X Season FEs	No	Yes	No
Foreign Buyer X Product X Season FEs	No	No	Yes

Notes: This table shows correlations between prices and market shares that are consistent with the hypothesis that buyers source beans through a competitive process between exporters in their set. The right hand side variable is the exporter-(buyer)-season fixed effect of residual prices as defined in Equation 1. The dependent variable is in column (1) the overall market share of a given exporter, in column (2) the within-buyer market share and in column (3) the within buyer and bean type market share. Standard errors clustered at the exporter level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.4: Exporter-buyer matches and divorces

Panel A: New exporter-buyer relationships

Dep. var:	Number of new matches per exporter			
	(1)	(2)	(3)	(4)
$\tilde{P}_{ijt}$ - Exporter X Season FE	−0.001		0.000	
	(0.002)		(0.002)	
$\tilde{P}_{ijt}$ - Exporter X Season FE - in previous season		−0.002		0.001
		(0.001)		(0.001)
Season FEs	Yes	Yes	Yes	Yes
Exporter FEs	No	No	Yes	Yes

Panel B: Survival analysis

Dep. var:	Exporter survival in relationship (dummy)					
	(1)	(2)	(3)	(4)	(5)	(6)
Length of relationship	0.103***	0.072***				
	(0.013)	(0.017)				
Average transaction price in previous season			−0.000	−0.000		
			(0.000)	(0.000)		
Market share with buyer in previous season					0.002	0.031
					(0.008)	(0.022)
Season FEs	Yes	No	Yes	No	Yes	No
Foreign buyer X Season FEs	No	Yes	No	Yes	No	Yes

Notes: This table documents stylized facts about matches and divorces between exporters and buyers. In Panel A, the dependent variable is the number of new buyers a given exporters matches to on a given season. The right hand side variable is the exporter-season fixed effect of residual prices as defined in Equation 1. It shows that exporters who sell or sold in the previous season at low prices do not necessarily make more connections with buyers in the current season. In Panel B, the dependent variable is a dummy equal to one if a given exporter survives in his relationship with a buyer to which he has made at least one trade in the previous season. *Length of relationship* is the number of seasons since the buyer-supplier relationship exists, *Average transaction price in previous season* is the exporter-buyer-season fixed effect of residual prices as defined in Equation 1 in the previous season and *Market share with buyer in previous season* is the market share a given exporter had with a buyer in the previous season. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.5: Estimated parameters

Parameter	Black Matpe (Urad)	Toor Whole (Pigeon Peas)	Mung Beans (Green grams)	All beans
μ	7	4	1	5.5
σ	22	14	13	19
σ_m	9.5	6.8	5.8	6.3
v	19.0	12.1	11.2	16.4
σ_e	66	76	76	67

Notes: μ is the mean of the distribution of marginal costs across all exporters and σ its standard deviation. σ_m is the standard deviation of the distribution of the mean marginal costs parameters ($mc_{i,s}$) across exporters. v is the standard deviation of the distribution of marginal costs across time for a given exporter. σ_e is the standard deviation of the measurement error in the data.

Table 1.6: Analysis of estimated $mc_{i,s}$

Panel A: Correlation with average transaction price				
Dep. var:		$mc_{i,s}$		
	Black Matpe	Toor Whole	Mung Beans	
$\widetilde{P}_{ijt}$ - Exporter X Season FEs	0.175^{***}	0.157^{***}	0.111^{***}	
	(0.004)	(0.003)	(0.002)	
Season FEs	Yes	Yes	Yes	
Panel B: Correlation across bean types				
Dep. var:		$mc_{i,s}$ - Black Matpe		
$mc_{i,s}$ - Toor Whole	0.251^{***}	0.282^{***}		
	(0.050)	(0.069)		
$mc_{i,s}$ - Mung Beans			0.082	0.135^{*}
			(0.055)	(0.072)
Season FEs	Yes	No	Yes	No
Exporter FEs	No	Yes	No	Yes
Panel C: Serial correlation				
Dep. var:		$mc_{i,s}$		
	Black Matpe	Toor Whole	Mung Beans	
$mc_{i,s-1}$ - Black Matpe	0.170^{***}			
	(0.042)			
$mc_{i,s-1}$ - Toor Whole		0.321^{***}		
		(0.040)		
$mc_{i,s-1}$ - Mung Beans			0.197^{***}	
			(0.043)	
Panel D: Correlation with visibility parameter				
Dep. var:		mc_i		
	Black Matpe	Toor Whole	Mung Beans	
γ_i	61.976	22.309	-49.472	
	(134.991)	(88.583)	(68.342)	

Notes: This table studies various correlations with the estimated exporter-level average marginal cost parameters $mc_{i,s}$ for each season and each bean type. Panel A looks at the correlation between the $mc_{i,s}$ and the exporter-season fixed effect of residual prices as defined in Equation 1. Panel B looks at the correlation between the estimated $mc_{i,s}$ across bean types. Panel C looks at the correlation between the estimated $mc_{i,s}$ for season s and season $s-1$. Panel D shows the absence of correlation between the estimated cost parameter $mc_{i,s}$ and the estimated visibility parameter γ_i. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

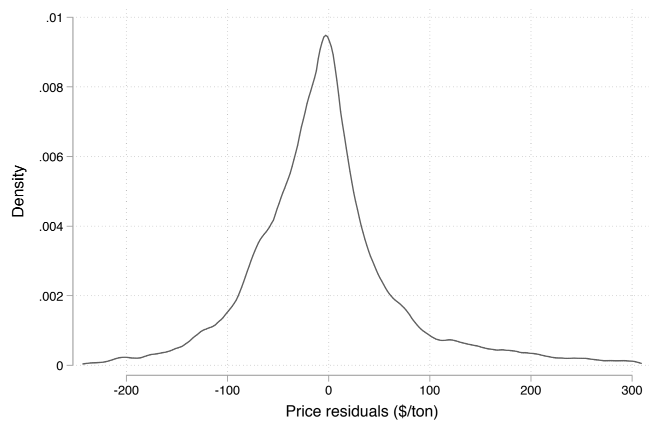
Table 1.7: Search costs

Percentiles	Entrant buyers	Continuing buyers
10th	$ 1,643	$ 520
25th	$ 1,725	$ 575
50th	$ 2,588	$ 1,725
75th	$ 4,313	$ 4,908
90th	$ 18,426	$ 12,077

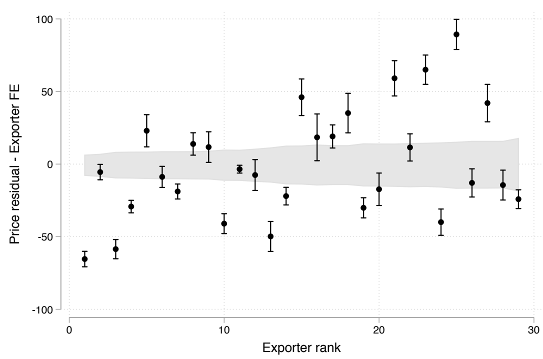
Notes: This table describes the distribution of estimated search costs across buyers for new and continuing buyers separately. These search costs are obtained from equation 2.1 and 2.2.

Figure 1.1: Distribution of price residuals and Exporter and Buyer FEs

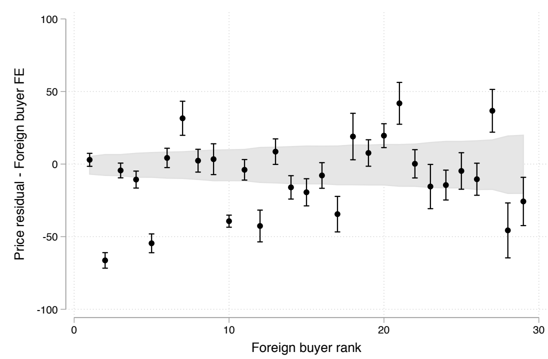
(a)



(b)



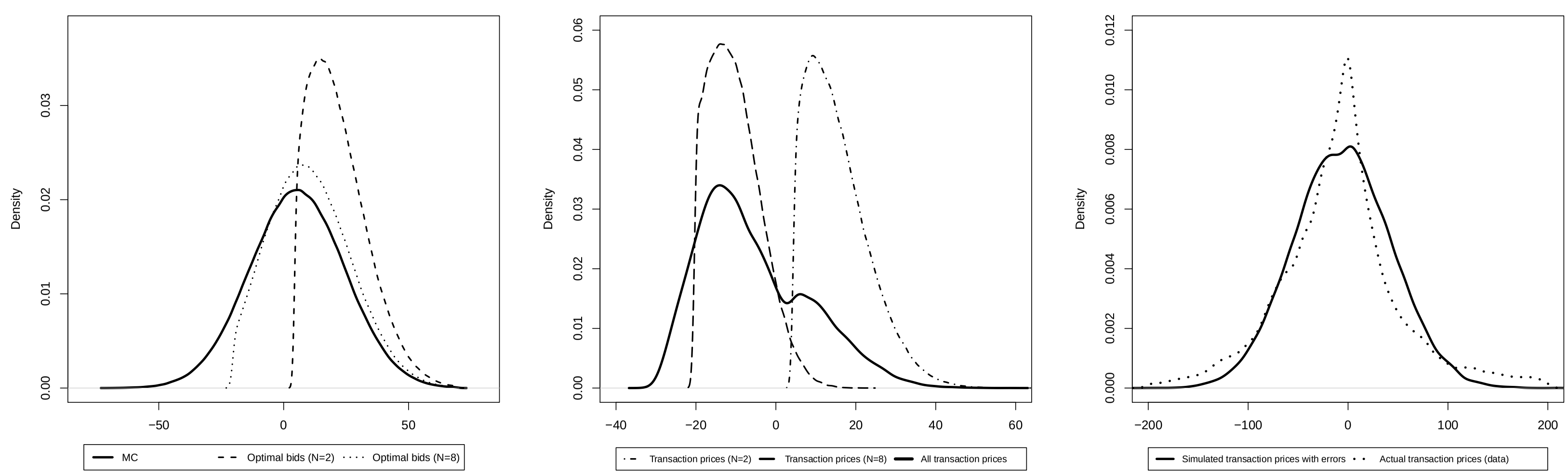
(c)



Notes: Panel (a) of this figure shows the distribution of the residual prices. Panel (b) and (c) show the price fixed effects for the top exporter and the top buyers. If residual prices were *iid* draws of reporting errors in the customs data, the fixed effects should lie in the 95 percent confidence interval (the grey area).

Figure 1.2: Simulated densities

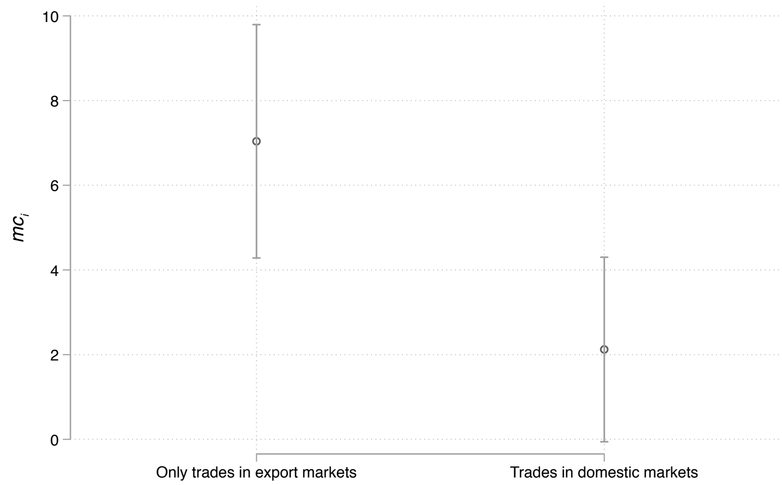
(a) (b) (c)



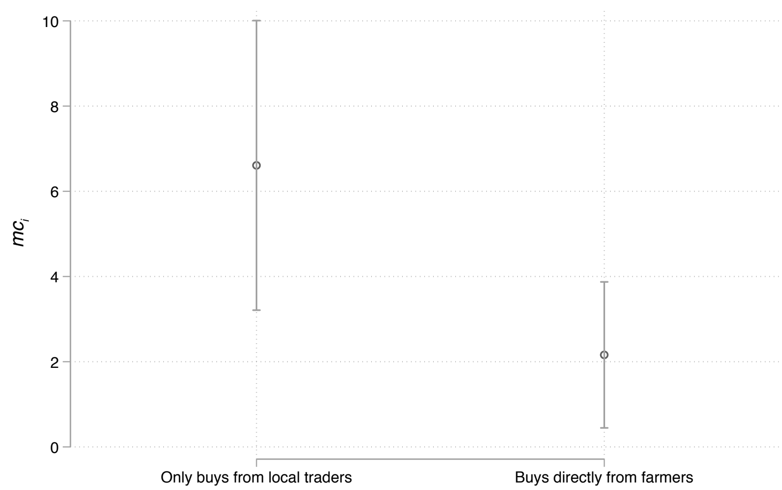
Notes: This figure shows the estimated distribution of various functions in the model presented in the main text. The thick line in panel (a) shows the distribution of marginal costs across exporters and across time. The two dashed line present the distribution of optimal bids for a buyer with respectively two and eight exporters in his set. Panel (b) plots the distribution of the lowest bid for buyers with two and eight exporters in their set and the overal distribution of transaction prices. Panel (c) presents the simulated price residuals as well as the distribution of price residuals in the data.

Figure 1.3: Correlation between mc_i and survey data

(a)



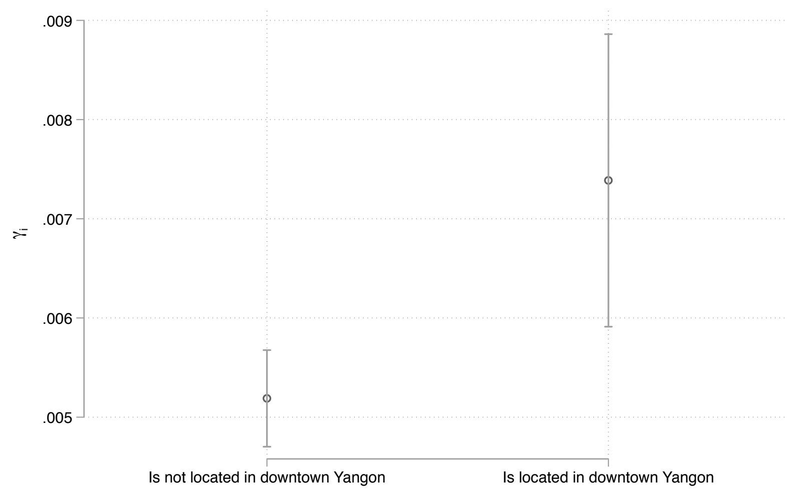
(b)



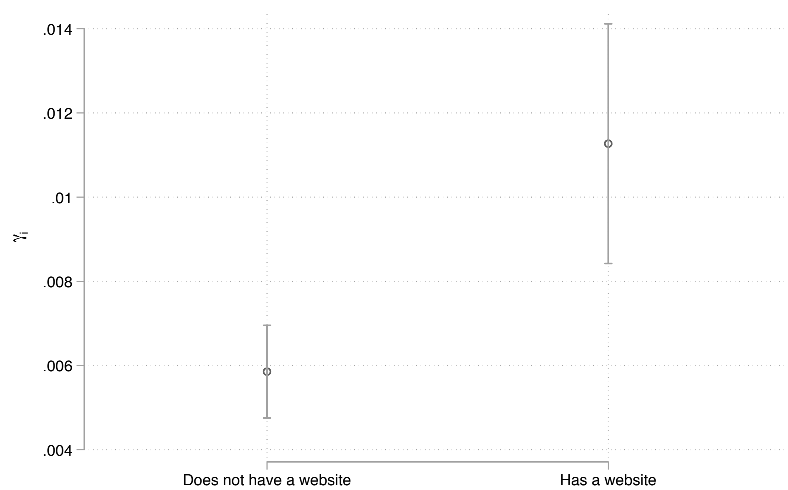
Notes: This figure illustrates the correlation between the estimated exporter cost parameters mc_i and survey data. Panel (a) shows a negative correlation between marginal costs and the likelihood that an exporter also trades in the domestic market. Panel (b) shows that exporters who buy some of the beans directly from farmers (as opposed to solely from traders) have lower estimated marginal costs. The bars represent the 90 percent confidence interval.

Figure 1.4: Correlation between γ_i and survey data

(a)



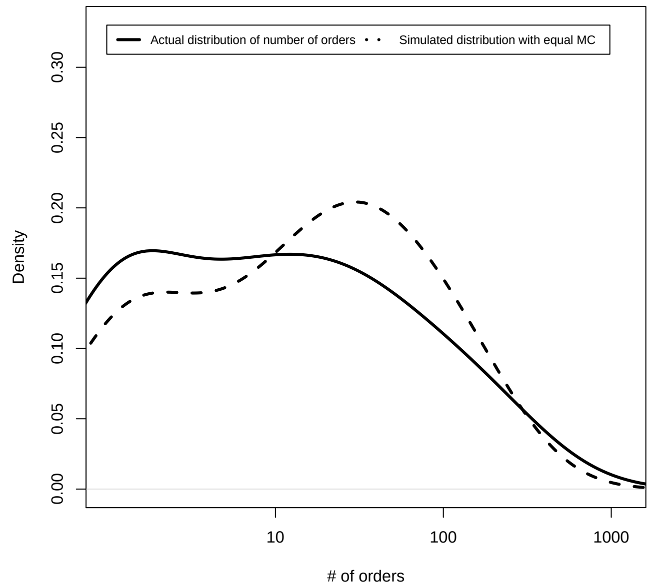
(b)



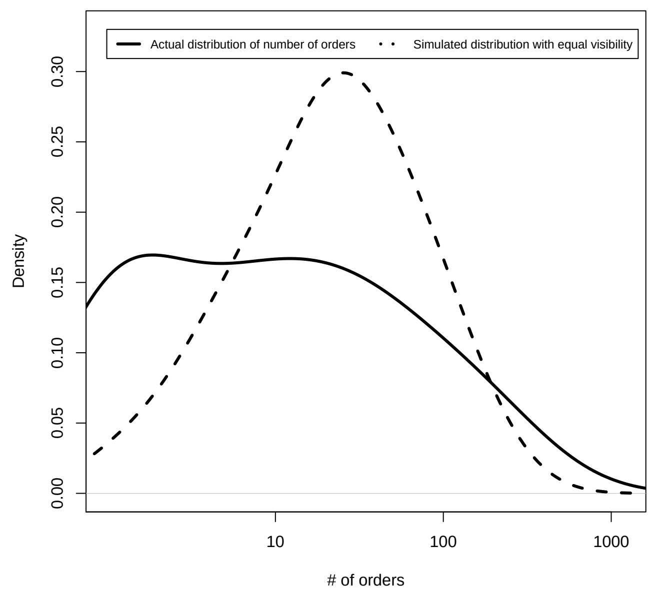
Notes: This figure illustrates the correlation between the estimated exporter visibility parameter γ_i and data collected to build the survey sample frame. Panel (a) shows that exporters who have an office in downtown Yangon have a higher visibility parameter. Panel (b) shows that buyers that are listed online are more likely to have a high visibility parameter. The bars represent the 90 percent confidence interval.

Figure 1.5: Counterfactual distributions

(a)



(b)



Notes: This figure plots the two counterfactual distributions used to quantify the importance of visibility versus cost in explaining the distribution of exporter size. Exporter size corresponds to the number of orders that a seller gets on a given season. Panel (a) plots the actual distribution with the distribution under the scenario that all exporters have the same marginal cost parameter. Panel (b) plots the actual distribution with the distribution under the scenario that all exporters have the same visibility.

Chapter 2

Vertical Integration, Supplier Behavior, and Quality Upgrading among Exporters

Joint with Christopher Hansman, Jonas Hjort and Gianmarco León

2.1 Introduction

Why do so many of our economic transactions occur within firm boundaries (Coase, 1937; Gibbons, 2005a; Lafontaine and Slade, 2007)? Vertical integration occurs for many different reasons, and motives vary by context.[1] However, as global incomes rise and barriers to trade fall, one potential motive has gained increased relevance: firms integrating in order to improve product quality. Access to wealthier, quality sensitive markets brings rising returns to output quality,[2] but producing high quality output typically requires high quality inputs (see e.g. Kugler and Verhoogen, 2012; Halpern et al., 2015a; Amodio and Martinez-Carrasco, 2018). Because input quality is often hard for firms to measure and contract over (Gibbons, 2005a; Lafontaine and Slade, 2007)—especially where institutions are weak (Woodruff, 2002; Nunn, 2007)—organizational structure may play a crucial role in firms' ability to meet demand for quality.

In this chapter, we test the hypothesis that firms vertically integrate in order to produce higher quality products. This hypothesis is inspired by classical theories characterizing how firm boundaries are expected to respond to output objectives (Baker et al., 2001, 2002; Gibbons, 2005a,b) when suppliers multitask (Holmstrom and Milgrom, 1991). However, given the rarity of data on product quality and internal firm structure, and the challenges of isolating firm strategy from confounding factors, causal evidence on the extent to which firms change their organizational structure to upgrade quality has remained elusive.

The context we study, the Peruvian fishmeal industry, enables progress. The simple structure of the sector creates an ideal setting in which to study the relationship between integration and output quality. Independent and integrated suppliers deliver inputs of hard-to-observe quality to manufacturers. Manufacturers convert these inputs into a

[1] Empirical work on the causes and consequences of firms' choice of organizational structure in developing countries began in earnest with Woodruff (2002). Gibbons (2005a); Lafontaine and Slade (2007); Bresnahan and Levin (2012) provide excellent overviews of the broader literature on firms' structure.

[2] See e.g. Sutton (1991, 1998); Hallak (2006); Verhoogen (2008); Manova and Zhang (2012); Atkin et al. (2017a); Bastos et al. (2018).

vertically differentiated but otherwise homogeneous product.[3] Furthermore, uniquely rich data on the sector's entire chain of production is available, including within-firm transactions and direct measures of output quality. Finally, there is substantial—and plausibly exogenous—variation in the *quality premium*, the price differential between high and low quality fishmeal. This allows us to isolate explicit strategic responses to incentives to quality upgrade.

Our analysis proceeds in four steps. We first present a simple theoretical framework that describes how and why a firm's choice of organizational structure may depend on its output quality objectives. We then ask if quality upgrading motives are—empirically—a direct determinant of integration decisions; that is, whether a manufacturer is more likely to integrate its suppliers when its returns to shifting from low to high quality production are higher. Next, we explore the mechanisms through which output quality objectives may impact integration decisions. We estimate how organizational structure affects supplier behavior, focusing particularly on "switchers"—suppliers who supply the same plant before and after being integrated (or sold). To conclude, we investigate whether integration ultimately raises output quality.

There are several reasons why unique data is available on the fishmeal industry in Peru. The regulatory authorities record all transactions between fishmeal plants and suppliers, and require firms to report each of their plants' production of fishmeal in the "prime" (high) quality and the "fair average" (low) quality range each month, providing a direct measure of quality.[4] We can link these input and output quantities to all export transactions, which are recorded in customs data. Furthermore, researchers—but not

[3]Fishmeal is a brown powder made by burning or steaming fish, and mostly used as animal feed. Peru's fishmeal industry accounts for around 3 percent of GDP (Paredes and Gutierrez, 2008; De La Puente et al., 2011). Price differentials across shipments of a given quality level in a given time period are negligible (see Sub-section 2.2.2). Our focus on a vertically differentiated but horizontally homogeneous product is inspired by influential earlier papers testing market power theories of integration in sectors producing homogeneous products (Syverson, 2004; Hortacsu and Syverson, 2007; Foster et al., 2008).

[4]See Goldberg and Pavcnik (2007); Khandelwal (2010); Hallak and Schott (2011) for discussion of the indirect quality proxies used in the existing literature, which risk conflating quality with mark-ups and horizontal differentiation, and Atkin et al. (2017a) for an example of direct measures of quality.

manufacturers—directly observe both independent and integrated suppliers' behavior because fishing boats are required to transmit GPS signals to the regulatory authorities.[5] In combination, these data sources allow us to track the flow of goods—from suppliers, to manufacturers, to foreign buyers—and provide the measures of output quality and firm-supplier transactions necessary to establish a correlation between the quality of a firm's output and the organizational structure of its production chain.

However, even if documented with ideal data, such a correlation may reflect third factors rather than an explicit organizational choice made *in order to* "climb" the quality ladder. It could be, for example, that productivity or demand affect both firms' choice of structure and products produced *without the two being directly related*. It could also be that firms integrate for reasons other than quality—for example to assure their own or restrict competitors' general access to inputs—but in the process coincidentally produce higher quality output. To identify a direct relationship between output quality objectives and integration, exogenous variation in incentives to upgrade quality—the quality premium—that firms are *differentially exposed to* is needed.

The quality premium facing Peruvian fishmeal firms varies considerably during the period we study. This allows us to test our causal hypothesis. Our empirical strategy exploits season-to-season variation in the quality premium that is due to fluctuations in the *regulatory fishing quota*-driven supply of high (and low) quality fishmeal in *countries other that Peru*. We construct an instrument for the quality premium, and test whether these fluctuations affect firms differently depending on their scope for upgrading quality.

We begin our analysis with a stylized model, which demonstrates how characteristics of the Peruvian fishmeal industry map directly to the existing theoretical work we build on. Fishmeal manufacturers face two important contracting challenges. First, the quality of the product's primary input—fish—is difficult to observe and, because

[5]The regulators do not allow manufacturers to access data on the behavior of independent suppliers. This is the primary reason why manufacturers and independent suppliers cannot contract over GPS-measured actions.

of its perishable nature, even harder to contract upon. Second, the presence of outside options—other fishmeal firms who may value input quality less—complicates controlling the incentives of an independent supplier (see also McMillan and Woodruff, 1999).[6] Holmstrom and Milgrom (1991) elegantly demonstrate how, when suppliers face a trade-off between producing inputs of high quality or in high volumes, weakening incentives over easier-to-measure *quantity* may be necessary to ensure that suppliers do not neglect *quality* (see also Holmstrom and Tirole, 1991; Holmstrom, 1999). In many situations, the best or only way to do so may be to bring the suppliers inside the firm (Baker et al., 2001, 2002; Gibbons, 2005a,b).

To take these textbook theoretical insights to the data, we first demonstrate that output quality is in fact significantly positively correlated with integration. Our primary measure of quality is the share of a firm's output that is of high quality grade—which we directly observe in production data. We also consider a fine-grained (but not-directly-observed) measure of the average quality grade of a firm's output. Similarly, we consider two different measures of vertical integration, one based on *use* of integrated versus independent suppliers and the other on supplier ownership. The relationship we establish holds across each of our measures of a firm's output quality and each of our measures of a firm's organizational structure.[7]

---

[6]It is in theory possible to imperfectly measure fish quality with chemical tests. As discussed in Section 2.2, industry insiders informed us that such tests were much too expensive and impractical to use during our data period. Alternatively, manufacturers and their suppliers could attempt to contract on plants' *output* quality. This would be difficult because of noise—input from multiple suppliers is, for technological reasons, typically used in a given batch of fishmeal, and other hard-to-measure exogenous factors also influence output quality realizations—and, more importantly, because outside inspectors would need to be able to determine if low output quality was due to poor input quality or actions taken by the fishmeal plant itself during the production process. The dynamic version of our model in Appendix B.2 demonstrates why the presence of other fishmeal manufacturers who value input quality less can make repeated interactions solutions to these challenges infeasible.

[7]Our primary measure of integration is the fraction of inputs that are sourced from integrated suppliers ("Share VI")—a measure that is motivated by our hypothesis that the characteristics of the inputs actually *delivered* by a supplier changes with integration so that integration and output quality are causally linked via firms' (and individual plants') *production* process. (On *use* of integrated versus independent suppliers, see also, among others, Baker and Hubbard (2003); Atalay et al. (2014); Breza and Liberman (2017)). Alternatively, we also consider the number of suppliers owned. Note that, in our setting, since boats' total seasonal catch is governed by a quota—and boats almost always exhaust their quota over the course of a

We then proceed to our central empirical analysis, which consists of three key pieces of evidence. The first of these—and this paper's main finding—shows that vertical integration is used by firms *as a strategy* for increasing output quality. To demonstrate this causally, we develop an IV for firm-specific incentives to upgrade quality. We instrument for the quality premium—the difference between the price of high and low quality grades—using the *quantities* produced by other top exporters. Because the other top exporters specialize in particular quality grades (e.g. Iceland produces primarily high quality fishmeal, while Thailand produces primarily low quality), and because their production quantities are driven by country specific regulatory fishing quotas, these quantities generate plausibly exogenous variation in the premium. We test if this variation *differentially* impacts Peruvian firms' integration decisions depending on their firm-specific scope for upgrading quality. Firms that are already producing mostly high quality output have little scope to improve quality further and are hence unlikely to respond strategically to an increase in the quality premium. Conversely, firms producing mostly low quality output have significant scope to upgrade. Thus, if our hypothesis is true—that is, if firms use integration as a strategy for upgrading quality—then firms producing primarily or exclusively high quality output will face weaker incentives to integrate when the quality premium rises.[8]

We find that Peruvian manufacturers integrate when their incentives to upgrade quality rise, and vice versa. *The industry as a whole* integrates when the quality premium increases for exogenous reasons and de-integrates when the quality premium falls. The integration (de-integration) response is stronger for firms with greater scope for upgrad-

---

season—Peruvian fishmeal manufacturers can generally increase the total amount of inputs they obtain from integrated suppliers in a given production season only by acquiring suppliers (see Section 2.2).

[8]A potential criticism of this argument is that firm-specific scope for quality upgrading might also correlate with some unobservable related to the marginal cost of integration and/or quality upgrading. The most natural forms of such arguments—that firms producing high quality output (a) have low marginal costs of further upgrading and are also more likely to integrate in general, or (b) have low marginal costs of integration and are also more likely to upgrade in general—predict the opposite of our findings. Furthermore, as our primary measure of quality is a share (rather than a level), arguments such as (a) cannot hold for firms that are already producing exclusively high quality, as they mechanically have no scope to improve further. Our results considering these firms are similar.

ing (downgrading) quality.[9] In an alternative approach shown in the Appendix, we exploit a different form of variation to show that firms similarly integrate when faced with greater firm-specific relative *demand* for high quality output. Finally, and crucially, we show that firms' organizational response to the quality premium does not reflect associated income shocks or general incentives to expand production of any-quality fishmeal: firms *do not* integrate suppliers when faced with higher average prices.

This first piece of evidence is hard to reconcile with alternative theories in which higher output quality is an unforeseen by-product of vertical integration driven by other motives and with stories wherein organizational structure and output quality are not causally linked in the "minds" of firms. Several of the most prominent, *specific* alternative explanations—such as firms integrating suppliers for general supply assurance reasons but coincidentally achieving higher quality in the process, and foreclosure motivations—are also inconsistent with other features of the context we study and auxiliary findings.[10] Our results indicate that quality upgrading itself is an important motive for integrating suppliers.

Next we explore *why* firms use integration as a strategy for upgrading quality. Our second key piece of evidence shows that integration changes suppliers' behavior, causing them to shift towards quality-increasing actions. We proxy for actions that increase input quality—i.e., fish freshness (FAO, 1986)[11]—using GPS-based measures. We show

[9]The long-term trend is towards more integration in the Peruvian fishmeal industry, and the long-term trends in demand for quality and average output quality in Peru are also positive. These broad patterns are consistent with our hypothesis. However, it is higher frequency variation *around* the long-term trends that we exploit to *test* our hypothesis. For example, we also observe de-integration during our data period—sales of boats from fishmeal firms to independent co-ops or captains, and from one fishmeal firm to another.

[10]We show that (i) Peruvian fishmeal manufacturers appear to achieve general supply assurance primarily through repeated interactions with independent suppliers (see also Martinez-Carrasco, 2017)—the quantity supplied by a given supplier to a given firm is in fact *lower* after integration (/before de-integration)—and (ii) repeated interactions with suppliers do not enable firms to produce higher quality output. Similarly, the relationship between output quality and integration holds when we control for a firm's share of the industry's total production, in contrast to what traditional "foreclosure" theories would predict.

[11]"Freshness of raw material is important in its effect on the quality of the protein in [quality of] the end product [fishmeal]. The importance of minimizing the time between catching fish and processing,

that—as the managers in the industry we interviewed reported to us[12]—a *given supplier supplying a given plant* delivers lower total quantities, but inputs whose quality has been better maintained, when integrated with the plant. We also show that, in the context we study, it is integration *per se*—not repeated interactions—that influences a supplier's quantity-versus-quality behavior. This result is consistent with a dynamic version of our model—which, along with the result, is shown in Appendix B.2—and with the fact that suppliers that de-integrate from a firm/plant supply that firm/plant almost as often after the change in status. Finally, we consider the possibility that integration affects behavior not via a supplier's quantity-quality trade-off, but instead via associated knowledge transfers, of the form that Atalay et al. (2014) convincingly show occur in the U.S. post-integration. We find that a given supplier behaves "as an integrated supplier" only when supplying its owner firm and not when owned by one firm *but supplying another*. Reconciling this finding with knowledge transfer theories would require such knowledge transfers to be useful only when supplying the parent firm. We ultimately cannot rule out that other incentives emanating from organizational structure itself than those our model focuses on also help explain the impact of integration on supplier behavior, but our *third* piece of evidence suggests that any such alternative incentives would need to ultimately benefit downstream output quality.

Our third piece of evidence indicates that vertical integration in fact increases output quality, as the managers in the industry we interviewed reported to us. (In the words of Ricardo Bernales Parodi, Managing Director of Pesquera Diamante, Peru's third largest fishmeal company: "Around 80 percent of my company's fishmeal is high-quality. If all my inputs came from integrated boats, around 95 percent of my fishmeal output would be of high-quality." (Authors' translation).) We first show that the firm level relationship between the share of inputs coming from integrated suppliers—Share VI—

---

and of keeping the fish at low temperatures by icing [which reduces the amount of fish a boat can fit], has already been mentioned" (FAO, 1986, sub-section 10.1.2).

[12]In the words of a prominent executive of Peru's National Fisheries Society: "as a consequence of integration, they must adopt my rules. Things like saying, 'hey, you must offload raw 24 hours after having caught it, at the maximum."' (Authors' translation).

and output quality holds also at the individual plant level, including *within* firms. We then instrument for a plant's Share VI using a leave-firm-supplied-out measure of the local presence of a particular type of supplier that is prohibited by regulation from being integrated. The presence of such suppliers fluctuates due to natural variation in fish density, weather, and decisions made by their captains. The logic behind the instrument is simply that a plant—holding fixed output quality objectives—will be forced to source a higher share of its inputs from integrated suppliers when there happens to be a local scarcity of independent suppliers.[13] The IV estimates are very similar to OLS estimates.

When viewed through the lens of our model, the three key pieces of evidence we present each follow logically from each other. We conclude that firms vertically integrate in order to produce higher quality products, and that the reason they do so appears to be that integration changes supplier behavior in a way that increases output quality.

A natural question is whether our results are specific to Peru's fishmeal industry. However, because input quality is so frequently difficult to observe (and hence incentivize), the challenges we describe—while far from universal—are likely typical of industries producing vertically differentiated output, particularly in settings where contracts are difficult to enforce.[14] In the concluding section of this chapter, we document a positive relationship between (a proxy for) the average quality of a given type of manufacturing product a country exports to the U.S. and the average degree of vertical integration among the exporters. This provides suggestive evidence that the relationship between firms' output quality and organizational structure we establish in this chapter may hold more broadly in export manufacturing industries.

[13]Plausible arguments against the exclusion restriction underlying our interpretation of the IV results would require a positive sign on the first stage (negative correlation between the use of independent suppliers by different plants in a locality), a negative sign on the second stage (use of independent suppliers increasing output quality), and/or a time-varying, location level component of output quality (that goes beyond the presence of independent suppliers)—none of which we find. The IV results are very similar if we instrument with the total number of independent suppliers present, rather than the subset that is independent by law.

[14]There is a robust relationship between countries' input-output structure and their level of contract enforcement (Nunn, 2007; Boehm, 2016), and vertical integration is more common in developing countries (Acemoglu et al., 2009; Macchiavello, 2011).

Our study bridges and advances the literatures on the boundaries of the firm and quality upgrading.[15] We make three contributions to the former. First, we identify an overlooked motivation for vertical integration. By showing that firms vertically integrate in order to raise output quality, we advance the body of work on the *causes* of organizational form (for seminal empirical work, see Hart et al. (1997); Baker and Hubbard (2003, 2004); Forbes and Lederman (2009)). Existing studies convincingly demonstrate how firms change their relative use of integrated suppliers in response to changes in e.g. available contracts (Breza and Liberman, 2017) or monitoring technology (Baker and Hubbard, 2003). We instead study how firms change their organizational structure when their output *objectives* change.[16]

Second, and building on earlier studies of the behavior of integrated and independent suppliers (Mullainathan and Sharfstein, 2001; Baker and Hubbard, 2003, 2004; Macchiavello and Miquel-Florensa, 2016), we provide what to our knowledge is the first evidence on how integration changes the quality-oriented behavior of a given supplier supplying a given firm. To do so we follow Atalay et al. (2014) and exploit changes in integration within supplier-firm pairs.

Finally, we show that vertically integrating raises output quality, which to our knowledge has not been done before. The one-dimensional nature of quality differentiation in our setting allows us to document this.[17] In general, there is little existing evidence on

[15] There is also a prominent literature studying the relationship between integration and international trade (Antràs (2014, 2016) provides excellent overviews of this literature), but our focus is on firms' *domestic* organizational structure.

[16] In a superficial sense, our finding that higher average fishmeal prices do not lead to more integration in the Peruvian fishmeal industry contrasts with the innovative work of Alfaro et al. (2016). We see our results as largely consistent with and complementary to theirs, however. Both their analysis and ours emphasize the impact of prices in the context of certain goods—in our case high quality products—where integration generates a gain in efficiency. We highlight that this efficiency gain is not generic, but rather depends on firms' quality objectives, while they emphasize that efficiency gains can also depend on the need to coordinate production stages.

[17] In settings where product differentiation is multidimensional, an analysis like ours would be difficult. Like this paper, the pioneering study by Forbes and Lederman (2010) exploits exogenous drivers of use of integrated suppliers, showing that routes airlines self-manage have fewer delays/cancellations (see also Gil et al., 2016; Gil and Kim, 2016). Other important evidence on the *consequences* of organizational structure includes, among others, Novak and Stern (2008); Gil (2009).

causal consequences of organizational structure for firm performance (see Gibbons and Roberts (2013), and Forbes and Lederman (2010) for a notable exception). Our results also imply that using *independent* suppliers is often efficient for producing output in *high volumes* rather than of high quality (see also Kosová et al., 2013). An especially unusual aspect of this paper is that the data and variation we exploit allow us to identify *both* the effectiveness of particular firm strategy and corresponding determinants of its use. We can therefore show that Peruvian fishmeal manufacturers vertically integrate when quality objectives indicate that they *should* do so.[18]

Both the friction—imperfect contracting over input quality—and the firm objective—producing high quality output—we focus on are especially relevant for poorer countries attempting to help meet growing global demand for quality. This connects our paper with a smaller empirical literature on the causes and consequences of firms' choice of organizational structure in the developing world that began with Woodruff (2002)'s landmark study (see also Natividad, 2014; Macchiavello and Miquel-Florensa, 2016; Martinez-Carrasco, 2017).[19]

The literature on quality upgrading is larger. It is now well-documented that producers of high quality goods use high quality inputs (Kugler and Verhoogen, 2012; Halpern et al., 2015a; Amodio and Martinez-Carrasco, 2018; Bastos et al., 2018), skilled workers (Verhoogen, 2008; Frías et al., 2009; Brambilla et al., 2012; Brambilla and Porto, 2016; Brambilla et al., ming), and export to richer destinations (Hallak, 2006; Verhoogen, 2008; Manova and Zhang, 2012; Atkin et al., 2017a; Bastos et al., 2018). Firms with such a pro-

[18]Existing empirical papers on firms almost exclusively study *either* the effectiveness of a strategy *or* the determinants of its use.

[19]Woodruff finds that forward integration is less common in the Mexican footwear industry when non-contractible investment by retailers is important, as the property rights framework predicts (Grossman and Hart, 1986; Hart and Moore, 1990). Macchiavello and Miquel-Florensa (2016) convincingly show how supply assurance motives influence organizational structure in the Costa Rican coffee industry by relating measures of ex post reneging temptations to ex ante choice of structure (see also Banerjee and Duflo, 2000; Macchiavello and Morjaria, 2015). We follow Natividad (2014) in studying organizational structure in the Peruvian fishmeal industry. He focuses on an earlier period when an unusual regulatory system—industry-wide fishing quotas—generated common pool incentives famously overshadowing other forms of supplier/plant incentives (see e.g. Tveteras et al., 2011), which lead to an "Olympic race" for fish.

file tend on average to be bigger, more productive, based in richer countries themselves, and to face foreign competition in low-quality segments (Schott, 2004; Hummels and Klenow, 2005; Baldwin and Harrigan, 2011; Johnson, 2012; Medina, 2017). We provide the first evidence linking quality upgrading to the boundaries of the firm.

2.2 Background on Peru's Fishmeal Manufacturing Sector

In this section we provide an overview of Peru's fishmeal manufacturing sector. We argue that three features are particularly salient for firms attempting to source high quality inputs: input quantity is measurable at the time of delivery, but input quality is not, and formal contracts appear to be difficult to write.

2.2.1 Sector profile

Fishmeal is a brown powder made by burning or steaming fish (in Peru, the anchoveta), and is primarily used as feed for agriculture and aquaculture. Peru makes up around 30 percent of the world's fishmeal exports. During our data period, 2009 to 2016, around 95 percent of the country's total fishmeal production was exported. The three largest buyers are China, Germany, and Japan, but many other countries also import Peruvian fishmeal (see Appendix Table B.1).

Fishmeal is produced in manufacturing plants located along the coast of Peru, of which there were 94 in 2009. These plants were in turn owned by 37 firms. There is heterogeneity in processing capacity, technology, and the share of production that is of high quality grade across both firms and individual plants in our sample. Firms differ considerably in their average number of export transactions per season, and in the size and value of their shipments. As seen in Appendix Figure B.1, firm size correlates positively with average quality grade produced.

Plants receive inputs of raw fish from their suppliers. The suppliers may be large steel boats—which may be independent or owned by the firm that owns the plant—or smaller wooden boats. Regulations prohibit fishmeal firms from owning wooden boats. There are on average 812 boats active in a given season, and significant heterogeneity in boat characteristics such as storage capacity, engine power, and average quantity caught per trip. Fishing trips last 21 hours (s.d. = 10 hours) and boats travel 76 kilometers away from the port of delivery (s.d. = 46 kilometers) on average. Changes in installed technology are observed in our data but rare both for boats and plants. Table 2.1 shows summary statistics, providing further detail on the sector.

Since 2009, boats in Peru have operated under Individual Transferable Quotas (ITQs), a common resource management system used in fisheries and natural resource sectors worldwide. Individual boats are assigned a share of an industry-wide quota. We limit our analysis to the time period after ITQs were implemented to avoid any potential changes in quality production or integration driven by the quota system.

2.2.2 Product differentiation and quality

An important feature of fishmeal is that *output* quality effectively depends on a single—measurable—dimension: protein content. Batches with protein content above a specified percentage are labeled "prime" quality, and plants report their production of prime and "fair average" (below prime) quality fishmeal to regulatory authorities each month. Price differentials across transactions for Peruvian fishmeal of a given quality grade in a given time period are negligible, highlighting the horizontal homogeneity of the product.

Fishmeal's protein content depends crucially on input characteristics, namely the freshness and integrity of the raw fish that boats deliver (see e.g. FAO, 1986). Freshness and integrity of the fish at the time of delivery in turn depends on choices made by the boat's captain before and during a trip, such as the amount of ice brought on board, the amount of fished packed on board, how tightly fish is packed, and the time spent between a catch and delivery to a plant (FAO, 1986). Because of the relationship between freshness and output quality, fish is processed as soon as possible after offload.

While it is easy to weigh and determine the quantity of fish a boat delivers, it is difficult to quantify or measure fish freshness directly. In theory, chemical tests of total volatile nitrogen content can be used to do so (imperfectly), but the managers in the industry we interviewed reported that such tests were too costly and time-consuming to be usable in Peru during our data period. In addition to the fixed cost of (the human and physical capital required for) adoption, this was due to high marginal cost of use and the value lost if fish was not processed immediately after offload. Footnote 6 discusses the extent to which input quality can be inferred from output quality post-production.

After offload, the fish is weighed, cleaned, and converted to fishmeal using one of two technologies: steam drying (hereinafter "High technology") or exposing the fish directly to heat (hereinafter "Low technology"). The technology used can matter for the protein content achieved.

Peru allows anchovy fishing for fishmeal production during two seasons each year and because of the need for fresh fish, fishmeal plants operate only during the fishing seasons. There were thus 14 fishing and fishmeal production seasons during our 2009-2016 study period. In theory fishmeal can be stored for a short period of time, but we find that almost all is sold before the next production season begins, as shown in Appendix Figure B.2 and discussed below.

2.2.3 Organizational structure

Consistent with our hypothesis, both integration and average output quality have slowly increased over time in the Peruvian fishmeal industry. However, these long-term trends are not the source of the relationship between organizational structure and quality upgrading we establish in this chapter. This is because our empirical strategy exploits variation *around* the long-term trends for identification.

There is significant buying and selling of suppliers during our sample period. As seen in Panel A of Table 2.2, we observe 317 instances where ownership of a steel boat changes hands. In 103 of these instances, a fishmeal firm acquires a supplier that is initially owned independently, that is, by a co-op or an individual captain. However, we also observe 32 instances where a supplier is sold from a fishmeal firm to an independent buyer, and 50 instances where a supplier is sold from one fishmeal firm to another. On average, 28 percent of the boats that are active in a given season are integrated with a fishmeal firm.

In our data, we observe not only supplier *ownership* but also *deliveries* from integrated and independent suppliers. We can therefore construct a measure of the vertical structure of firms' *production* process, namely the share of inputs coming from integrated suppliers ("Share VI"). Peruvian fishmeal manufacturers' Share VI is on average 45 percent. Firms can generally increase or decrease the total amount of inputs that come from integrated suppliers only by buying or selling boats. The reason is that a boat's total catch in a given season is governed by a regulatory quota, and each boat typically exhausts its quota. Of course, a firm may vary its Share VI also by increasing or decreasing its use of independent suppliers. As seen in Appendix Figure B.3, and following the trend in ownership, Share VI slowly increased during our data period—by 2.9 percent from season to season. Approximately 77 percent of this growth came solely from increasing the amount of input coming from integrated suppliers, and the rest from lower total input purchases, as shown in Panel B of Table 2.2.

Importantly for our purposes, Share VI can be defined not just for firms, but also for individual plants within firms. A *plant's* Share VI at a given point in time depends mostly on the organizational structure of the firm the plant belongs to, but there is significant variation across plants within the same firm. This variation depends both on the extent to which firm managers direct integrated suppliers to deliver to one plant over another, and on the presence of independent suppliers near a given plant. The latter varies considerably over time, and depends on variation in weather, fish density, and independent captains' decisions.

Figure 2.1 shows that integration and de-integration primarily represents a change in the formal status of the relationship between a firm/plant and a supplier engaged in frequent and continuing interactions. The figure displays the fraction of trips suppliers deliver to various firms and plants. The top part of the figure focuses on all boats, while the bottom part of the figure restricts attention to the "switchers" we focus on in our empirical analysis of supplier behavior in Section 2.6. These switchers—suppliers that get integrated or sold—deliver to the plant (within the acquiring/selling firm) they interact with most frequently around 41 percent of the time *when independent* (i.e. before getting acquired or after getting sold), and around 45 percent of the time when integrated. Similarly, switchers deliver to the acquiring/selling firm around 63 percent of the time when independent and around 81 percent of the time when integrated.[20]

2.2.4 Contracting and supplier incentives

There is no centralized spot market for fish purchases: plants are spread out along the coast, both because the fish move around and because of geography's influence on the location of ports. Similarly, the movements of boats are a complex function of fishing conditions, weather (winds, swell, etc), and the captains' incentives. Because of the importance of fish freshness, independent captains typically begin contacting plants over the radio on their way to a port after fishing.

We interviewed fishmeal industry associations, a major company's Managing Director, another major company's Chief Operating Officer, and others in the sector to gain a qualitative understanding of the characteristics of the contracts used and the incentives suppliers face. The interviewees reported that captains of boats owned by fishmeal firms generally are paid a fixed wage, in some cases with a bonus tied to some measure of performance.[21]

We are not aware of formal contracts between independent suppliers and firms over when, where, or what quality of fish to deliver. Interviewees reported that payments

[20] In the top part of the figure, we see that, as a whole, integrated suppliers deliver to the firm they deliver to most often (i.e., the parent firm) about 90 percent of their trips, and the plant they deliver to most often 38 percent of trips. Independent suppliers deliver to the firm they deliver to most often around 65 percent of trips, and the plant they deliver to most often 45 percent of trips.

[21] The fishmeal industry associations reported that payment schemes vary across firms; that some pay bonuses tied to measures of performance; but that these are on top of a fixed wage and usually small.

to independent suppliers—while agreed upon case by case—are typically simply the quantity multiplied by a going price. We use internal data on payments to suppliers from a large firm to confirm this. These indicate that independent suppliers at a given point in time are paid a price per metric ton of fish delivered that is essentially fixed: Port×Date fixed effects explain 99 percent of the price variation across transactions.

Our data on suppliers' behavior—discussed in Section 2.4—come from a map the regulators update roughly every hour using the GPS signals all boats are required to transmit to authorities while fishing. Firms are allowed to access information on their integrated suppliers' whereabouts if they install the required technology, but not the GPS data of independent suppliers or those owned by other firms. This is a primary reason why manufacturers and independent suppliers cannot contract over GPS-measured actions.

2.3 Theoretical Framework

2.3.1 Description

In this section we present a simple model to highlight how vertical integration may resolve the contracting issues facing downstream firms that aim to produce high quality output. The intuition of the model is based on two insights. First, high powered incentives to produce quantity can lead to actions that are wasteful and even harmful to quality. Second, the open market provides independent suppliers strong incentives to produce quantity and, in a setting where contracts are difficult to write, the only way to temper those incentives may be to integrate.

The first point of intuition above—the tradeoff between quality and quantity—is one of the classic examples of the challenges of designing incentives in a multitask environment, and in fact is used by Holmstrom and Milgrom (1991) to motivate their seminal work. This is for the simple reason that input quantity is typically straightforward to measure and reward, while quality is not. As a result, care must be taken not to over-incentivize quantity to the detriment of quality.

Of course, the difficulty of determining quality is somewhat of a stereotype: there are goods for which quality depends on something like strength or size or durability that is just as easy to measure as quantity. However, in our setting, this stereotype seems broadly accurate. While the quantity of fish that suppliers deliver is easily measured, the quality of that fish is difficult to ascertain for a purchasing manager examining several tons of anchoveta.

A few pieces of context are helpful to understand the second point of intuition above. Firstly, it appears that contracts are difficult to write ex-ante: independent suppliers retain their right to deliver their catch where they choose. Additionally, while some firms primarily produce high protein content fishmeal, others primarily produce low quality grades, and hence provide a (presumably less quality sensitive) alternative for suppliers to deliver their catch.[22]

With this in mind, a logic applies that is familiar from the models presented in Baker et al. (2001, 2002), based on the notion of integration as asset ownership that follows Grossman and Hart (1986). Even if a firm interested in sourcing high quality inputs has no interest in high volumes, the fact that an independent supplier has the option to sell its inputs to an alternative downstream firm that values quantity creates powerful incentives. The independent supplier will then invest in producing quantity—although it may be wasteful or detrimental—if only to improve its bargaining position with the quality focused firm. By acquiring the supplier, the manufacturer removes this outside option, and hence any incentive for wasteful or harmful investment in quantity. In this sense, integration is valuable precisely because it mutes the power of market incentives, a notion that has been described by Williamson (1971), Holmstrom and Milgrom (1994), and Gibbons (2005a), among others.

[22]A question that our model abstracts from is why firms might want to produce different quality levels simultaneously. We return to this question at the end of Section 2.3.

2.3.2 Model details

We consider a static game with two actors: suppliers and high quality firms. Suppliers take costly actions to produce a good that is valuable both to the firms and in an alternative use. They may be integrated or independent. If the suppliers are integrated, the firms that own them have the right to the good after the actions are taken. If the suppliers are independent, they retain the right to the good. They bargain with the high quality firms over whether to deliver the good or consign it to its alternative use.

We assume that suppliers have two potential actions $\{a_1, a_2\}$, with costs $c(a_1, a_2) = \frac{1}{2}a_1^2 + \frac{1}{2}a_2^2$. These actions impact the surplus created by delivering their inputs to a downstream quality focused firm. We denote this surplus by Q, and refer to it as the quality surplus. Suppliers' actions also impact the surplus they receive by delivering the inputs to an alternative—quantity focused—downstream firm. We denote this by P, and refer to it as the quantity surplus. We assume that the good is specific, in the sense that $Q > P$. In particular, we define:

$$P = a_1$$
$$Q = Q_0 - \gamma a_1 + \delta a_2.$$

with $\gamma, \delta \geq 0$.[23] In this sense, a_1 is a quantity focused action, while a_2 is a quality focused action. While this is a simplified model, a_1 can be thought of along the lines of fishing for extended periods to catch the maximum amount, traveling long distances to find fish in high volumes, or packing the hold tightly with fish. On the other hand, a_2 can be thought of as carrying extra ice on board to keep the catch cool, or taking care to ensure that the fish are not crushed. Q_0 is a baseline level of quality surplus.[24] Note also that a_1 enters negatively in Q, to capture the notion that actions taken to increase the quantity

[23] More specifically, we assume that $0 \leq \delta \leq 1$ and $0 \leq \gamma \leq 1 - \alpha$. Also, note that P could itself be the result of a bargaining process between the boat and a quantity focused firm.

[24] This can be thought of as the amount that suppliers will catch before exerting any costly action, or perhaps more reasonably as the result of some limited contractual agreement that we abstract from.

caught, such as packing the hold tightly with fish, often adversely affect quality.

We assume that neither P nor Q is contractible, but that P—the quantity surplus—is perfectly observable at the time of bargaining and Q—the quality surplus—is not. All parties know the value of Q_0, and because $P = a_1$ is observable, Q in effect has an observable portion: $\tilde{Q} = Q_0 - \gamma a_1 = Q - \delta a_2$.

Integrated suppliers

If a supplier is integrated, the firm has rights to the supplier's catch. However, because the firm cannot write contracts over Q and P, it cannot credibly commit to rewarding the supplier's actions. As a result, the supplier chooses $a_1 = 0$ and $a_2 = 0$, and the total surplus is simply Q_0.

Independent suppliers

Although neither Q nor P is contractible[25], the firm and supplier may bargain ex-post over the price of the delivery. We assume a Nash bargaining concept, with the supplier's bargaining coefficient equal to α. Because the supplier can always deliver its catch to the alternative quantity focused firm and receive P, the supplier must always receive at least P. The supplier additionally receives a share α of the observable portion of the surplus $\tilde{Q} - P$ that accrues to the firm: $\alpha(Q_0 - \gamma P - P)$. As a result, an independent supplier solves the problem:

$$\max_{a_1, a_2} \alpha Q_0 + (1 - \alpha\gamma - \alpha)a_1 - \frac{1}{2}a_1^2 - \frac{1}{2}a_2^2$$

This gives: $a_1 = (1 - \alpha\gamma - \alpha), a_2 = 0$, and social surplus is

$$Q_0 - \gamma(1 - \alpha\gamma - \alpha) - \frac{1}{2}(1 - \alpha\gamma - \alpha)^2 < Q_0$$

[25] Alternatively, we could assume that only a portion of Q and P is non-contractible, and that we consider only this portion as in Baker et al. (2002).

Because of the counterproductive actions to increase quantity $(a_1 > 0)$, and the adverse effects of those actions on the quality surplus, the surplus is lower when the suppliers are independent. As a result, the more efficient organizational structure to produce quality is vertical integration.

It is worth noting that a number of assumptions made in this model are not strictly necessary to get this result. The relative efficiency of integration holds whether or not quantity focused actions directly negatively impacts the quality surplus (because of the inefficiency of quality actions), and would hold even more strongly if, for example, there were complementarities in the costs of quality and quantity actions.

2.3.3 Discussion

The theoretical role of vertical integration is a contentious topic. Our framing follows Baker et al. (2001, 2002) in combining elements of the incentives based theories in the tradition Holmstrom and Milgrom (1991) and the property rights theories in the vein of Grossman and Hart (1986). Such a framing is not the only type of model that would produce a relationship between integration and output quality. In actuality, integration is a complex organizational change whose causes and consequences operate through multiple mechanisms. However, because the foundations of the model above depend on a series of salient features of our context—unobservable quality, observable quantity, and alternative buyers that are less concerned with quality—and because we are able to directly test the predictions of the model, we see these alternative theories as complementary to the mechanisms our framework focuses on, rather than contradictory.

Our model presents a highly stylized, and somewhat stark, example to highlight a key intuition: that integration can act as a valuable tool for muting the incentives provided in the open market. We believe this starkness most simply portrays why firms in our context might want to integrate in order to produce high quality output. That said, this oversimplification does have a few drawbacks, most notably the lack of incentive to take quality focused actions, and to take any actions at all when integrated. This is in some sense a strong version of what are sometimes called the drone employees (Gibbons, 2005a) that appear in property rights theories of the firm that follow Grossman and Hart (1986). However, this feature may be easily remedied in more complex models that preserve the basic intuition and result. For example, assuming observability over Q induces quality focused actions among independent suppliers and—for sufficiently small values of δ—does not affect the main result. Perhaps more realistically, introducing dynamics into the model, with long-term relationships between firms and suppliers, creates an environment in which the incentives of the downstream and the upstream parties can be aligned through repeated interactions.

In Appendix B.2, we present and test the empirical implications of exactly such a dynamic model, in which we allow the downstream party to use relational contracts to incentivize the quality action. We posit that Q—the quality surplus—can be observed to the downstream party but only with some lag (e.g. once the inputs are processed and output quality is measurable). The firm can then offer the supplier a (delayed) reward contingent on this surplus, but can only credibly promise to pay this reward if it interacts repeatedly with the upstream party. In this context, we show that the value of the relationship can incentivize the supplier to take the first best actions, but that this sort of relational contract may be difficult to sustain if the supplier is independent. The intuition for this result is similar to our static baseline: independent suppliers own the rights over the inputs, and when the value of these inputs in their alternative use is high, they face incentives to renege on the relational contract and sell the goods in their alternative use.

Our model above also implicitly demonstrates the *costs* of integration. The market provides strong incentives for quantity, and for a low quality firm that is aiming to produce quantity, integration would only interfere with and lessen the strength of these incentives. Accordingly, quantity focused firms prefer independent suppliers. A similarly formulated model, with the roles of high and low quality firms switched (e.g. $P >> Q$), provides precisely this result.

In our stylized model, firms are either quality-oriented or not. In reality, a firm's output objectives are likely a combination of quality surplus and quantity surplus in which the weight attached to each depends on the demand the firm faces at a particular point in time. In this case, firms should not source all inputs from either integrated or non-integrated suppliers, but choose an intermediate organizational structure—that is, an intermediate level of integration—that depends on the relative importance of Q and P in the firm's *current* objective function.

The framework presented in this section motivates three empirical predictions that we test in the remainder of the chapter:

1. Firms' organizational structure responds to variation in the *relative* profitability of producing high quality output. An increase in the quality premium—for example due to increased demand for high quality grades—leads to more integration, that is, a higher share of inputs from sourced from integrated suppliers.

2. The reason is that the actions of a supplier differ when the supplier is integrated. In particular, suppliers that get integrated *reduce* their effort to produce quantity, especially in ways that benefit quality.

3. As a result, the degree to which a firm or plant uses integrated suppliers affects output quality. Firms that use inputs from integrated suppliers produce higher quality output.

2.4 Data, Variables, and the Relationship of Interest

2.4.1 Data

The primary datasets we use to test our three predictions are the following:

Plant production. Administrative data on all plants' production come from Peru's Ministry of Production, which regulates the fishmeal industry. Every month plants are required to submit information on how much prime (high quality) and fair average (low quality) fishmeal they produce. Quality grade is thus directly reported in the plant production data, and subject to auditing by government inspectors. As discussed in Sub-section 2.2.2, the distinction between prime and fair average quality is based on the fishmeal's protein content. From these records, we construct each individual plant's and each firm's "high quality share of production" in a given month or production season.

Plant registry. We link the production data with an administrative plant registry that contains monthly information on each plant's (i) technological production capacity and (ii) owner, typically a multi-plant fishmeal firm.[26] We also use this registry to link the production data to export data. We can do so for almost all firms, but not the smallest firms, which use intermediaries to export.

Export transactions. Detailed data on the universe of fishmeal exports at the transaction level come from Peru's customs authority. We observe the date of the transaction, the export port, the destination country, the weight of the fishmeal, the value of the transaction, and the exporting firm.

Internal data from a large firm. One of the largest fishmeal firms in Peru shared its internal sales records with us. The firm owns many plants along the coast. The sales records include information on the shipment's packing, its free-on-board value, the price per metric ton, the buyer, destination country, date of the contract, and the terms. Most importantly for our purposes, the specific plant that produced a given shipment of fishmeal is reported.

Supply transactions. The Ministry of Production records all transactions between the fishmeal plants and their suppliers of raw materials, i.e. fishing boats. Information on the date of the transaction, the boat, the plant, and the amount of fish involved (though not the price), is included.

Boat registry. We merge the supply transactions data with an administrative boat registry that provides information on a boat's owner, the material the boat is made of,

[26]The data contains information on the number of metric tons that can be produced per hour with currently installed Low and High technology. As very few firms in our sample only have the Low technology, we define a High technology firm as one for which the High technology share of total processing capacity is higher than the median (0.67).

its storage capacity and engine power, and whether it has a cooling system installed.[27]

Boat GPS data. Peruvian fishing boats that supply fishmeal plants are required to have a GPS tracking system installed, and to continuously transmit their GPS signal to the Ministry of Production while at sea. The ministry stores the transmitted information—the boat's ID, latitude, longitude, speed, and direction—each hour on average, and shared the resulting dataset with us.[28]

2.4.2 Variables of interest

Our primary measure of an individual plant's output quality is the share of the fishmeal the plant produces in a given month that is of "prime" quality grades—a direct measure of quality whose interpretation requires no assumptions. We aggregate this measure up to firm level to construct a corresponding measure of a firm's "high quality share of production".

We also construct a granular measure of the average quality grade—protein content—of the fishmeal a firm produces. While we do not directly observe the exact protein content of each export shipment, we can go beyond simply using unit prices and approximate the precise quality grade. This is because we observe quality grade-specific fishmeal prices in detailed (week$\times$export port$\times$protein content level) data recorded by a fishmeal consulting company. We infer the protein content of each of a firm's export shipments by comparing the corresponding unit values to this price data. To construct a firm$\times$season level measure, we average protein content across export shipments, weighting by quantity.[29] A priori, we have little reason to believe that this in-

[27]Information on engine power is only available for 2004-2006. However, changes in engine power are extremely rare in that period, so we treat this characteristic as fixed over time.

[28]Only about half of the observations in the Supply transactions dataset can be matched to a GPS recording. Some boat owners, for example, disappear from the GPS data for a complete calendar year. However, such missingness is unlikely to be of concern for within-boat analysis, the level at which we use the GPS data.

[29]The export transaction records do not report the specific plant that made the fishmeal so the inferred quality grade is only available at the firm level—except for data covering the fishmeal firm that shared internal data with us, including information on the plant that produced a given export shipment. One

ferred protein content measure could be systematically biased.[30] Empirically, it is highly correlated with the "high quality share of production" directly observed for a firm's plants in production data, and with the exact quality grade reported in the sales records of a firm that shared its data with us.

To quantify vertical integration, we consider both the *number of suppliers a firm owns*, and the corresponding *share of inputs used in its production process* that come from integrated suppliers ("Share VI"). Share VI is our preferred measure of integration for a number of reasons. Because we observe all transactions between *plants* and suppliers, we can construct Share VI in a consistent manner for both firms and individual plants. This allows us to move from the across firm comparisons we make in Section 2.5 to the within firm comparisons we make in Section 2.7. Furthermore, Share VI automatically captures suppliers' size, allowing us to avoid assumptions on "scale effects"—e.g. how the benefit of one big integrated supplier compares to two small ones. Finally, Share VI is the more relevant measure when asking whether organizational structure and output quality are causally related: if firms vertically integrate when the quality premium rises because doing so allows them to increase input quality, then it should matter not just if a firm owns suppliers, but the degree to which the firm as a whole and its individual

potential concern is that fishmeal can be stored for a short period, and hence firms could attempt to strategically time their export transactions. In practice the product is almost always sold before the next production season starts. (The reason why inventories are small—between +10 and -10 percent of total season production (see Appendix Figure B.2)—is likely that many contracts are entered into before the production season starts (which helps the fishmeal manufacturers and their foreign buyers reduce demand/supply uncertainty), and because firms' ability to strategically "time" their sales is in actuality limited). A shipment can thus be traced back to a specific production season (but not a specific production month; constructing the inferred protein content measure at month level would require an assumption about how firms manage their inventories—for example, first-in-first-out versus first-in-last-out). A related concern is that firms that are about to end operations and close down might sell off their fishmeal, in which case a lower unit price might not reflect lower quality but rather a "going-out-of-business" discount. We thus exclude data from any firm×season observations that correspond to a firm's last season producing and exporting fishmeal, but the results are robust to including these observations. These issues are not relevant for our directly observed "high quality share of production" measure of output quality.

[30]Fishmeal is a vertically differentiated but otherwise homogenous product, and price differentials across shipments of a given quality level (and across firms producing a given quality level) in a given time period are negligible (see Sub-section 2.2.2). This implies that pricing-to-market, bulk discounts, etc, are not a concern.

plants actually *source inputs* from those suppliers at the time of production. Recall from Sub-section 2.2.3 that Peruvian fishmeal manufacturers generally use integrated suppliers to capacity over the course of a season, so in the context we study, Share VI and supplier ownership are closely related measures of integration.

2.4.3 Relationship of interest

In Section 2.5 we begin our analysis of how exogenous changes in incentives to quality upgrade affect integration decisions. Before doing so, we first demonstrate that the basic relationship predicted by our model holds empirically: integration and output quality are positively correlated. To do so, we estimate regressions of the form:

$$\text{Quality}_{it} = \alpha + \beta_1 \text{VI}_{it} + \beta_2 \text{HighTech}_{it} + \gamma_i + \delta_t + \varepsilon_{it} \tag{2.1}$$

where Quality_{it} and VI_{it} respectively measure the quality of the output produced by firm i in season t and how vertically integrated the firm's organizational structure is in the same season. We control for the technology the firm uses to convert fish into fishmeal,[31] HighTech_{it}, and firm and season fixed effects γ_i and δ_t. We thus estimate *changes* in output quality for those firms that vertically integrate in a given season, relative to other firms that do not. We cluster the standard errors at the firm level.

The results in Panel A of Table 2.3 point towards a strong baseline relationship between owning suppliers and output quality. The estimates in column 3 imply, for example, that moving from the 25th to 75th percentile of number of boats owned is associated with an increase in protein content of just under 10 percent of the range observed in Peru.[32]

[31] A firm's production technology is an important potential determinant of output quality, and one that could plausibly correlate with organizational structure (Acemoglu et al., 2007, 2010). We thus control for installed HighTech_{it}, i.e., steam drying (High) technology. At the firm level, HighTech_{it} is equal to the share of installed capacity that is of the high type.

[32] Firms in the 25th percentile own four boats, while firms in the 75th percentile own 36 boats. $\hat{\beta} \times (\text{asinh}(36) - \text{asinh}(4)) = 0.43$. The range of protein content observed in Peru is approximately 63-68 percent.

In Panel B we show that, beyond simply owning suppliers, what matters for output quality is Share VI: *the share of a firm's supplies coming from integrated suppliers at the time of production*. The results imply that a firm that uses inputs coming entirely from integrated suppliers rather than inputs entirely from independent suppliers sees a share of high quality output that is 50 percent higher, and an average protein content that is higher by about 20 percent of the range observed in Peru.

In Panel C we show that these results are not not driven by observable, time-varying supplier or firm characteristics. We control for the firm's share of total industry production and a series of supplier characteristics. Doing so has little impact on the estimated coefficient.[33]

In Appendix B.2, we consider whether the relationship between output quality and integration might be the result of long-term supplier-firm relationships, rather than ownership per se. This does not appear to be the case, as we do not observe the association between quality and the share of inputs coming from suppliers in long-term relationships that we do for Share VI. In other words, it is integration itself, not the relationship, that co-varies with output quality. This is in line with the predictions of a dynamic version of our model, also shown in Appendix B.2.

The relationship between a firm's organizational structure and its output quality that we established in this sub-section is the starting point of our empirical analysis. This basic relationship is consistent with this paper's hypothesis. However, it is also consistent with the alternative theories discussed in the introduction. In these theories, a correlation between integration and output quality arises, but the relationship is either not causal or not known to (or ignored by) firms. We rule out such explanations in the next

[33]Controlling for the share of inputs coming from steel boats, high capacity boats, and boats with a cooling system leaves the magnitude and significance of the coefficient on share of inputs coming from VI suppliers essentially unchanged. Note that two of the supplier characteristics variables included—Share of inputs from high capacity boats and Share of inputs from boats with cooling system—*are* significantly correlated with output quality *in the cross-section* of firms. One reason why the coefficients on these characteristics are not significant is that we observe little change in these boat characteristics over time. Controlling for the firm's share of total industry production also leaves the magnitude and significance of the coefficient on Share VI essentially unchanged.

section.

2.5 The Quality Premium and Organizational Structure

We now show that the relationship between output quality and vertical integration we established in the previous section reflects an explicit organizational choice firms make *in order to* "climb" the quality ladder. Specifically, Peruvian fishmeal manufacturers integrate suppliers when the returns they earn from upgrading quality rises for exogenous reasons. This finding provides empirical support for the prediction that a vertically integrated organizational structure is efficient for producing high quality output. We provide additional evidence that makes clear why this finding is difficult to reconcile with alternative theories.

2.5.1 Estimating how the quality price premium affects vertical integration

We begin by showing that firms as a whole *do* vertically integrate when the quality premium is high and de-integrate when the quality premium is low. We quantify Integration_{it}—a firm's decision to integrate (de-integrate)—as a season-to-season increase (decrease) in the share of inputs the firm obtains from integrated suppliers.[34] We measure the quality premium as the difference between the (log) price of high and low quality fishmeal.

We first show results from simple descriptive regressions of the form:

$$\text{Integration}_{it} = \alpha + \beta \text{QualityPremium}_t + \eta_i + \varepsilon_{it}. \tag{2.2}$$

Here, $\beta > 0$ indicates that firms vertically integrate when the quality premium is high, and η_i represents a firm fixed effect.

[34] Specifically, if VI_{it} is defined as Share VI, we analyze $\text{VI}_{it} - \text{VI}_{it-1}$, where t indicates a season. In Appendix Table B.3 we show all of our results using VI_{it}, rather than the difference, with qualitatively similar results.

The first column of Table 2.4 shows the results from this specification, omitting firm fixed effects. There is a strong baseline correlation between the quality premium and firm integration. The second column shows that this relationship holds both when the quality premium increases and when it decreases. Firms vertically integrate when the quality premium is high and de-integrate when the premium is low. In this specification, we replace QualityPremium_t with two dummy variables: an indicator equal to one when the quality premium is above the mean, and an indicator equal to one when the quality premium is below the mean.[35] The signs on the two coefficients suggest that firms respond to both high and low levels of the quality premium, but the coefficient on the high quality premium indicator is larger and significant.

In column 3 we show that firms' response to the quality premium is not due to associated income shocks or general incentives to expand or reduce production. We repeat the regression from column 1, but with the Log(Average Price)—the average fishmeal price in season *t*—replacing the quality premium as the regressor of interest. The estimated coefficient is near zero and insignificant, indicating that firms are not more likely to vertically integrate when the overall price level is high. Figure 2.2 highlights that the quality premium in Peru is at most weakly correlated with average prices in Peru. It is thus not surprising that firms respond differently to the two. Column 4 includes the firm fixed effect, with effectively identical results to column 1. In sum there is a strong relationship in the time series between the quality premium and vertical integration.

We next consider how the quality premium differentially impacts firms' integration decisions depending on their *capacity* to upgrade quality. Firms that are already producing exclusively high quality output have no scope to upgrade the average quality they produce. For these firms, an increase in the quality premium should not lead to a change in organizational structure. Conversely, a firm that produces some mix of high and low quality output has capacity to raise quality. If vertical integration indeed enhances output quality, we expect to see these firms integrate when the quality premium

[35]We omit the constant in this specification.

rises.

To investigate this differential response, we interact QualityPremium_t with two measures of a firm's capacity to upgrade. First, in each season t, we consider a binary indicator for the firms who produced exclusively high quality output in season $t-1$ (*High Quality Producers*) versus those who produced at least some low quality output in $t-1$ (*Low Quality Producers*). Second, we explicitly consider each firm's share of low quality production in season $t-1$, which we refer to as the *Upgradable Share of Production*. For example, a firm currently producing 25 percent high quality output has a 75 percent *Upgradable Share of Production*. We consider specifications of the form:

$$\begin{aligned} \text{Integration}_{it} = \alpha + \beta(\text{UpgradingCapacity}_{it-1} \times \text{QualityPremium}_t) \\ + \gamma\text{UpgradingCapacity}_{it-1} + \eta_i + \delta_t + \varepsilon_{it}. \end{aligned} \tag{2.3}$$

where $\text{UpgradingCapacity}_{it-1}$ refers to either *Low Quality Producer* or *Upgradable Share of Production*. The approach in (2.3) is a generalized difference-in-differences in which firms that are more versus less exposed to changes in quality upgrading incentives are compared in each of 13 different production seasons, and in each of these 13 seasons the quality premium may be relatively high or relatively low. The season-to-season variation in the quality premium is shown in Figure 2.2. While the long-term trend is weakly positive during our data period, the quality premium fluctuates substantially from season to season, sometimes rising and other times falling. We control for $\text{UpgradingCapacity}_{it-1}$ itself, firm and production season fixed effects (γ_i and δ_t), and cluster the standard errors at the firm level.[36] Our hypothesis implies that $\beta_1 > 0$.

Columns 5 and 6 of Table 2.4 show OLS results from these specifications for *Low Qual-*

[36]The fact that we control for the characteristic that defines a firm's exposure to a "treatment" variable that varies across time—here $\text{UpgradingCapacity}_{it-1}$—distinguishes our approach from traditional Bartik instrument approaches (Goldsmith-Pinkham et al., 2017). Note also that firms' upgrading capacity evolves over time. A strength of our approach is thus that the high and low Upgrading Capacity_{it-1} firms being compared across one rise or fall in the quality premium may differ from those being compared across another rise or fall.

ity Producer and *Upgradable Share of Production*, respectively. Under both definitions, we find that firms with greater scope to shift from low to high quality production are more likely to vertically integrate when the quality premium is high—consistent with our hypothesis and the model in Section 2.3. It is worth noting that if—not implausibly—firms producing a high share of high quality output face a lower marginal cost of either quality upgrading or vertical integration, then such a countervailing force would if anything *strengthen* the support for our hypothesis implied by the results in columns 5 and 6 of Table 2.4.[37]

A potential concern is that some omitted factor might be influencing both Peruvian firms' incentive to integrate and the quality premium (or, alternatively, that the quality premium is itself influenced by integration decisions). To address this concern, we develop an instrumental variable strategy based on quantities produced by other top fishmeal producing countries. The total quantities these other countries produce are ideal instruments because they are driven by *aggregate fishing quotas* set by each country's (and European) regulatory authorities. Aggregate fishing quotas are set based on sustainability considerations so the quantities of fishmeal these other countries produce are unlikely to correlate with factors influencing Peruvian firm's integration decisions, except via their influence on market prices.[38] In addition, the production volumes of countries that specialize in high or low quality grades generate meaningful variation in the quality premium in Peru—and we know from column 2 of Table 2.4 that any impact these volumes may ultimately have on integration decisions in Peru does not arise

[37]Suppose firms producing a higher share of high quality output, in addition to their mechanically lower *scope* for further quality upgrading, also face a lower marginal cost of quality upgrading. This would only be a concern for our strategy of this lower cost of upgrading was also related to those firms' ease of integration. (Similarly, a potential lower marginal cost of integration for firms producing a higher share of high quality output would only be a concern if the lower cost of integration was also related to those firms' ease of upgrading). The logic of our approach—the argument that firms with high *scope* for quality upgrading face stronger incentives to upgrade quality when the quality premium is high—would then need to not only hold, but to outweigh any such countervailing forces for this empirical strategy to yield evidence supporting our hypothesis.

[38]Chile, Denmark and Iceland have had aggregate fishing quotas in place for the relevant fish species throughout our sample period (IFFO, 2014; Tanoue, 2015; IRF, 2017; European Commission, 2018), while Thailand introduced such a system in 2015.

through the general price level.

Specifically, our instruments are the quantities exported in each production season by the top 5 fishmeal exporters, excluding Peru: Chile, Denmark, Iceland, and Thailand, all of which specialize in high or low quality grades.[39] We interact these instruments with Upgrading Capacity_{it-1} to instrument for Upgrading $\text{Capacity}_{it-1} \times$ Quality Premium_t as follows:

$$\begin{aligned}\text{Upgrading Capacity}_{it-1} \times \text{Quality Premium}_t &= \beta_2 \text{ Upgradable Share Of Production}_{it-1} \\ &+ \sum_c \beta_c \text{ Upgrading Capacity}_{it-1} \times \text{Quantity}_{ct} + \gamma_i + \delta_t + \varepsilon_{it} \qquad (2.4)\end{aligned}$$

where c is an exporter country, and Quantity_{ct} is quantity of fishmeal exported from country c in season t. The first stage results—shown in Appendix Table B.2—are strong, confirming that the total quantities of high and low quality fishmeal produced by countries with aggregate fishing quotas impact the quality premium in Peru.

The second stage results with Upgrading Capacity_{it-1} defined as *Low Quality Producer* are shown in Column 7 and as *Upgradable Share of Production* in Column 8. The IV estimates of $\hat{\beta}$ are of very similar magnitude to the OLS estimate in columns 4 and 5. For example, the estimates in columns 7 and 8 both imply that, when the quality premium in Peru rises by 10 percent, a firm with a high upgradeable share of production—one that produces only low quality output—increases its Share VI by about 30 percent when compared to a firm producing only high quality output.

Our instrumental variables strategy rests on the assumption that the quantities produced by other top fishmeal exporting countries affect integration decisions in Peru only through their impact on market prices. If this argument is correct, we would expect these production volumes to manifest themselves in the price of high and low quality fishmeal *locally*, and ultimately in the quality premium in Peru. This is what we find in

[39] Information on total monthly exports from these countries is available from COMTRADE (a dataset described in footnote 52).

columns 9 and 10, where we repeat the strategy in equation 2.4, but instrument using the *quality-grade specific price* in other top exporting countries, rather than the quantities exported themselves. The estimates are again qualitatively similar to the OLS estimates in columns 5 and 6.[40]

In Appendix B.3 we exploit a different form of variation and find results consistent with those discussed in this sub-section. We show that manufacturers respond the same way to variation in firm-specific, quality-differentiated demand shocks as they do to analogous shocks to the quality price premium—integrating suppliers and increasing Share VI when relative demand for high quality grades increases, and selling boats and decreasing Share VI when relative demand for high quality grades decreases. To do so we construct instruments for firm-specific demand shocks that exploit the fact that each importer country tends to import very specific quality grades; that importer countries' relative demand fluctuates over time; and that changes in demand from a given country matter more for firms that previously exported to that country.[41]

2.5.2 Interpretation

The results discussed in this section are consistent with this chapter's hypothesis and the theoretical framework in Section 2.3. In our model, a firm integrates suppliers when its returns to upgrading quality rise because it is difficult to ensure that independent suppliers deliver high quality inputs when the quantity they produce is valued by other buyers in the market. We now consider whether firms' decision to integrate suppliers when the benefits of quality upgrading rise can be explained by alternative theories.

[40]Specifically, we use the *high quality price* in other top exporting countries. Price data is available—from IFFO, an industry association—only for three of the four exporters used in our strategy above, all of which specialize in high quality fishmeal: Chile, Denmark, and Iceland. In Equation 2.4We replace Quantity_{ct} with $\text{HighQualityPrice}_{ct}$, the high quality price in exporting country c in season t. High quality fishmeal prices in Chile, Denmark, and Iceland—shown in Appendix Figure B.4, along with Peru's price—are highly correlated with those in Peru. The first stage results in Appendix Table B.2 are thus strong. (Note that the reason why high quality prices are not exactly the same across countries is that production seasons differ and the distance between countries make it difficult to conduct arbitrage.)

[41]We follow many fruitful applications of such an approach in the trade literature (see e.g. Park et al., 2009; Brambilla et al., 2012; Bastos et al., 2018; Tintelnot et al., 2017).

A first possibility is that firms simultaneously choose their organizational structure and output quality, and shocks—for example to demand (Legros and Newman, 2013; Alfaro et al., 2016)—affect both without the two being directly related. Such a story is difficult to reconcile with the fact that Peruvian fishmeal manufacturers integrate suppliers in response to increases in the *relative* price of high quality output, but not in response to increases in the average price of fishmeal.

The same is true for a second possibility, namely that firms, when the benefits of producing high quality output rise, buy suppliers so as to restrict competitors' access to independent suppliers and thereby capture a higher share of a newly appealing market segment that happens to be the high quality one (Ordover et al., 1990). If such a story explained our results, we should see manufacturers integrating suppliers also when the price of low (or any) quality fishmeal rises—unless integrated suppliers are more useful when producing high quality output (as we conjecture).

A third—and related—possibility is that the integration decisions we observe are driven by supply assurance motives (Macchiavello and Miquel-Florensa, 2016; Martinez-Carrasco, 2017). One supply assurance story—namely that firms integrate suppliers to secure general access to inputs but in the process coincidentally produce higher quality output—cannot explain our findings. Such a story is inconsistent with the fact that manufacturers vertically integrate in response to the relative price of high quality output, but not average prices. Another form of supply assurance—integrating to secure access to suppliers who are incentivized to deliver the high quality inputs that are needed to meet the demand for high quality output—*is* exactly the interpretation we favor.

We conclude that manufacturers vertically integrate when the quality premium rises for exogenous reasons *in order to* produce a higher share of high quality output.

2.6 Firms' Organizational Structure and Supplier Behavior

The model in Section 2.3 predicts that integration is an efficient organizational structure for producing high quality output for a specific reason: because integration weakens suppliers' incentives to maximize quantity in ways that might be detrimental to the quality of the inputs they produce. As a result, we expect to see suppliers reduce behavior that increases quantity but is harmful to quality when integrated.

2.6.1 Estimating how vertical integration affects suppliers' quality-enhancing actions

We analyze three measures of behavior that capture the tradeoff between input quantity and quality: the total quantity supplied, the maximum distance travelled from the delivery port, and the total time the supplier spends at sea on a given trip. The first of these three we observe in supply transactions data, while the second two are constructed from boat GPS data. The total quantity supplied is a direct measure of actions taken by the supplier to increase quantity. However, this variable also relates to input *quality*. This is because the supplier may need to forego quality-increasing actions—such as bringing a lot of ice on board to keep it fresh, not stacking fish high on top of each other to prevent smashing it, etc—in order to bring back a high quantity of fish. The maximum distance travelled and total time spent at sea are chosen because they explicitly capture quality-decreasing actions that will tend to increase quantity. Fish freshness—which depends on the time between catch and delivery—is paramount for the protein content of fishmeal. As the Food and Agriculture Organization of the United Nations puts it, "Freshness of raw material is important in its effect on the quality of the protein in the end product [fishmeal]. The importance of minimizing the time between catching fish and processing, and of keeping the fish at low temperatures by icing [which reduces the amount of fish a boat can fit], has already been mentioned" (FAO, 1986, sub-section 10.1.2). Captains must thus balance traveling further and longer to catch more fish against ensuring freshness. Because all three of these measures of behavior increase quantity but decrease quality, we expect them to decrease post-integration (or increase post-separation).

Our empirical strategy focuses on "switchers". Switchers are suppliers that are either bought or sold by a fishmeal firm during our data period and observed supplying the same plant within the firm in question both before and after the change in status. We include supplier×plant fixed effects and hence compare the behavior of a *specific* supplier within a *specific* relationship before versus after integration (or de-integration).

As discussed in Section 2.2, we observe 103 instances in which a fishmeal firm acquires a supplier that is initially owned independently; 32 instances where a supplier is sold from a fishmeal firm to an independent buyer; and 50 instances where a supplier is sold from one fishmeal firm to another. Conveniently, a subset of our qualifying switches—in which the supplier is observed supplying the firm in question both before and after the change in status—comes from this last set of firm-to-firm supplier transitions. This is because integrated suppliers sometimes supply other fishmeal firms.[42] We exploit these transitions in which an always-integrated supplier's relationship with a specific firm changes below.

We do not observe any significant *changes* in suppliers' characteristics when switching in or out of integration with the plant supplied. Thus, while any average differences between the behavior of independent and integrated suppliers might be attributable in part to boat characteristics,[43] our analysis of *within* supplier changes in behavior is unlikely to be influenced by these attributes. Recall also that we saw in Figure 2.1 that suppliers that get integrated or sold deliver to the acquiring/selling firm 63 percent of the time *before integration* (or after de-integration): integration typically implies a sim-

[42] A firm's output objectives may vary across time within seasons, and fish move around and the location of a catch constrains the set of plants a boat can deliver to. As a result, and as seen in Figure 2.1, Panel (a), integrated suppliers on average deliver to other firms just over 10 percent of the time.

[43] As shown in Appendix Table B.4, the characteristics of integrated suppliers unsurprisingly differ from the characteristics of independent suppliers. On observable features such as the size of the boat, the power of its engine, and whether or not it has a cooling system installed, the average switcher falls in between the average always-independent boat and the average always-integrated boat, but closer to the latter.

ple change in the formal status of the relationship between a firm/plant and a supplier engaged in frequent and continuing interactions.

We estimate regressions of the following form:

$$B_{ijt} = \alpha + \beta \mathrm{I}[\text{VI} \times \text{supplies owner firm}]_{ijt} + \gamma_{ij} + \delta_t + \varepsilon_{ijt} \tag{2.5}$$

where B_{ijt} is a measure of the behavior of supplier i, delivering to plant j, on date t. $[\text{VI} \times \text{supplies owner firm}]_{ijt}$ is an indicator for the supplier being integrated with the plant it delivers to on date t. We include date fixed effects (δ_t) to control for potential date specific effects and Supplier×Plant fixed effects (γ_{ij}) to focus on how integration affects the behavior of a specific supplier supplying a specific plant. We cluster the standard errors at the boat level.

Column 1 of Panel A shows that, when integrated and supplying a parent plant, a boat delivers on average about ten percent less per trip compared to when it supplies the same plant while independent. This result is clearly consistent with integration offering lower powered incentives to produce quantity, and also suggests that integrated suppliers dedicate more of their storage capacity to ice and/or are more concerned with crushing fish. Columns 2 and 3 show that boats fish approximately five percent closer to the port of delivery, and spend on average three percent less time at sea on a trip when integrated with the plant supplied. These results suggest that, when integrated, suppliers reduce costly actions associated with long trips, and bring back fresher fish as a result—as the managers in the industry we interviewed reported to us. (In the words of a prominent executive of Peru's National Fisheries Society: "Independent boats prefer to extend their fishing trips [until] they are at full hold capacity, so as to maximize quantity, and this is not good for fish quality...as a consequence of integration, they must adopt my rules. Things like saying, 'hey, you must offload raw 24 hours after having caught it, at the maximum."' (Authors' translation).)

In our model, integration is defined by asset ownership, as in Grossman and Hart (1986). Indeed, suppliers' change in behavior appears to be the result of integration itself, as opposed to any long term relationship that coincides with integration. In Appendix Table B.8, we show that—absent integration—repeated interactions with the same plant do not lead to a change in quality-increasing actions, consistent with the predictions of the dynamic version of our theoretical framework also shown in Appendix B.2. Thus, while repeated interactions help fishmeal manufacturers and independent suppliers exchange supply and demand assurance (Martinez-Carrasco, 2017), they appear not to offer an alternative way to achieve the change in quality-conducive incentives associated with integration in the context we study.

2.6.2 Interpretation

In this section we have seen that a given supplier supplying a given plant takes more quality-oriented and less quantity-oriented actions when the two are vertically integrated. Our interpretation is that integration dampens high-powered incentives to prioritize quantity over quality that suppliers face on the open market. Other changes in incentives that arise due to integration could also play a role. Perhaps the most plausible possibility is that what constrains suppliers' input quality is not their incentive to prioritize quality but their knowledge of how to do so. If so, firms may be reluctant to "teach" a supplier how to upgrade input quality if the supplier is independent (Pigou, 1912). We can shed some light on the likelihood that such a story explains our results in this section by exploiting the fact that integrated suppliers occasionally deliver inputs to other firms. We analyze the behavior of suppliers that are *always* integrated with a fishmeal firm, but sold from one firm to another during our sample period, and that supply a plant belonging to the acquiring and/or the selling firm both before and after the sale. We thus continue to focus on changes in supplier behavior *within* a supplier×plant pair.[44]

As seen in Panel B of Table 2.5, we find quite similar—even slightly larger—effects

[44] To implement, we run the same specification as in Equation 2.5, but define I[VI×supplies owner firm] to be equal to one if the supplier is (i) always owned by a fishmeal firm, and (ii) currently delivering to its parent firm.

compared to Panel A. If acquired, a supplier changes its behavior consistent with prioritizing quantity less—to the benefit of quality—while delivering to the acquiring firm. This pattern is identical to how previously independent "switchers" change their behavior once integrated, suggesting that a story in which integration enables knowledge transfer from Peruvian manufacturers to their suppliers is unlikely to be the primary explanation behind the difference in supplier behavior when integrated. In other contexts, such knowledge transfers may provide an additional—or the primary—motivation for vertical integration (see Atalay et al., 2014).

The results in Panel B of Table 2.5 also underscore that it is not the case that firms simply choose to integrate suppliers that have already begun changing their behaviors, providing support for the parallel trends assumption that underlies a causal interpretation of the results in Panel A.

Another alternative explanation of the change in supplier behavior when integrated is that our results simply reflect the fact that integrated suppliers face low-powered incentives, the behaviors we see not generating any input quality benefits that manufacturers are aware of and *act on* them. Such a story is difficult to reconcile with this chapter's central finding that firms integrate suppliers when the quality premium rises.[45]

2.7 Vertical Integration and Output Quality

In Section 2.5 we saw that firms vertically integrate when the benefits of shifting from low to high quality production rise. In Section 2.6 we saw that suppliers that get integrated take more input quality-increasing and less input quantity-increasing actions. In this section we show that plants' *output* quality responds to integrating suppliers in exactly the manner we expect if the integration-induced change in supplier behavior improves input quality. This provides empirical support for our model's third prediction, namely that vertical integration is an *effective* organizational strategy for producing

[45] Additionally, such a story would raise a conceptual question: if there is no known input quality benefit, and integration lowers input quantity, then why integrate at all?

high quality output—as the managers in the industry we interviewed reported to us. (In the words of Ricardo Bernales Parodi, Managing Director of Pesquera Diamante, Peru's third largest fishmeal company: From the boat to the factory, and to the commercialization, the flour has quality A, B, C and D. If I only bought from my boats, I would make an effort so that 95 percent would be A and B, and only 5 percent of C and D. But when buying from third parties, I end up with 20 percent of C and D." (Authors' translation).)

We first show that there is a robust relationship between changes over time in the share of inputs *individual plants* obtain from integrated suppliers—Share VI—and changes in their output quality that goes beyond the firm level evidence discussed in Sub-section 2.4.3. We then attempt to isolate shifts in a plants Share VI that occur for exogenous reasons. We show evidence from an IV approach that exploits geographic variation in the local concentration of a particular type of supplier that is prohibited from being integrated by regulation. In sum the results we present suggest that the Share VI-output quality relationship arises because integration increases output quality.

2.7.1 Estimating how vertical integration affects output quality

If integration increases output quality because integrated suppliers deliver higher quality inputs, then the relationship between Share VI and output quality we observe at the firm level should hold at the *plant* level as well. This is what we find in Table 2.6. We repeat regression (2.1) from Sub-section 2.4.3, but now at plant (i) $\times$ month (t) level, the lowest level at which we directly observe output quality.

The sample consists of all 94 plants we observe across Peru. We include plant and month fixed effects and thus focus on variation in Share VI across months within a given plant.[46] The results in columns 1 and 2 of Table 2.6 imply that the share of a plant's output that is of the high quality type would be 8-12 percent higher if its parent firm were to integrate all (relative to none) of the plant's suppliers. We also find the same

[46]We observe whether each plant has any high technology installed so HighTech_{it} is now a dummy variable.

integration-quality relationship across different plants *within the same firm* over time, as shown in Appendix Table B.5. There we use internal data provided to us by a single major firm.[47]

In combination with Table 2.3, the first two columns of Table 2.6 establish a positive, statistically significant, and quantitatively consistent association between Share VI and directly observed output quality at the firm and plant levels. Of course, the fact that these correlations hold for individual plants does not rule out non-causal interpretations. It may be that plant specific shocks, for example to productivity,[48] occur and *independently* affect the quality of a plant's output and the share of the plant's supply coming from integrated suppliers.

To evaluate this alternative explanation, we construct an instrument for a plant's use of integrated suppliers at a particular point in time. We use the local presence of wooden fishing boats—which are, by law, independently owned—as a source of variation in a plant's Share VI. These, and other independent boats, move up and down the coast as a function of weather, presence of fish, and other factors. The logic of our instrument is simply that, at times when there happens to be an abundance of independent suppliers in a given area for exogenous reasons, firms are more likely to use those suppliers. A plant's choice of suppliers is the result of a complex optimization process involving output quality objectives on the one hand and the relative cost of using integrated versus independent suppliers on the other. At times when input from independent suppliers is relatively cheap, optimizing plants will tend to decrease their Share VI—even holding their incentives to produce quality constant. When independent suppliers are scarce, the cost of their inputs is likely to be high, and vice versa. This suggests that measures of the presence of independent suppliers may serve as instruments for a plant's Share

[47]The firm's data reports which plant produced the fishmeal included in a given export shipment. In addition to "share high quality", for this firm's plants we can thus measure output quality also as the fine-grained quality grade inferred from exports unit values and auxiliary price data, as we do for firms in columns 3 and 4 of Panel B of Table 2.3. The magnitude and significance of the estimates are very similar to those in Panel B of Table 2.3.

[48]Another example of a shock that may affect different plants within a firm differently is El Niño, which hit Peru in late 2009.

VI.

With this in mind, we consider the number of wooden—and hence independent by law—suppliers active in a port (cluster of plants) in a given month as a proxy for the relative cost of using independent suppliers. Of course, a plant's quality objectives may themselves influence independent suppliers' whereabouts. The plant may for example request deliveries from independent suppliers. We thus use a *leave-firm-out* measure of the presence of independent-by-law suppliers in a given port *during a given period*. In particular, our instrument for Share VI is the number of wooden boats present, excluding any that supply the firm to which the plant in question belongs. We also show results for an analogous instrument using all independent suppliers, not restricting to wooden boats.

The first stage, shown in Appendix Table B.6, is strong: the number of wooden (or independent) boats supplying other plants in the port is highly correlated with the share of integrated supply to the plant in question during the same period. The sign is negative, suggesting that—even using our leave-out proxy—the availability of independent suppliers influences Share VI in the manner we expect. A plant substitutes towards integrated suppliers when independent suppliers are relatively scarce, and vice versa.

Results from the IV specifications are in columns 3-6 of Table 2.6. The IV estimates are of the same sign, statistical significance, and general magnitude as the corresponding OLS estimates, only slightly bigger. This holds whether we restrict attention to suppliers that are independent by law or include all independent suppliers. Additionally, the same is true in a similar specification show in Appendix Table B.5, which utilizes internal data from the firm that shared its data with us.

Might the composition of neighboring plants' suppliers correlate with the quality of a given plant's output for other reasons than having comparable access to independent suppliers? A time-varying, *port level* component of output quality that correlates with our instrument for other reasons than independent suppliers' inputs lowering output quality is a possible concern. However, beyond the presence of independent suppliers, we find no evidence of a relationship between changes in output quality across different plants within the same port.[49] This result, in combination with the sign we find on the first stage—greater presence of independent suppliers increases use of such suppliers—and the second stage—use of independent suppliers lowers output quality—suggests that our instrument's exclusion restriction holds.[50]

In this section we began by documenting that the firm level relationship between inputs coming from suppliers that are integrated at the time of production and output quality holds also at the plant level, including within firms. We then showed that instrumenting for Share VI yields the same positive, estimated relationship with the quality of a plant's output as OLS regressions.

2.7.2 Interpretation

Our interpretation of the results in this section is that access to inputs from integrated suppliers directly increases output quality because a manufacturer can incentivize suppliers to engage in less quality-decreasing behavior once the suppliers are integrated. This follows the model in Section 2.3, and is consistent with the results in Section 2.6. A priori, output quality may of course co-vary with organizational structure without necessarily reflecting a causal relationship. Perhaps the most plausible non-causal links between quality upgrading and integration—for example, that growing firms both produce higher quality output and acquire more suppliers for independent reasons—are ruled out as explanations of our findings by the simple OLS regressions in Table 2.6 and Appendix Table B.5: output quality correlates with *use* of integrated suppliers at the

[49]For example, consider a regression of the share of high quality output at the plant level on the average share of high quality output of other plants in the port, controlling for month and plant fixed effects, as well as the presence of independent suppliers. If a given plant's output quality and that of other plants were perfectly positively or negatively correlated across time, the coefficient on the average share of high quality output of other plants in the port would be respectively one and minus one. We find a coefficient of 0.04, with a standard error of 0.080.

[50]A priori, it could be that a plant's use of independent suppliers itself affects the number of independent suppliers supplying other plants in the port because firms compete for access to suppliers, or that high fish density near a cluster of plants simultaneously enables plants to produce higher quality fishmeal and attracts independent fishing boats. The first of these scenarios would imply a positive sign on the first stage and the second a negative sign on the second stage—the opposite of what we find.

time of production across *plants*, including within firms. The IV regressions go a step further by documenting that the same relationship holds when we restrict attention to fluctuations in the use of integrated suppliers that is driven by variation in the local presence of independent suppliers.[51]

Combining these findings with those found in Section 2.5, we conclude that it is not the case that higher output quality in vertically integrated Peruvian fishmeal manufacturers is simply an ignored by-product of integration decisions made for other reasons, nor that integration and output quality are causally unrelated in the "minds" of the firms in our sample. In Section 2.5 we showed that one of firms' explicit motives for integrating suppliers is to produce a higher share of high quality output. Our evidence indicates that vertical integration increases output quality and that, as a result, firms integrate suppliers when the quality premium rises.

2.8 Conclusion

Guided by Holmstrom and Milgrom (1991)'s classical ideas and subsequent theories of the firm characterizing how we expect firm boundaries to respond to the multitasking nature of suppliers' work (Baker et al., 2001, 2002; Gibbons, 2005a,b), this chapter identifies an overlooked motivation for and consequence of vertical integration in incomplete contracts settings: downstream firms integrate to be able to produce output of high enough quality to sell to high-paying consumers abroad. Integration allows manufacturing firms to incentivize quality-increasing behavior from existing suppliers and better control input quality.

We first present a simple theoretical framework that captures how suppliers and the downstream firms they supply are expected to behave in sectors where firms produce vertically differentiated goods and contracts are incomplete. The model motivates three predictions that follow logically from each other: on how the quality premium—the

[51]In Table 2.3 we also showed that the firm level relationship between vertical integration and output quality holds when we control for the firm's share of total industry output and supplier characteristics.

difference between the price of high and low quality output—affects firms' choice of organizational structure; how suppliers' behavior changes with integration; and how integration consequently affects output quality.

We test these predictions using transaction level data and direct measures of the quality grades manufacturers produce in Peru's fishmeal industry. We show that, when firms' returns to shifting from low to high quality production rise for exogenous reasons, they acquire more of their suppliers. This strategy appears to be effective because fishing boats change their behavior in a way consistent with delivering fresher fish when they are acquired by the downstream firm they supply—which helps firms produce higher quality fishmeal. Finally, we show that firms ultimately produce higher quality output when their organizational structure is more vertically integrated.

These results are inconsistent with alternative theories in which the integration-quality relationship reflects third factors that affect both firms' choice of structure and products produced without the two being directly related. They are also inconsistent with explanations in which firms integrate for reasons other than quality—for example to assure their own or restrict competitors' general access to inputs—but in the process coincidentally produce higher quality output. Instead, the evidence we present suggests that—while firms vertically integrate for many different reasons—in settings where output quality is vertically differentiated and contracts incomplete, one motive for integration is quality upgrading. That is, integration is an explicit organizational choice made *in order to* "climb" the quality ladder.

A natural next question is the generality of this finding. In Figure 2.3, we plot a proxy for average quality that is available for most exporter countries—the average unit value of manufacturing products exported to the U.S.—against the share of those exports that is imported by "related party" downstream firms located in the U.S. (a measure of vertical integration). The figure shows clear evidence of an upward-sloping relationship between average unit values and related party import shares. The same relationship holds also within product categories.[52] This suggests that our findings reflect an association between vertical integration and manufacturing output quality that tends to hold on average across countries and industries. We find this unsurprising, as theory suggests that integration can help address the contracting problems that are typical when producing high quality goods. Given this—and despite vertical integration *overall* being common in developing countries (Acemoglu et al., 2009; Macchiavello, 2011)—it may thus be that the extent of vertical integration observed among firms in the developing world is actually suboptimally *low*, since upgrading output quality is essential for export-driven economic development. Of course, in a world with perfect contracting, there might be no need for integration. As such, our paper's results conversely imply that improvements in contract enforcement may reduce the need for firms to rely on organizational structure to align their suppliers' incentives.

[52]We show this in Appendix Table B.7. In Figure 2.3, the variable plotted on the y-axis is $\hat{\gamma}_c$ from the regression $\log(\text{unit value})_{cpt} = \alpha_{pt} + \gamma_c + \varepsilon_{cpt}$, where $\log(\text{unit value})_{cpt}$ is the average log unit value of products exported from country c, of HS6 code p, in year t to the U.S.; α_{pt} is a product×year fixed effect; and γ_c is an origin country fixed effect. This regression is estimated using COMTRADE data from BACI (See Gaulier and Zignago (2010) for a description of the data). The variable plotted on the x-axis is $\hat{\delta}_c$ from the regression $\text{Related party share of U.S. imports}_{cpt} = \beta_{pt} + \delta_c + v_{cpt}$, where $\text{Related party share of U.S. imports}_{cpt}$ is the share of products exported from country c, of NAICS code p, in year t to the U.S. that are imported by related parties (usually other units of the same firm (Ruhl, 2015)); β_{pt} is a product×year fixed effect; and δ_c is an origin country fixed effect. This regression is estimated using data from the U.S. Census Bureau.

Table 2.1: Summary statistics

		Mean	Sd
Firms	Total number of firms in sample	37	
	Export shipment (metric tons)	380	(351)
	Export Price ($/metric ton)	1454	(303)
	Number of destinations per season	7.05	(5.30)
	Number of export transactions per season	85	(99)
Plants	Total number of plants in sample	94	
	Has high technology	0.85	(0.36)
	High quality share of production	0.85	(0.35)
	Monthly production (metric tons)	3116	(3266)
	Processing capacity (metric tons/hour)	106	(54)
Boats	Number of boats operating per season	812	92
	Fraction owned by a downstream firm per season	0.28	(0.45)
	Fraction of boats made of steel per season	0.44	(0.50)
	Storage capacity (m3)	187	(165)
	Power engine (hp)	432	(343)
	Number of fishing trips per season	24.6	(13.3)
	Number of delivery ports per season	3.49	(1.90)
	Offload weight (metric tons) per trip	110	(110)
	Time at sea per trip (hours)	20.85	(9.96)
	Max. distance from the plant's port (kms)	76	(46)

Notes: This table gives summary statistics over our sample period. *Has high technology* is a dummy equal to 1 if the plant is equipped with steam drying technology. Plants' *processing capacity* measures the total weight of fish that can be processed in an hour. *Steel* is a binary variable equal to 1 if a boat is a steel boat (which tend to be bigger, better suited for industrial fishing, and are subject to different regulations). *Offload weight per trip* is the amount fished and delivered to a downstream firm on each trip. *Time at sea per trip* is the total time spent at sea on a fishing trip. *Max. distance from the plant's port* is the maximum distance between the boat and the port it delivers to on any trip.

Table 2.2: Summary statistics on Integration

Panel A: Boat purchases and sales	
Total number of steel boats registered	741
Number of steel boat transactions	317
Number of transactions Indep. $\rightarrow$ VI	103
Number of transactions VI $\rightarrow$ Indep.	32
Number of transactions VI $\rightarrow$ VI	50
Number of transactions Indep. $\rightarrow$ Indep.	132

Panel B: Decomposition of the growth rate of Share of inputs from VI suppliers

$$\text{Growth (Share VI)}_{i,t} \approx \log\left(\frac{\text{Share VI}_{i,t+1}}{\text{Share VI}_{i,t}}\right) = \log\left(\frac{\frac{\text{VI}_{i,t+1}}{\text{Total}_{i,t+1}}}{\frac{\text{VI}_{i,t}}{\text{Total}_{i,t}}}\right) = \underbrace{\log\left(\frac{\frac{\text{VI}_{i,t+1}}{\text{Total}_{t+1}}}{\frac{\text{VI}_{i,t}}{\text{Total}_{t}}}\right)}_{A} - \underbrace{\log\left(\frac{\frac{\text{Total}_{i,t+1}}{\text{Total}_{t+1}}}{\frac{\text{Total}_{i,t}}{\text{Total}_{t}}}\right)}_{B}$$

	Total	A (Boats purchases or sales)	B (Buying less from Indep.)
Growth	2.1%	1.4%	0.7%
Relative Contribution		67%	33%

Notes: Panel A displays basic statistics on boat purchases and sales. In Panel B, the growth rate of "Share $\text{VI}_{i,t}$" – the share of the inputs sourced by firm i during production season t that comes from vertically integrated suppliers – can be decomposed as presented in the first row of this table. $\text{VI}_{i,t}$ and $\text{Total}_{i,t}$ is respectively the *amount* of inputs firm i sources from vertically integrated suppliers and in total during season t, and Total_t is the total amount of inputs sourced by the industry as a whole during season t. Term A can then be interpreted as the contribution to the growth rate of Share $\text{VI}_{i,t}$ that comes from increasing solely the (relative) amount of inputs coming from integrated suppliers. Since boats fish all their individual quota during the course of a season, the only way to increase (decrease) this term is by acquiring (selling) suppliers. Term B can be interpreted as the contribution of a firm decreasing the (relative) amount of inputs sourced from all suppliers. The table gives the growth rate of "Share $\text{VI}_{i,t}$", Term A and Term B.

Table 2.3: Output quality and vertically integrated suppliers

Panel A: Output quality and number of suppliers owned

Dep. var:	High Quality share of prod.		Protein content	
	(1)	(2)	(3)	(4)
Asinh(Number of suppliers owned)	0.056	0.043	0.197**	0.167**
	(0.060)	(0.042)	(0.083)	(0.066)
High technology share of capacity	No	Yes	No	Yes
Season FEs	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.75	0.75	65.7	65.7
N	275	275	208	208

Panel B: Output quality and Share of inputs from VI suppliers

Dep. var:	High Quality share of prod.		Protein content	
	(1)	(2)	(3)	(4)
Share of inputs from VI suppliers	0.377**	0.375**	0.982***	1.040***
	(0.166)	(0.164)	(0.293)	(0.283)
High technology share of capacity	No	Yes	No	Yes
Season FEs	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.75	0.75	65.7	65.7
N	275	275	208	208

Panel C: Output quality and Share of inputs from VI suppliers

Dep. var:	High Quality share of prod.			Protein Content		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of inputs from VI suppliers	0.313*	0.373**	0.313*	1.153**	0.936***	1.190**
	(0.159)	(0.148)	(0.159)	(0.457)	(0.260)	(0.461)
Share of inputs from steel boats	−0.098		−0.092	−1.152		−1.153
	(0.164)		(0.161)	(0.808)		(0.794)
Share of inputs from boats with high capacity	0.152		0.139	0.988		1.036
	(0.166)		(0.164)	(1.003)		(0.976)
Share of inputs from boats with cooling system	0.191		0.202	−0.232		−0.282
	(0.123)		(0.124)	(0.979)		(0.986)
Share of industry's production		−0.807	−0.748		1.799	1.889
		(0.925)	(0.891)		(3.869)	(3.930)
High technology share of capacity	Yes	Yes	Yes	Yes	Yes	Yes
Season FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.75	0.75	0.75	0.75	65.7	65.7
N	275	275	275	208	208	208

Notes: One observation is a firm during a production season. *High Quality share of prod.* is the share of a firm's total production during a fishing season that is reported as high quality ("prime") output. *Protein content* is the quantity weighted average of a measure of quality inferred from a database that provides weekly prices by quality. *Share of inputs from VI suppliers* is the share of a firm's inputs that come from VI suppliers during a season. Steel boats tend to be bigger, better suited for industrial fishing, and are subject to different regulations. High capacity boats are boats whose hold capacity is in the upper quartile of the distribution. Boats without integrated cooling system must use ice to keep fish fresh. *High technology share of capacity* controls for the share of the firm's total processing capacity (measured in metric tons per hour and averaged across all active plants within the firm) that uses steam drying technology. Standard errors clustered at the firm level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.4: Vertical integration and the Quality price premium

Dep. var:	Share of inputs from VI suppliers (t) - Share of inputs from VI suppliers (t-1)									
	OLS	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Quality premium	0.248**			0.249**						
	(0.089)			(0.095)						
Quality premium is high		0.018**								
		(0.007)								
Quality premium is low		−0.008								
		(0.008)								
Log(average price)			0.040							
			(0.052)							
Low quality producer (t-1)					0.784*		0.767		0.831*	
× Quality premium					(0.451)		(0.479)		(0.434)	
Upgradable share of production (t-1)						2.151***		1.724*		2.430***
× Quality premium						(0.584)		(0.932)		(0.436)
Low quality producer (t-1)	No	No	No	No	Yes	No	Yes	No	Yes	No
Upgradable share of production (t-1)	No	No	No	No	No	Yes	No	Yes	No	Yes
Season FEs	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	191	191	191	191	191	191	191	191	191	191
Kleibergen-Paap LM p-value							0.05	0.01	0.06	0.01
Kleibergen-Paap Wald F statistic							78	178	30	471

Notes: One observation is a firm during a production season. *Share of inputs from VI suppliers (t) - Share of inputs from VI suppliers $(t-1)$* is the change between season $t-1$ and season t of the share of inputs sourced from integrated suppliers. As shown in Table 2.2, most of the variation in Share VI is driven by acquisition or sales of suppliers. *Quality premium* is equal to $Log(\text{High Quality}) - Log(\text{Low Quality})$ where High and Low Quality are the average price of "Prime" and "FAQ" fishmeal in the month preceding the current fishing season. We choose to take the month preceding the fishing season rather than the fishing season itself as integration decisions are typically decided in the month preceding the season and integration within a season is extremely rare in the data. *High Quality premium* (*Low Quality premium*) is equal to 1 is the Quality Premium is above (below) the sample average value. $Log(\text{average price})$ is the Log of the average price of Peruvian fishmeal, again computed in the month preceding the current fishing season. Low quality producer$(t-1)$ is equal to 1 if a firm's share of low quality output in the preceding season was at least 1 percent. Upgradable share of production$(t-1)$ is the share of a firm's production that was of low quality in the previous season. A firm that produces almost only low quality output has more potential to upgrade than a firm already producing almost only high quality output. In Columns 7 and 8, the instruments are interactions between Low quality producer$(t-1)$ or Upgradable share of production$(t-1)$ and the quantity produced by top fishmeal exporters. In Columns 9 and 10, the instruments are interactions between Low quality producer$(t-1)$ or Upgradable share of production$(t-1)$ and the price of fishmeal in other top fishmeal exporting countries. Standard errors clustered at the firm level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.5: Supplier behavior and vertical integration

Panel A: Identified from all switchers (Idependent to VI, VI to Independent and VI to VI)

Dep. var:	Log(Quantity supplied)	Log(Max. distance the plant's port)	Log(Total time spent at sea)
	(1)	(2)	(3)
I[VI × supplies owner firm]	-0.096^{***}	-0.054^{***}	-0.030^{*}
	(0.023)	(0.019)	(0.016)
Date FEs	Yes	Yes	Yes
Supplier × Plant FEs	Yes	Yes	Yes
N	315,442	137,278	159,724

Panel B: Idenfified only from VI switchers changing ownership (VI to VI)

Dep. var:	Log(Quantity supplied)	Log(Max. distance the plant's port)	Log(Total time spent at sea)
	(1)	(2)	(3)
	(1)	(2)	(3)
I[Always VI × supplies owner firm]	-0.147^{***}	-0.082^{***}	-0.073^{***}
	(0.027)	(0.026)	(0.023)
Date FEs	Yes	Yes	Yes
Supplier × Plant FEs	Yes	Yes	Yes
N	315,442	137,274	159,724

Notes: One observation is a boat during a fishing trip. *Quantity supplied* is the amount of fish the boat delivers to the plant per trip. *Max. distance from the plant's port* is maximum distance a specific boat is observed away from port. *Max. distance from the plant's port* can only be measured if the boat leaves from and arrives at the same port. *Total time at sea* is the amount of time the boat is away from port per trip. The number of observations varies from one column to the next as GPS variables for a given trip are sometimes missing. In panel A, we define I[VI×supplies owner firm] to be equal to one if the supplier is (i) currently vertically integrated (ii) currently delivering to its parent firm. In panel B, we define I[Always VI×supplies owner firm] to be equal to one if the supplier is (i) always owned by a fishmeal firm, and (ii) currently delivering to its parent firm. Because we include Supplier × Plant FEs, I[VI×supplies owner firm] and I[Always VI×supplies owner firm] are identified based only on suppliers who change ownership during our sample period. Standard errors clustered at the boat level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.6: Output quality and share of inputs from vertically integrated suppliers

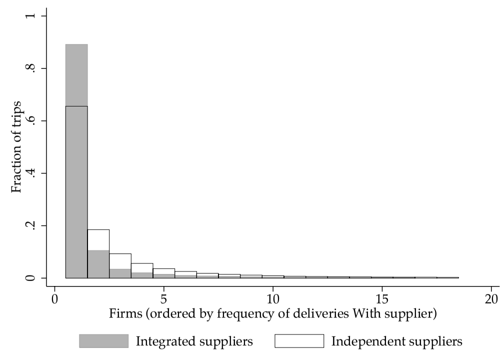
	Impact of Share of VI Inputs on Quality					
Dep. var:	High Quality Share of Production					
	OLS	OLS	IV: Ind. Boats		IV: Wooden Boats	
	(1)	(2)	(3)	(4)	(5)	(6)
Share of inputs from VI suppliers	0.102^{**}	0.064^{**}	0.213^{***}	0.169^{**}	0.160^{***}	0.142^{**}
	(0.038)	(0.030)	(0.080)	(0.068)	(0.060)	(0.060)
Has high technology	No	Yes	No	Yes	No	Yes
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Plant FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.85	0.85	0.85	0.85	0.85	0.85
N	2647	2647	2487	2487	2647	2647

Notes: One observation is a plant in a particular month. *High Quality share of production* is the share of a firm's total production during a fishing season that is reported as high quality "prime" output. *Share of inputs from VI suppliers* is the share of a firm's inputs that come from VI suppliers during a season. *Has high technology* is a dummy variable equal to one if the plant in question has any steam drying technologies installed. Columns 3 and 4 instrument for *Share of inputs from VI suppliers* with the number of independent boats present locally (in the plant's port) in the season in question, excluding those that interact directly with the plant itself. Columns 5 and 6 instrument for *Share of inputs from VI suppliers* with the number of wooden boats present locally (in the plant's port) in the season in question, excluding those that interact directly with the plant itself. Standard errors are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

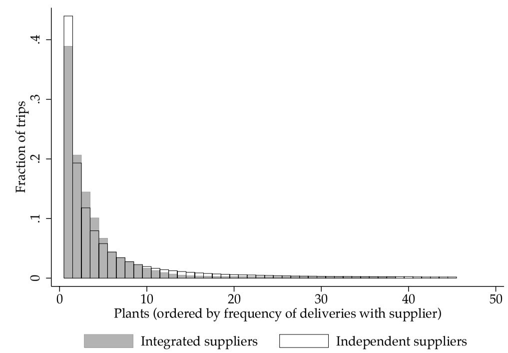
Figure 2.1: Interactions between manufacturers and integrated and independent suppliers

All Suppliers

(a) Distribution of interactions across firms

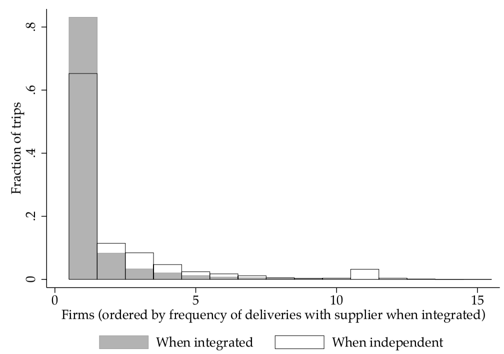


(b) Distribution of interactions across plants

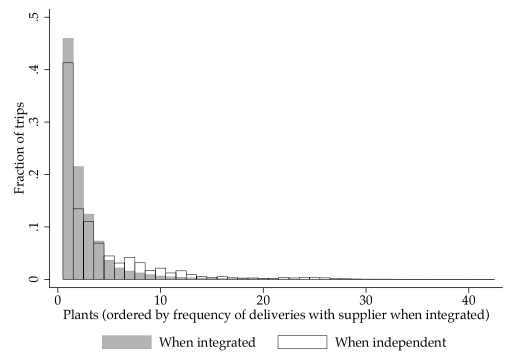


Switchers Only

(c) Distribution of interactions across firms



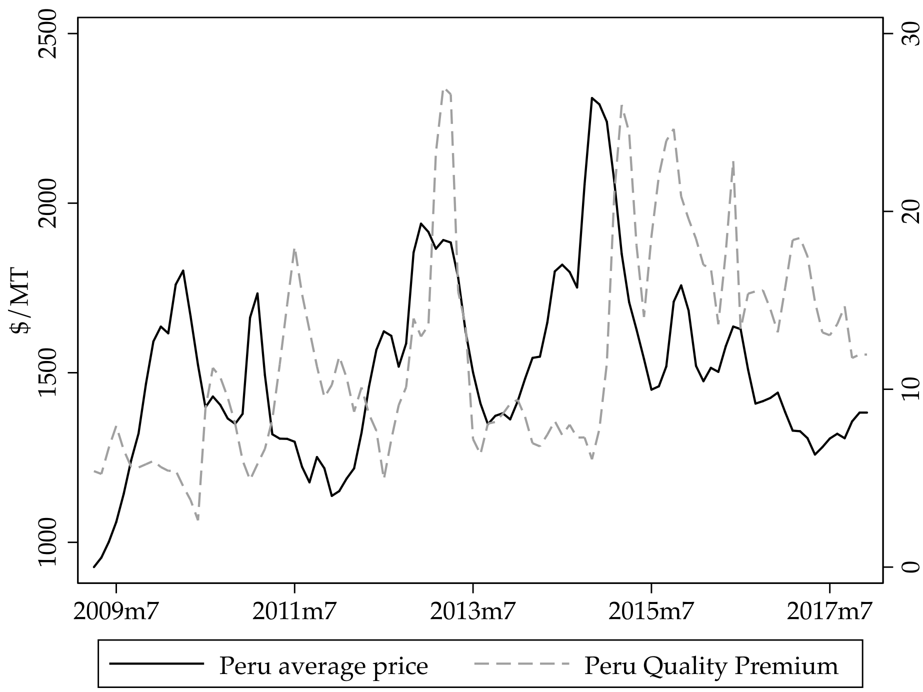
(d) Distribution of interactions across plants



Notes: Figures above show the average fraction of deliveries at each firm and plant for independent and integrated suppliers. In figures (a) and (b), plants or firms are ordered based on the frequency of deliveries for each boat x boat owner pair: the plant or firm that receives the highest number of deliveries by the boat in question while owned by the owner in question is ranked one, the next highest ranked two, and so on. In figures (c) and (d) plants or firms are ranked based on the frequency of deliveries for each boat *while it is integrated*. Figures (a) and (b) include all suppliers, while figures (c) and (d) include only *switchers*: boats that were independent at one point and integrated at another point during our sample.

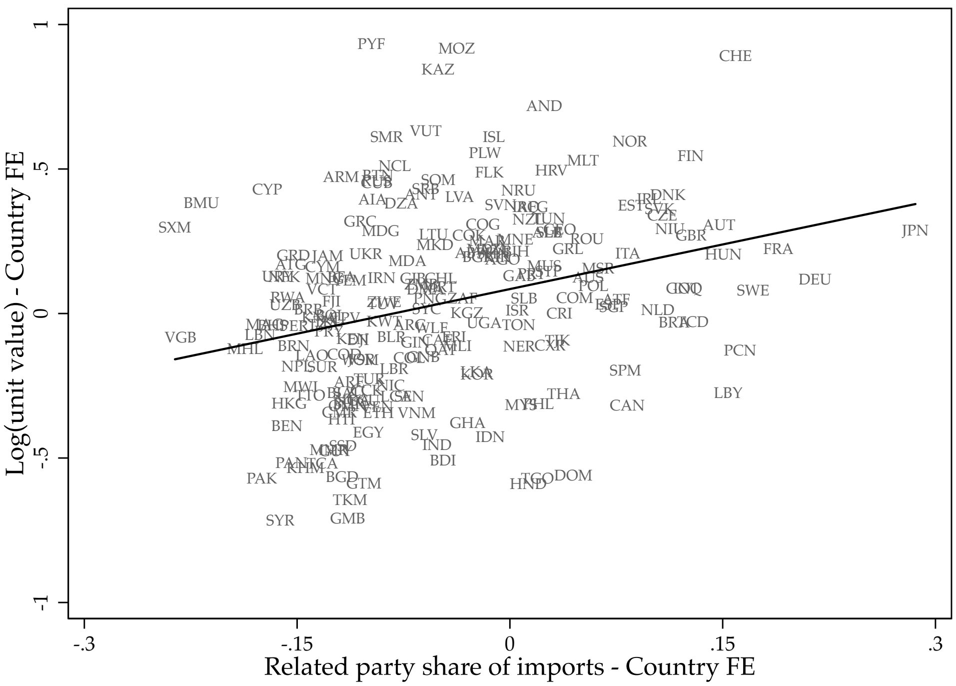
Figure 2.2: Average Fishmeal Price and Quality Premium in Peru

(a)



Notes: This figure shows the evolution over time of the average fishmeal price in Peru (the average between the price of "Prime" and "FAQ" fishmeal grades) and the Quality Premium in Peru. *Quality premium* is equal to $Log(\text{High Quality}) - Log(\text{Low Quality})$ where High and Low Quality are the average prices of "Prime" and "FAQ" fishmeal grades respectively.

Figure 2.3: Countries' output quality and vertical integration in export manufacturing



Notes: In this Figure, the variable plotted on the y-axis is $\hat{\gamma}_c$ from the regression $\log(\text{unit value})_{cpt} = \alpha_{pt} + \gamma_c + \varepsilon_{cpt}$, where $\log(\text{unit value})_{cpt}$ is the average log unit value of products exported from country c, of HS6 code p, in year t to the U.S.; α_{pt} is a product×year fixed effect; and γ_c is an origin country fixed effect. This regression is estimated using COMTRADE data from BACI (See Gaulier and Zignago (2010) for a description of the data). The variable plotted on the x-axis is $\hat{\delta}_c$ from the regression Related party share of U.S. $\text{imports}_{cpt} = \beta_{pt} + \delta_c + v_{cpt}$, where Related party share of U.S. imports_{cpt} is the share of products exported from country c, of NAICS code p, in year t to the U.S. that are imported by related parties (usually other units of the same firm (Ruhl, 2015)); β_{pt} is a product×year fixed effect; and δ_c is an origin country fixed effect. Related party share of U.S. imports_{cpt} is constructed using data from the U.S. Census Bureau. The data is from 2005 to 2014.

Chapter 3

Import Liberalization and Political Connections: Evidence from Myanmar

Joint with Felix Forster and Rocco Macchiavello

3.1 Introduction

Trade liberalization has been widely recognized as a key driver of productivity improvements and economic growth [1]. Yet, as institutional environments shape the gains from trade, regulatory or economic distortions can protect inefficient incumbent enterprises and so prevent reallocation of economic activity from low to high productivity firms.

In this chapter, we study how politically connected firms benefit from a competitive landscape tilted in their favour. This generally includes subverting the liberalization process itself. In developing country contexts, it furthermore includes market failures, such as incomplete credit markets, that connected firms are better positioned to overcome. The prevailing institutions and domestic market conditions may then protect rents and undermine the gains from trade liberalization.

Several features of the political environment in Myanmar and the unique data we gathered allows us to investigate these issues. First, after decades of military control, Myanmar recently went over a long process of economic and political liberalization. This provides a unique opportunity to follow how individuals connected to the previous regime were impacted by this transition. This provides a unique opportunity to follow how individuals connected to the previous regime were impacted by this transition. Second, the data environment allows us to study firms' connections with military power beyond just state-owned entreprises. The Directorate of Investment and Company Administration (DICA) publicly releases the list of all firms registered in Myanmar, their board members and their national identification number. We use this information to match firms in the import data to the list of individuals sanctioned by the US, the EU and Australia, for their connection to the military. We thus have three separate groups of firms for this study: SOEs, privately connected firms and non connected firms. Third, in 2013, the country implemented a trade liberalization reform, which removed the need to get a license prior to importing a shipment of goods. However, this liberalization

[1] See e.g. Pavcnik (2002); Topalova and Khandelwal (2011); Khandelwal et al. (2013) all inspired by Melitz (2003)-type theoretical predictions

was only limited to a subset of the products imported. Fourth, the government body in charge of the *de jure* liberalization is historically less connected to the military than the one implementing the reform *de facto*. We observe that the set of goods liberalized *de jure* is wider than the list of products liberalized *de facto*, which allows us to study the determinants of *de facto* liberalization.

In the first part of this chapter, we quantify the impact of the import liberalization reform. We follow closely the methodology developed in Khandelwal et al. (2013) in separating the effect of removing import licenses along the intensive and the extensive margins. In the sectors where connected firms are only marginally present, we show that liberalization had a significant impact on import volumes, almost entirely driven by entry of non connected firms. However, in the sectors populated by connected firms, our results depart from the existing literature on the subject in that we document no effect of the liberalization and no subsequent reallocation.

The second part of this chapter is dedicated to explaining these patterns. We first abstract from the import liberalization event and show that higher presence of connected firms is associated with less entry of non connected firms across sectors. This correlation holds both in the short and in long run. The low entry rates observed in sectors populated by connected firms is thus not specific to the impact of import license liberalization. We then document that the set of product codes in which connected firms operate exhibit important economies of scale. Importing these goods in large quantities or through shipments of high value offers significant unit price discounts. The results suggest the following interpretation. Economies of scale provides opportunities for rent-seeking activities. Wholesalers can import large quantities of these products at low prices and resell them in the domestic market. Even if importers charge significant markups when selling downstream, manufacturing firms and small final good distributors would choose to purchase these products from connected firms rather than importing the goods directly at higher costs. In turn, credit constraints or poor access to capital act as a barrier to entry and prevent other large firms or wholesalers to compete with connected firms in these sectors.

Finally, we show that only a subset of the goods liberalized *de jure* were liberalized *de facto*. The ministry deciding on the set of products to be liberalized is historically less connected to the military than the ministry implementing the reform. We document that products connected firms were importing were less likely to be *de facto* liberalized. However, sectors with connected firms presence which were liberalized *de facto* were more likely to be products with low economies of scale. This last result suggests that the institution implementing the reform maintained the licensing scheme to protect connected firms in the sectors where they had to rely more on the licensing process to act as a barrier to entry.

Our paper relates to two broad strands of the literature: (i) studies from international economics and development on the impact of trade liberalization; and (ii) the economics of political connections.

The trade literature highlights reallocation of economic activity as a key characteristic of liberalization episodes. The seminal work by Melitz (2003) provides the general narrative that low-productivity incumbent firms exit due to increased competitive pressure from imports as existing high-productivity (exporting) firms and new entrants grow their share. These dynamics are broadly reflected in the empirical literature on developing countries (Pavcnik, 2002; Topalova and Khandelwal, 2011; Brandt et al., 2017, 2019) as well as advanced economies (Trefler, 2004; Bernard et al., 2006). There is also growing evidence, however, of differences from this standard narrative. Khandelwal et al. (2013), for instance, find that that the removal of inefficient export quota allocations in China resulted in higher-than-expected growth (through entrant rather than incumbents). In a related paper, Bai et al. (2018) show that other types of distortions can lead to the opposite: welfare losses from trade liberalization. Most closely related to our work is a paper by Baccini et al. (2019) who show that private firms in Vietnam experience the predicted aspects of trade liberalization during accession to the WTO (higher exit rates, lower profitability and increases in productivity) while state-owned enterprises (SOEs) did not. Overall, the "consensus" that emerges from this literature is that entry of new firms and reallocation to high-productivity firms are very important for the gains from trade liberalization to materialize, but that barriers to entry (regulatory or economic) can undermine this process. Our paper's contribution picks up exactly this point by expanding the scope of barriers beyond SOEs to private politically connected firms.

Our focus on imports relative to the trade literature is noteworthy for three reasons. First, imports have been recognized as a key contributor to the productivity gains from trade liberalization. They arise within firms that benefit from the above-mentioned reallocation at the sector level, but also within firms from access to cheaper and/or higher quality imported inputs (e.g., Kugler and Verhoogen, 2009; Goldberg et al., 2009, 2010; Topalova and Khandelwal, 2011; Halpern et al., 2015b). Second, in most of the recent theoretical trade literature and the description of many related empirical findings, imports are either assumed to be consumed (final goods) or used as production inputs (intermediates); but they are not traded domestically. One of the simpler ways in which connected firms, however, might extract rents is by restricting access to certain imports or positioning themselves in sectors with high barriers to entry and imposing high markups on domestic resale. We therefore regard "importing" as a separate step in the international supply-chain that can be conducted by wholesalers, in contrast to the direct sales by the producer as it is common in the literature. Third, our paper is the first, to our knowledge, to distinguish import sectors according to their economies of scale. This may be particularly relevant in developing country contexts, where capital constraints preclude many businesses from operating at a large scale and thereby limit competition and potentially the gains from trade within the "importing segment" of the supply chain.

Our research also relates to the empirical literature on the relationship between political connections and private sector outcomes in developing countries. The most prominent studies in this literature aim to quantify the value of political connections, either for firms (Fisman, 2001; Faccio, 2006; Chekir and Diwan, 2014; Rijkers et al., 2017) or for politicians (Fisman et al., 2014). Firm value in the former case is, however, typically measured as equity value and taken from public companies' financial data. This restricts the set of firms included in the analysis and requires that financial markets accurately incorporate the value of connections into company share prices. While the value of connections is not the primary focus of our study, we contribute to this literature by assessing the differential performance by political connectedness on a different (and potentially broader) set of firms and by introducing alternative performance measures, including entry/exit rates, trade value.

This literature also provides some insight into the relevance of the institutional context for the success of economic reform efforts and macroeconomic performance. Most closely related to our work is a paper by Rijkers et al. (2017) who study the relative performance of firms connected to President Ben Ali's family in Tunisia and the associated sector licensing requirements, FDI restrictions, and market structure[2]. A few case studies of *crony capitalism* in other countries with strong military regimes, including Egypt (Chekir and Diwan, 2014) and Turkey (Demir, 2005), also document how privatization of state-owned assets and trade liberalization can facilitate the establishment of opaque rent-seeking networks. Firms and sectors intertwined in these networks can be shielded from competition and thereby undermine key benefits from economic liberalization. The evidence provided in these case-studies, however, is generally qualitative or exclusively focused on describing the setting. Overall, international trade is a dimension that, to our knowledge, is almost entirely lacking from this literature.

3.2 Background and data

3.2.1 Historical Background

After over 50 years of authoritarian control, Myanmar military junta was officially dissolved in 2011 and a nominally civilian government was installed. Alongside the release of political prisoners, improved foreign relations led to the easing of western trade and

[2] Also notable are several studies on the Chinese economy that attempt to explain the country's growth experience despite the prominence of SOEs, political connections, and rent-seeking relationships, e.g., through reduced frictions from market imperfections (Kang, 2003) or competition between locally favored businesses (Bai et al., 2014)

targeted economic sanctions. In the following years, the military-backed government implemented a series of economic and political reforms. The 2013 announcement that Aung San Suu Kyi's National League for Democracy (NLD) party could run for elections and their widely expected landslide victory held the promise of change for many, including a new business environment. While the economic liberalization has continued, albeit at a slower pacethere is widespread speculation/anecdotal evidence among observers that the patronage networks established during the military rule, still loom large in the business and political spheres (James, 2010; Jones, 2014; Larkin, 2015).

Despite the expectation of changing crony relationships during Myanmar's transitions, their persistence is not entirely surprising given the country's history of power relations. First, one key justification for military control in Myanmar since it first took power in 1962 has been the presence of insurgent armed groups, particularly in the border regions with Thailand and China, several of which still remain in conflict with the central state today. The military negotiated bilateral "ceasefire agreements" with many of these groups in the early 1990s (and later from 2011). In exchange for relinquishing varying degrees of territorial control, armed groups were permitted to continue their activities, retain arms and access government services. The relative stability brought about by ceasefires also allowed such groups to commence or dramatically increase activities in lucrative illicit trade, including drug trafficking (Callahan, 2007). It is alleged that several of today's largest enterprises in Myanmar build their capital base on profits derived from illicit trade, and such companies were long "permitted" to launder the proceeds through state banks (Meehan, 2011). In addition, the ceasefire agreements, in conjunction with nascent pro-market reforms, opened up the possibility for large-scale resource extraction businesses - prominently mining, logging and rubber - in the resource-rich border regions. Monopoly licenses for these extractive industries were granted to a select group of individuals with close personal connections to the military regime, ensuring a tight network of beneficiaries comprised of ethnic leaders, military officers and crony entrepreneurs Woods (2011). Overall, the backdrop of Myanmar's state fragility and the military's engagement with ethnic leaders created a system of entrenched interests among the elites that has in part survived until today.

A second contributing factor to the persistence of cronyism in Myanmar is the continued direct involvement of the military in many parts of state and business affairs, despite its partial retreat from government. Importantly, as per the 2008 constitution, a quarter of the seats in both houses of parliament and states and regions parliaments are reserved for military personnel appointed by the army, guaranteeing the former regime's continued influence on the legislative process in the country and veto power over constitutional amendments. They also fully control three powerful ministries - Home Affairs, Defense and Border Affairs. In addition, two military-owned conglomerates, the Union of Myanmar Economic Holdings Limited (UMEHL) and the Myanmar Economic Corporation (MEC), still belong to the largest companies in the country despite partial divestitures. UMEHL and MEC were established in 1990 to directly finance the army's operations and personnel, including retired veterans, and the companies own interests in a broad range of sectors in the country, including outright ownership and joint ventures Myoe (2009). During the 1990s, all major foreign investments were required to enter a joint venture with such military firms These large conglomerates continue to extend the military's reach deep into the business community of Myanmar and foster close connections not only to subsidiaries but also affiliate companies.

Finally, several waves of privatization in Myanmar, most importantly those staring in 2008 and in 2011, concentrated private asset ownership in the hands of the business elite with personal connections to the military regime. These included primarily family members and close prior business associates, like the "national entrepreneurs" that were already the key beneficiaries of the first private businesses and "ceasefire capitalism" in the 1990s. During the reign of successive military regimes, the state was long reliant on the support of the private sector, with enterprises often supplementing state capacity through the provision of public goods in exchange for import permits or monopolistic concessions. For example, connected firms contributed to the construction of the sprawling new capital city of Naypyitaw in the early 2000s. This co-dependence became dramatically apparent in the aftermath of the 2008 cyclone Nargis when relief efforts were partially provided by individual publicly denounced as “cronies” in coordination with international donors and non-governmental organizations Jones (2014). More recently, the NLD government has asked private individuals including the wealthiest who gained their fortunes under the former military regime to contribute to the humanitarian and rebuilding efforts in the wake of Rakhine State’s Rohingya crisis. Despite the shifting political landscape, the concentration of resources in a few industries therefore appears to have limited the government’s alternatives in some areas to working with individuals publicly regarded as cronies of the former regime.

3.2.2 Measures of connectedness

Our measures of firm connectedness cover two broad categories: (i) firms with direct connections to the former military regime via company board membership or commonly known business connections, which we refer to as “Privately connected firms” and (ii) state-owned enterprises, which we label “SOEs” for the remainder of this chapter. Former military regime members, their family, and their business associates have access the strong (patronage) network of the powerful elite that effectively governed the country for half a century. State-owned enterprises undoubtedly also have access to these networks, but may face political constraints to use them to their advantage under the new democratic regime. Nonetheless, SOEs may benefit more directly from current government policy and access to similarly valuable resources, including financing. The key common features for our analysis are that both types of connections can facilitate access to capital and help overcome regulatory barriers (such as licensing requirements).

Private firms’ connectedness to the former military regime is measured according to information from international sanctions lists and investigative work based on public sources by a local research company. In the aftermath of the 1990 elections in Myanmar and the military’s refusal to relinquish power, trading partners around the world gradually imposed sanctions targeted at senior figures in the military regime, their family members and close business associates. We focus on the targeted sanctions by Australia,[3] the EU,[4] and the US[5] that were in effect at the outset of the recent democratic transition, the November 2010 election, and through their revocation in 2012, 2013 and 2016 for Australia, the EU, and the US, respectively.[6] While these sanctions typically prohibited trade with named entities and froze any assets held in the sanctioning jurisdiction[7], we do not think of them as a variable prohibiting these firms to operate in the import market. Ninety percent of the goods imported in Myanmar come from neighboring and non sanctioning Asian countries. Instead, we consider these sanctions lists as comprehensive records of the most important actors with connections to the former military regime, compiled by some of the most intelligence agencies.[8] In aggregate, these sanctions lists provide a total of 669 individuals and 136 firms with connections to the former military regime.

Sanctioned businesses from these lists were matched to those from the customs data directly by name whereas sanctioned individuals were first matched to firm directors using a nearly comprehensive company registry maintained by DICA. For the identification of the sanctioned individuals, the sanctions lists often contain detailed informa-

[3]These were initially under the *Banking (Foreign Exchange) Regulations 1959* by the Reserve Bank of Australia and since 2011 under the *Autonomous Sanctions Regulations 2011* by the Minister for Foreign Affairs.

[4]These were implemented via various regulations of the Council and Commission of the European Union .

[5]These were implemented under the *Burma sanctions program* by the Office of Foreign Assets Control (OFAC).

[6]Some targeted sanctions were re-introduced in response to the Rohingya crisis in Myanmar's Rakhine state, but these were not considered in our analysis.

[7]The Australian sanctions were first introduced in October 2007 and we refer to the October 2008 amendment (which includes an extended list) and subsequent amendments to construct a comprehensive list of entities in Myanmar sanctioned by Australia for their connection to the former military regime. Targeted sanctions by the EU against entities in Myanmar were first introduced in April 2003 and, similar to the Australian case, we consider the most substantially expanded list from May 2010 and subsequent amendments for our analysis. Targeted US sanctions against individuals and firms connected to the former military regime were initiated in 2007 and we capture all entities from the Specially Designated Nationals (SDN) list from that point until their removal in October 2016.

[8]The sanctions lists notably change over time and we consider any individual or firms as being connected to the military regime if she/he is listed at least once during the periods given above.

tion, including various aliases, passport numbers, addresses, affiliated businesses, and family members. This information was used in corroborating any potential matches between the entities from the sanctions lists and DICA company directors. In a second step, the businesses from DICA were then matched to the import data by name. The matching of sanctioned individuals to company directors in DICA was implemented by the mentioned local research company and the matching of company directors to companies in the customs data conducted or confirmed by a team of research assistants in Myanmar. In addition, this research company also maintains its own database of companies and individuals connected to the former military regime and we also classify these companies as connected whenever they were not already among the sanctioned entities (or entities with sanctioned directors).

The second category of connected firms, "SOEs", is identified directly in the import customs data – which is available at the transaction level from April 2011 to March 2016 [9]– by means of a variable distinguishing between private and government enterprises.[10]

This process resulted in a total of 300 importers that are connected to the former military regime and 420 SOEs. Tables 3.1 provides summary statistics on their relevance among importers, in terms of number and size. As a share of the total number of importers in the country, connected firms and SOEs are a very small fraction (less than 1% taken together), but they are substantially larger on average than not connected firms and include some of the largest importers by value, accounting for nearly 30% of total imports (11.6% and 17.3% for connected and government entities, respectively). Figure 3.1 documents that there is significant dispersion in the presence of connected firms across sectors. They are not present in approximately 40 percent of product codes and appear marginally (market share lower than 20 percent) in 45 percent of sectors. Connected firms have more than 80 percent of the market share in only a handful of

[9]The data was provided by the Ministry of Commerce. It has a total of approximately 4 million transactions.

[10]As a few companies are both sanctioned and government, for this analysis, we define any importer as an SOE if the entity is listed as a government company in the customs data and it is not classified as firm with connections to the former military regime according to the process described above.

product codes. Appendix Table C.1 reports the share of import value for HS-chapters by the different importer classifications and overall. It is evident that the activity of connected firms is primarily focused on a few sectors: mineral products, especially mineral fuels/oils (approximately 40%), machinery and electrical equipment (13%), metals and metal products (9%), transportation equipment (8%), and animal/vegetable fats and oils (8%). Overall, over 90% of imports by value among connected entities is accounted for 17 HS2-codes.

3.3 Import license liberalization: implementation and impact

3.3.1 *De facto* vs. *de Jure* liberalization

In February 2013, the Government of Myanmar removed the requirement to obtain a license prior to importing 166 broad categories of goods[11]. Prior to this reform, all companies had to apply for a license from the Ministry of Commerce before being able to clear a shipment from customs. As such, the removal of licenses should have reduced the fixed cost of making individual import shipments. The reform was implemented by the customs department in April 2013, the beginning of the new fiscal year in Myanmar. Historically, the Ministry of Commerce in Myanmar is known to be reformist and pro-liberalization while the customs department, under the control of the Ministry of Finance and Revenue, is more tied to the former military regime.

We compare the *de jure* list of products liberalized which was decided upon and published by the Ministry of Commerce, [12] to the *de facto* set of product codes liberalized from the customs data. Customs indicates in the transaction-level data if a license was requested for clearing a specific shipment. The *de jure* list of liberalized products is wider

[11]This corresponds to roughly 2,700 HS6 product codes.

[12]The list published by the Ministry of Commerce is actually a negative list which provides the product codes that were still under the licensing scheme post-reform. (Announcement Order no 16/2013)

than the *de facto* list. Of the products which were meant to be liberalized by the Ministry of Commerce and appear at least once in the trade data over our sample period, 58 percent ended up not being liberalized in practice. We do not find evidence in the data that this partial liberalization was implemented differentially across firms. The *de facto* set of products which were liberalized was the same for all companies, connected or not. We asked the Ministry of Commerce in Myanmar who was aware of the situation and attributed it to poor communication or lack of training of customs agents. Post-liberalization, the Ministry of Commerce was still issuing licenses for products for the products liberalized *de jure* to satisfy the requirements made by customs agents.

As can be seen in Figure 3.2, the choice of import goods to be liberalized *de jure* was mostly driven by a choice of broad product categories or HS chapters. Almost none of animal, vegetable and mineral products were liberalized, while for most of the other HS chapters, more than 70 percent of product codes were liberalized. The patterns in the share of products that were *de facto* liberalized is more challenging to interpret. We revisit the determinants that render a product code *de jure* but not *de facto* liberalized in section 3.4.3.

3.3.2 The effect of import liberalization

In this section, we closely follow the methodology employed by Khandelwal et al. (2013) to measure the impact of the removal of import licenses on trade volumes and the contribution of incumbents versus entrants in generating these patterns. We restrict our analysis to the set of products which are imported pre and post reform (excluding product entry and exit) and the set of products labeled as "not liberalized" and "liberalized" by both the Ministry of Commerce and the customs data. We do so as we know little about the conditions under which licenses were still delivered by the Ministry of Commerce for the *de jure* but not *de facto* set of products liberalized, so that we cannot easily compare the cost of obtaining a license between these products and those that were not liberalized *de jure*.

We estimate a simple difference-in-differences model that compares changes in outcomes among liberalized and non liberalized products in 2013, the first year post-reform:

$$G_{pt} = \gamma_p + \alpha_t + \beta \cdot 1\{\text{Lib}\} \times 1\{\text{Post}\} + \epsilon_{pt} \tag{3.1}$$

where $G_{pt} = \dfrac{Y_{p,t} - Y_{p,t-1}}{Y_{p,t} + Y_{p,t-1}}$ is the centered growth rate between year t-1 and year t. 1{Lib} is a dummy equal to one if the HS6 code belongs the list of liberalized products, and 1{Post} is a dummy that indicates if the fiscal year belongs to the post liberalization period. The terms α_t are time fixed effects and γ_p are product fixed effects.

The regression includes both fiscal years 2012 and 2013. As we do not have sufficient time periods pre-liberalization to test the parallel trends assumption, we add HS6 fixed effects to account for the heterogeneity in growth rates across product codes. Our preferred setting replaces time fixed effects with HS2-time fixed effects to control for demand shocks. As illustrated in Figure 3.2, the choice of products liberalized *de jure* was based on broad product categories, which could be subject to differential shocks in demand. We thus consider the variation within years and broad product categories to compare the trade volumes between liberalized and non liberalized products.

We split sectors according to the prevalence of connected firms. We do so to differentiate the impact of the liberalization between sectors with high and low incidences of connected firms. We define a sector to be *connected* if privately connected firms and SOEs represent at least 15 percent of the market share in the pre-liberalization period (66th percentile) [13]. We then decompose the effect of the liberalization into the intensive and extensive margins. The intensive margin corresponds to incumbents, firms that import the same product code in fiscal years 2012 and 2013. The extensive margin comprises entrants and exiters. Entrants are firm-product pairs which import in fiscal year 2013 but not in 2012, and the converse holds for exiters. For each margin, we distinguish

[13] In Appendix Table C.2, we show a robustness test where sectors are defined as *connected* if at least 30 percent of the market share in the pre-liberalization period is comprised of privately connected firms and SOEs. The patterns found are similar under that definition.

connected and non-connected firms. We thus estimate 8 regressions (two connected firm categories multiplied by two margins for *connected* and *non connected* sectors) similar to the estimation equation given in 3.1.

We choose to work with the centered growth rate for two reasons. First, as opposed to working with market shares as in Khandelwal et al. (2013), it facilitates the interpretation of the decomposition exercise as the sum of each margin's contribution corresponds to the overall impact of the liberalization. Second, we take the centered growth rate rather than the classical growth rate to have a bounded variable centered on zero. There is enormous heterogeneity in growth rates across product codes but the classical growth rate is not bounded above which could lead to overestimating the impact of the liberalization.

Finally, for the decomposition to reflect the effect per dollar imported rather than the average across sectors, we weight each observation by its product code share in total trade in the pre-liberalization period. This is necessary for the sum of the margins' contributions to match the overall estimate of the liberalization effect - reallocation patterns in sectors worth tens of millions of dollars must proportionately more in the regression than sectors worth several thousand dollars.

Table 3.2 presents the results of the overall impact of the liberalization. Panel A shows that import license liberalization had a important effect on import volumes, but only in non connected sectors. Column 1 shows a negative and significant coefficient for connected sectors which is smaller in magnitude and no longer significant when adding HS2-year pair fixed effects. Panel B shows that there is no differential effect in connected sectors whether the majority of the market share of connected firms is composed of privately connected firms or SOEs.

Table 3.3 presents the decomposition across the different margins. Given the large number of regressions ran for this table, we only report the coefficients and their significance. In non connected sectors, the positive effect of the liberalization is almost entirely driven by the net entry of non connected firms. The results are similar to Khandelwal et al. (2013), in that incumbents do not seem to benefit from the liberalization[14] while entrant firms, who may have been unable to import due to the fixed cost of getting a license prior to 2012, benefit greatly from the liberalization. In connected sectors however, there is a small reallocation effect between connected and non connected firms, where non connected firms appear to grow marginally while connected firms shrink but not significantly. Importantly, while it is almost entirely net entry that contributes to the positive impact of the liberalization in non connected sectors, there was no significant entry in the liberalized connected sectors relative to non liberalized connected sectors.

Analyzing these results, we postulate that there are higher barriers to entry in the sectors populated by connected firms. This hypothesis would explain both why there was limited entry in the liberalized connected product codes post-reform but also why connected firms operate in these sectors in the first place. Through their connections, these firms may be able to overcome the high barriers to entry in these sectors more easily and thus capture rents from limited competition in importing these goods. In the next section, we explore this hypothesis.

3.4 Connected firms and economies of scale

3.4.1 Patterns of entry and presence of connected firms

We first abstract from the import license liberalization to explore the patterns of entry of non connected firms across sectors as a function of the presence of connected firms, in the short and in the long run. We do so by running the following cross-sectional regression at the product-code level:

$$MS_{p,2012}^{\text{non conn, ent}} = \alpha + \beta \cdot MS_{p,2012,inc}^{\text{conn}} + \epsilon_p \tag{3.2}$$

[14]In Appendix Table C.3, we show that even if incumbents did not grow significantly post liberalization, they started making smaller shipments but more of them on a given year, consistent with the remark made earlier that removing the licensing process reduced the fixed cost of making an individual shipment.

where p is a product-code, $MS_{p,2012}^{\text{non conn, ent}} = \frac{value_{p,2012}^{\text{non conn, ent}}}{value_{p,2012}}$ is the market share of non connected firms that are entrants in 2012 (that did not import the good in 2011 but import it in 2012) and $MS_{p,2012,inc}^{\text{conn}} = \frac{value_{p,2012}^{\text{conn, inc}}}{value_{p,2012}^{\text{inc}}}$ is the market share of connected firms *among* incumbent firms in 2012 (firms that imported the product code in 2011 and 2012). We use the market share among incumbents rather than simply the overall market share of connected incumbents to avoid a mechanical correlation between the market share of incumbents and the market share of entrants. The parameter β thus captures potential barriers to entry associated with sectors in which connected firms are also more likely to be present.

We run another similar regression replacing the dependent variable by $MS_{p,2015}^{\text{non conn, ent}}$, the market share of non connected entrant firms in 2015 to capture how the presence of connected firms in sectors affects entry of non connected firms in the long run. The estimates from these two regressions are reported in column 1 of Table 3.4. Panel A reports the short run correlation while Panel B reports the same coefficient in the long run. Both panels show that a higher incidence of connected firms negatively impacts entry of non connected firms. These results are consistent with our hypothesis made at the end of the previous section that connected firms operate in sectors with higher barriers to entry. Low entry rates of firms in sectors in which connected firms are present is thus not specific to the license liberalization.

From these basic regression settings, we gradually add sector characteristics as controls and explore how adding them affects our estimate of the β parameter. We look for intrinsic sectoral features that would generate high barriers to entry and would then make it more likely for connected firms to operate in these sectors. The negative correlation observed in column 1 is robust to adding HS2 fixed effects, which could potentially capture the differences in fixed costs associated with entry in various broad product categories. Column 3 adds BEC product types which classifies HS6 codes according to intermediates, capital and final goods. In column 4, we add fixed effects associated to the conservative Rauch (1999) classification, which codes products in three separate categories: goods traded on an organized exchange market, goods with a reference price and differentiated products. These last two product classification could capture differences in fixed costs associated with knowing demand for specific products downstream to the importing activity or setting up the distribution networks domestically to resell imported goods. Our estimates are robust to including these fixed effects [15].

Column 5 adds in a control which corresponds to the logarithm of the average shipment value of non connected firms in the considered sector. Our rationale for doing so is that some sectors may require that shipments be of a higher value because of minimum scale required for logistics reasons (it may be difficult to import small quantities of a particular good) or because importing large quantities of the good may give a competitive advantage (e.g. by obtaining lower unit prices). In both the short-run and the long-run, controlling for shipment value annuls the negative correlation observed previously. This means that connected firms are more prominent in sectors with higher capital requirements and suggests that they have easier access to the capital needed to make imports in these sectors. As capital may be more difficult to access for non connected firms, these sectors then tend to exhibit less entry. In the next sub-section, we test this idea that connected firms may be importing products with high economies of scale. If importing high quantities of a good translates into lower average unit costs, connected firms may leverage their access to capital to import goods at low prices and extract significant rents from doing so.

3.4.2 Measuring economies of scales

We test the hypothesis that connected firms occupy sectors with higher internal economies of scale. To measure economies of scale, we estimate the price discount for buying larger quantities or importing shipments of higher value. For each HS6 code, we first divide

[15]In an alternative specification, not presented here, we add controls for the main origins of the goods imported and do not find that it affects the correlation between entry and presence of connected firms

the sample of quantities imported into five bins of equal sizes[16] and run the following regression:

$$\log(\text{unit val})_{p,o,t} = \alpha_{p,o,t} + \sum_{i=1}^{5} \delta_i \cdot \mathbb{1}\{q_{p,o,t} \in Q_i\} + \epsilon_{p,o,t} \tag{3.3}$$

where $\log(\text{unit val})_{p,o,t}$ denotes the log unit value of a shipment of product p imported during month t from origin o, $q_{p,o,t}$ is the corresponding quantity, and $\alpha_{p,o,t}$ are product-origin-month fixed effects. The Q_i's are the quantity bins defined above and the δ_i correspond to the average unit value residual for quantity bin Q_i. The importance of economies of scale is assessed by the degree to which the estimated δ_i decrease with i.

Sectors with limited economies of scale might potentially never make very large shipment in value over our sample period, so that quantity bins in different sectors would not correspond to shipments of similar (comparable) values. We therefore also estimate a similar regression as the one above with value bins rather than quantity bins, the bins being defined in absolute value rather than relative value:

$$\log(\text{unit val})_{p,o,t} = \alpha_{p,o,t} + \sum_{i=1}^{5} \gamma_i \cdot \mathbb{1}\{v_{p,o,t} \in V_i\} + \epsilon_{p,o,t} \tag{3.4}$$

where the bounds of the value bins are: \$1,000; \$10,000; \$100,000 and \$1,000,000.

We estimate these two equations for *connected* and *non connected* sectors, as defined in the previous section, to assess if economies of scale are more important in the sectors populated by connected firms. We exclude connected firms from these regressions, because they might potentially benefit from lower prices through their connections that are unrelated to sector-specific scale economies.

As the unit value of a shipment is just its value divided by its quantity, there is a mechanical negative correlation between our dependent variable and the quantity bins and a positive one with the same left-hand side variable and the value bins. As such, our δ_i estimates are a lower bound of the actual economies of scale in the considered

[16] The bins are not of equal density as product codes with higher economies of scale should have the distribution of shipment sizes concentrated on high values

sectors while the γ_i constitute an upper bound. Importantly, we are not interested in the level of these estimated economies of scale per-se, but rather the relative level between connected and non connected sectors. Thus, we need to assume that the bias to which our δ_i and γ_i estimates are subject to is the same in sectors in which connected firms are prominent and sectors where they are not.

The coefficients δ_i and γ_i are reported in the top and bottom part of Figure 3.3. Both graphs display larger economies of scale for connected sectors compared to non connected sectors, and especially so for the highest quantity and value bins. These results confirm the hypothesis that connected firms are more prominent in sectors with high economies of scale. From the combination of all these empirical patterns, we posit that these firms have better access to the capital required to make large shipments and as such, capture the benefits from importing goods with high price discounts. These high economies of scale generate opportunities for rents to be captured by importing these goods in bulk at low prices reselling to the domestic market. Firms looking to compete with connected firms as wholesalers may not be able to do so from binding credit constraints while individual firms or small distributors face lower costs from buying their inputs or final goods from connected firms than directly importing the goods in small quantities.

3.4.3 Connected firms, economies of scale and *de facto* liberalization

In this final section, we investigate the determinants that made some sectors liberalized *de jure* but not *de facto*. We test the assumption that connected firms may have been able to prevent *de facto* liberalization in sectors where they needed protection from entrants. Our previous analysis has shown that connected firms tend to operate in sectors in which imports are subject to high economies of scale. Firms with limited access to capital may not be able to enter and benefit from these economies of scale. Thus, the liberalization of sectors with high economies of scale should not lead to entry, as observed in section 3.3.2, and would not jeopardize the rents captured by connected firms in these sectors. However, in sectors which do not exhibit high economies of scale, connected firms may rely more heavily on the licensing process to act as a barrier to entry and have incentives to prevent their liberalization if they are able to do so.

We test these hypotheses by running a set of probit regressions. We keep only the set of products which were liberalized *de jure* to estimate the probability that a given product code in that list is liberalized *de facto*. We run several regressions based on the following estimation equation:

$$\begin{aligned} y_p^* = \alpha_h + \beta \cdot \mathbb{1}\{\text{p=connected}\} + \gamma \cdot \mathbb{1}\{\text{p=economies of scale}\} \\ + \delta \cdot \mathbb{1}\{\text{p=connected}\} \cdot \mathbb{1}\{\text{p=economies of scale}\} + \epsilon_p \end{aligned} \tag{3.5}$$

where p indexes a product code, α_h are HS2 fixed effects, $\mathbb{1}\{\text{p=connected}\}$ is a dummy equal to one if product p is a *connected* sector as defined in section 3.3.2 and $\mathbb{1}\{\text{p=economies of scale}\}$ indicates whether sector p is subject to economies of scale on the import side. y_p^* is the standard probit model latent variable that drives the likelihood that a product be liberalized *de facto*. Our hypothesis is that β should be negative and significant while δ should be positive and significant.

To measure economies of scale in a particular sector, we use the logarithm of the average shipment size for non connected firms as in section 3.4.1. Additionally, we run the regression described in Equation 3.4 for each product code and then categorize a specific product code as subject to economies of scale if the estimate γ_5 in that sector is negative [17].

Results are presented in Table 3.5. Column 1 presents the pseudo-R^2 in a version with only HS2 fixed effects as a benchmark. Column 2 shows that connected sectors where less likely to be liberalized. Column 3 uses a continuous version of the right-hand side variable, where the dummy for connected sector is replaced by the market share of connected firms in the pre-liberalization period. Column 4 splits connected

[17] For sectors in the sample which do not contain at least one shipment which value is higher than a million dollars, we extrapolate the coefficient from the slope of the γ_i's estimated for smaller bins as a fuction of i.

firms into our two sub-categories: privately connected firms and SOEs. The presence of both decreases the likelihood that a sector be *de facto* liberalized. Column 5 follows closely equation 3.5. It shows that while connected sectors were less likely to be *de facto* liberalized, connected sectors which exhibit high economies of scale were more likely to be *de facto* liberalized relative to connected product codes with lower economies of scale. Column 6, using our measure of economies of scale derived from section 3.4.2, shows similar results [18]. In both columns 5 and 6, economies of scale, in the absence of connected firms, do not appear to be important in explaining the likelihood that a sector be liberalized. These results suggest that connected firms may have used their relationships with parties in power to limit *de facto* liberalization and especially so in sectors with low economies of scale that could act as "natural" barriers to entry and prevent competition.

3.5 Conclusion

In this chapter, we investigate the impact of import license liberalization in Myanmar, a complex political economy environment. By contrast to the literature on this issue, we find that liberalization in the sectors where firms that are politically connected operate did not lead to substantial entry and so lack an important reallocation effect. We explore potential explanations for these patterns and document that connected firms tend to trade goods which offer important price discounts if imported in large quantities. These economies of scale act as a "natural" barrier to entry to protect connected firms from potential competition. Finally, this chapter documents interactions between the strategic decisions made by connected firms in the choice of the sectors in which they operate and the role of institutions in shaping economic liberalization while protecting the rents

[18] One caveat of these results is that the presence of connected firms in a sector or the importance of economies of scale in a given sector only explain a small portion of the total variation in the likelihood that a sector be *de facto* liberalized, especially relative to other factors such as 2 digits HS codes. However, both our measures of connectedness and economies of scale are noisy, so doing a variance decomposition exercise with these variables is challenging.

of connected firms. These observed patterns motivate the developments of a unifying theoretical framework to rationalize them.

The findings in this study also point to potentially important follow-up research. Our paper is, to our knowledge, the first to investigate the role of economies of scale in the import sector. Possible leads of research in that area could be developments of a stronger methodological approach to measuring economies of scale as well as documenting their role in shaping market power and concentration in import markets worldwide.

Tables and Figures

Table 3.1: Connected and Not Connected Importers

	Not Connected	Privately connected	SOEs	Total
Number of firms	118,412	300	420	119,132
Share of firms (%)	99.4	0.3	0.4	100.0
Value (Million USD)	45,720	7,455	11,156	64,330
Share of import value (%)	71.1	11.6	17.3	100.0

Notes: This table presents summary statistics for our connectness variables. Privately connected firms are firms which are either directly sanctioned by the US, the EU or Australia or firms who have on their board an individual that is sanctioned. SOEs are government companies. Total value is the total import value in our sample.

Table 3.2: Import license liberalization and connected sectors

Panel A: Privately connected firms and SOEs combined				
Dep. var:	Standardized growth rate			
	Connected sectors		Non connected sectors	
	(1)	(2)	(3)	(4)
1{Lib} × 1{Post}	-0.148^{*}	-0.026	0.233^{***}	0.467^{***}
	(0.083)	(0.155)	(0.036)	(0.098)
Year FEs	Yes	No	Yes	No
HS2 X Year FEs	No	Yes	No	Yes
HS6 FEs	Yes	Yes	Yes	Yes
N	818	818	1544	1544

Panel B: Distinguishing sectors with majority of privately connected vs. SOEs						
Dep. var:	Standardized growth rate					
	Connected sectors Privately conn.		Connected sectors SOEs		Non connected sectors	
	(1)	(2)	(3)	(4)	(5)	(6)
1{Lib} × 1{Post}	-0.086	0.204	-0.308	-0.165	0.233^{***}	0.467^{***}
	(0.082)	(0.240)	(0.187)	(0.254)	(0.036)	(0.098)
Year FEs	Yes	No	Yes	No	Yes	No
HS2 X Year FEs	No	Yes	No	Yes	No	Yes
HS6 FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	464	464	354	354	1544	1544

Notes: One observation is a 6 digit product code during fiscal year. 1{Lib} is a dummy equal to one if the product is in the group of products which are *de facto* liberalized. The sample only includes the set od products which are *de facto* and *de jure* liberalized or not liberalized. 1{Post} is a dummy equal to one if the fiscal year considered is greater than 2013, the first fiscal year when the liberalization was implemented. HS2 X Year FEs are 2 digit product codes - fiscal years pairs of fixed effects. Sectors are defined to be connected if the market share of connected firms in that sector in the pre-reform period is higher than 15 percent. Appendix Table C.2 provides a robustness check. Connected sectors where the privately connected firms constitute most of the market share of connected firms are presented in Panel B, columns 1 and 2. Connected sectors where SOEs constitute most of the market share of connected firms are presented in Panel B, columns 3 and 4. Standard errors are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: Decomposition of the effect of import license liberalization

	Connected sectors	Non connected sectors
Incumbents	0.007	0.003
Connected firms	−0.034	−0.008
Not connected firms	0.041**	0.011
Net Entry	−0.034	0.464***
Connected firms	−0.089	0.024
Not connected firms	0.056	0.441***
All	−0.026	0.467***
Connected firms	−0.123	0.016
Not connected firms	0.097	0.451***

Notes: This table reports the coefficients from the regressions detailed in section 3.3.2 and equation 3.1. The coefficient reported is the coefficient for the interaction $1\{\text{Lib}\} \cdot 1\{\text{Post}\}$, so each coefficient is estimated in a separate regression. For example, the incumbents non connected coefficient in non connected sectors corresponds to the estimates from regression equation 3.1 for a sample that only includes incumbent firms which are non connected and operate in non connected sectors. Standard errors are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: Connected sectors and entry

Panel A: Entry in the short run					
Dep. var:	Market share of non connected entrants in 2012				
	(1)	(2)	(3)	(4)	(5)
Connected firms MS - 2012	-0.080^{***}	-0.076^{***}	-0.072^{***}	-0.072^{***}	0.013
	(0.025)	(0.026)	(0.027)	(0.027)	(0.028)
Log(Av. shipment size)					-0.046^{***}
					(0.006)
HS2 FEs	No	Yes	Yes	Yes	No
BEC-type FEs	No	No	Yes	Yes	No
Rauch Class FEs	No	No	No	Yes	No
Panel B: Entry in the long run					
Dep. var:	Market share of non connected entrants in 2015				
	(1)	(2)	(3)	(4)	(5)
Connected firms MS - 2012	-0.089^{***}	-0.072^{***}	-0.069^{***}	-0.068^{***}	-0.027
	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)
Log(Av. shipment size)					-0.030^{***}
					(0.005)
HS2 FEs	No	Yes	Yes	Yes	No
BEC-type FEs	No	No	Yes	Yes	No
Rauch Class FEs	No	No	No	Yes	No

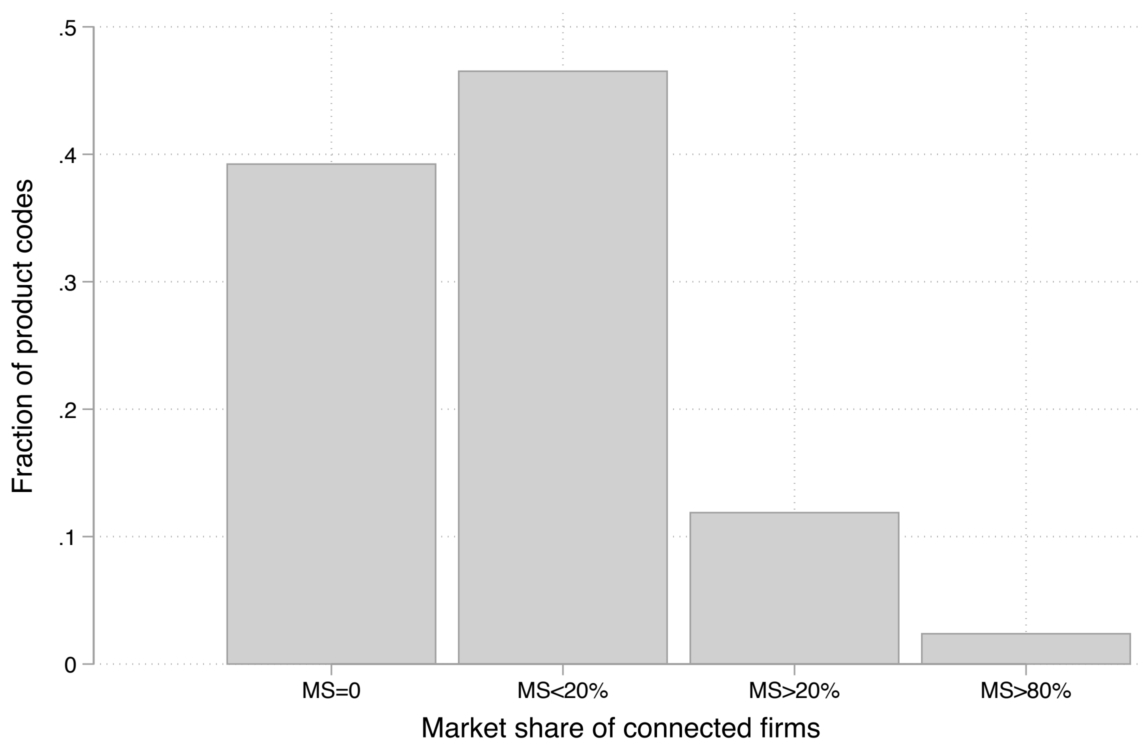
Notes: One observation is a 6 digit product code. In Panel A, the dependent variable is the market share of non connected firms which are entrants in 2012 (they did not import that product in 2011). In Panel A, the dependent variable is the same except that the year considered is now 2015, the last year in our sample. In both panels, the right hand side variable is the market share of connected firms among incumbent firms in 2012, that is the value of imports made by connected incumbent firms in 2012 divided by the value of imports made all incumbents in 2012 in the considered product code. Standard errors are included in parentheses. HS2 FEs are 2 digit product code fixed effects. BEC-type FEs are fixed effects that classify goods according to whether they are intermediates, final or consumption goods. Rauch Class FEs are fixed effects for the Rauch (1999) classification. Log(av. shipment value) is a leave out connected firms average shipment value by sector. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: *De facto* liberalization

Dep. var:	$\mathbb{1}$ {*de facto* liberalized}					
	(1)	(2)	(3)	(4)	(5)	(6)
Connected sector		-0.206^{***}			-1.481^{***}	-0.417^{***}
		(0.078)			(0.568)	(0.112)
Market share of connected firms			-0.591^{***}			
			(0.135)			
Market share of privately connected firms				-0.491^{**}		
				(0.208)		
Market share of SOEs				-0.652^{***}		
				(0.166)		
Log(av. shipment value)					-0.045	
					(0.037)	
Connected sector × Log(av. shipment value)					0.145^{**}	
					(0.064)	
Sector with economies of scale						-0.021
						(0.095)
Connected sector × Sector with economies of scale						0.377^{**}
						(0.158)
HS Chapter FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	1407	1407	1407	1407	1407	1407
Pseudo - R^2	0.144	0.147	0.154	0.154	0.150	0.152

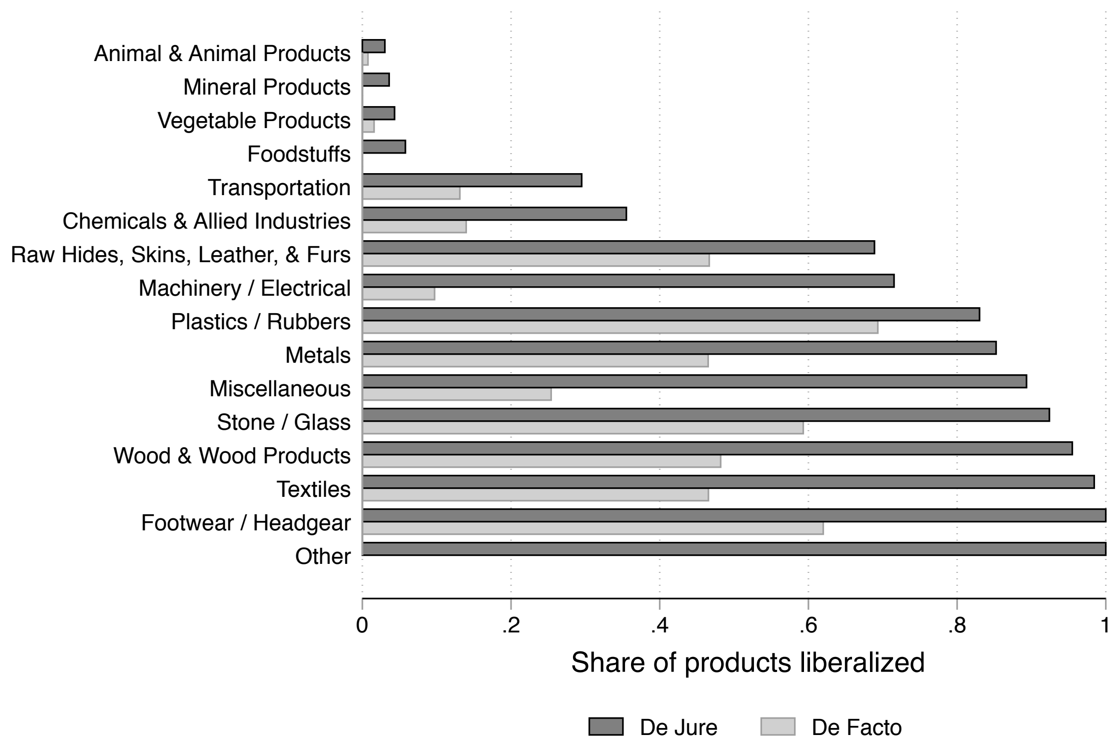
Notes: One observation is a 6 digit product code. The sample is restricted to the set of goods which are liberalized *de jure*. The table reports the estimation results from a probit model where the left hand side variable is a dummy equal to one if the product considered is liberalized *de facto*. A sector is defined to be connected is the market share of connected firms is greater than 15 percent. Log(av. shipment value) is a leave out connected firms average shipment value by sector. Sector with economies of scale is a dummy equal to one if the sector exhibits economies of scale as defined in section 3.4.3 and estimated through equation 3.4. HS2 FEs are 2 digit product code fixed effects. Standard errors are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 3.1: Distribution of the presence of connected firms across sectors



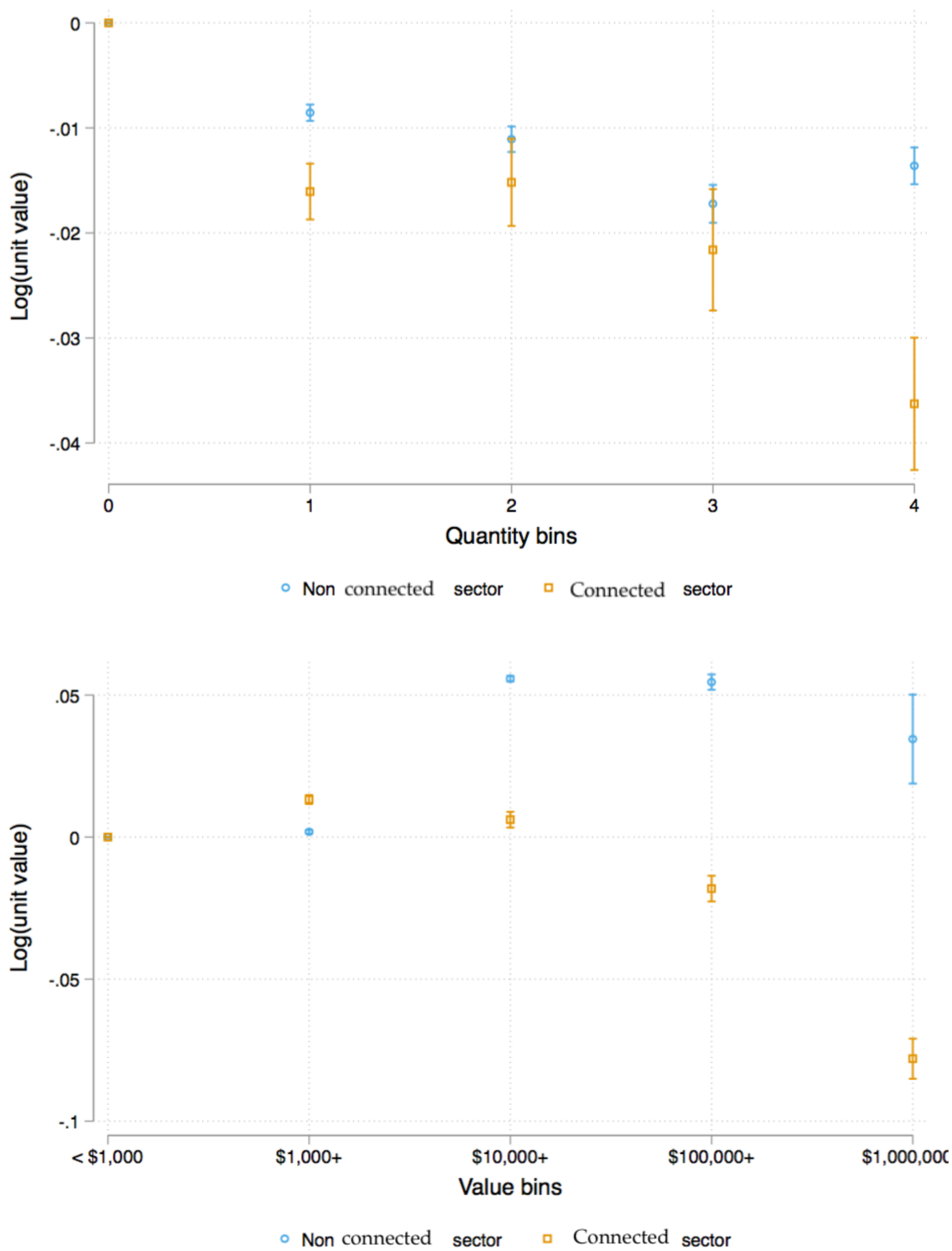
Notes: This figure shows the distribution of product codes in the pre-liberalization period according to the importance of the presence of connected firms. Approximately 40 percent of sectors do not have any connected firms trading the good, while more than 45 percent of the product codes have a connected firms market share higher than zero percent but less than 20 percent.

Figure 3.2: Share of products liberalized by HS chapter



Notes: This figure shows the percentage of products which are liberalized *de facto* or *de jure* and so by HS chapter. The share of products which are liberalized *de facto* is a subset of the set of products which are liberalized *de jure*. There is significant heterogeneity in the shares of products liberalized across HS chapters.

Figure 3.3: Economies of scale



Notes: This figure reports the point estimates of the quantity bins and value bins estimated in equations 3.3 and 3.4 respectively. The bars represent the 95 percent confidence intervals. Both equations are estimated separately for connected and non connected sectors.

Bibliography

Acemoglu, D., Aghion, P., Griffith, R., and Zilibotti, F. (2010). Vertical integration and technology: Theory and evidence. *Journal of European Economic Association*.

Acemoglu, D., Aghion, P., Lelarge, C., Van Reenen, J., and Zilibotti, F. (2007). Technology, information and the decentralization of the firm. *Quarterly Journal of Economics*.

Acemoglu, D., Johnson, S., and Mitton, T. (2009). Determinants of vertical integration: financial development and contracting costs. *Journal of Finance*.

Alfaro, L., Conconi, P., Fadinger, H., and Newman, A. F. (2016). Do prices determine vertical integration? *Review of Economic Studies*, 83:855–888.

Allen, T. (2014). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Amodio, F. and Martinez-Carrasco, M. A. (2018). Input allocation, workforce management and productivity spillovers: Evidence from personnel data. *Review of Economic Studies*.

Anderson, S. P. and Renault, R. (1999). Pricing, product diversity, and search costs: A bertrand-chamberlin-diamond model. *The RAND Journal of Economics*, pages 719–735.

Antràs, P. (2014). Grossman-hart (1986) goes global: Incomplete contracts, property rights, and the international organization of production. *Journal of Law, Economics, and Organization*.

Antràs, P. (2016). *Global Production: Firms, Contracts, and Trade Structure*. Princeton University Press.

Atalay, E., Hortacsu, A., and Syverson, C. (2014). Vertical Integration and Input Flows. *American Economic Review*, 104(4):1120–48.

Athey, S. and Haile, P. A. (2002). Identification of standard auction models. *Econometrica*, 70(6):2107–2140.

Atkin, D., Khandelwal, A., and Osman, A. (2017a). Exporting and firm performance: Evidence from a randomized trial. *Quarterly Journal of Economics*.

Atkin, D., Khandelwal, A. K., and Osman, A. (2017b). Exporting and Firm Performance: Evidence from a Randomized Experiment*. *The Quarterly Journal of Economics*, 132(2):551–615.

Baccini, L., Impullitti, G., and Malesky, E. J. (2019). Globalization and state capitalism: Assessing Vietnam's accession to the WTO. *Journal of International Economics*, 119:75–92.

Bai, C.-e., Hsieh, C.-t., and Song, Z. (2014). Crony Capitalism with Chinese Characteristics. *Working paper*, (May):1–12.

Bai, Y., Lu, D., and Jin, K. (2018). Misallocation under trade liberalization. *Working Paper*.

Baker, G., Gibbons, R., and Murphy, K. J. (2001). Bringing the market inside the firm? *American Economic Review, Papers and Proceedings*.

Baker, G., Gibbons, R., and Murphy, K. J. (2002). Relational contracts and the theory of the firm. *The Quarterly Journal of Economics*, 117(1):39–84.

Baker, G. and Hubbard, T. (2003). Make versus buy in trucking: Asset ownership, job design and information. *American Economic Review*.

Baker, G. and Hubbard, T. (2004). Contractibility and asset ownership: On-board computers and governance in us trucking. *Quarterly Journal of Economics*.

Baldwin, R. and Harrigan, J. (2011). Zeros, quality, and space: Trade theory and trade evidence. *American Economic Journal: Microeconomics*.

Banerjee, A. and Duflo, E. (2000). Reputation effects and the limits of contracting: A study of the indian software industry. *Quarterly Journal of Economics*.

Bastos, P., Silva, J., and Verhoogen, E. (2018). Export destinations and input prices. *American Economic Review*.

Benaroya, H., Han, S. M., and Nagurka, M. (2005). *Probability models in engineering and science*. CRC press.

Benguria, F. (2015). The matching and sorting of exporting and importing firms: theory and evidence. *Available at SSRN 2638925*.

Bernard, A. B., Jensen, J. B., and Schott, P. K. (2006). Trade costs, firms and productivity. *Journal of Monetary Economics*, 53(5):917–937.

Bernard, A. B., Moxnes, A., and Ulltveit-Moe, K. H. (2018). Two-sided heterogeneity and trade. *The Review of Economics and Statistics*, 100(3):424–439.

Blum, B. S., Claro, S., and Horstmann, I. (2010). Facts and figures on intermediated trade. *American Economic Review*, 100(2):419–23.

Blum, B. S., Claro, S., and Horstmann, I. (2011). Intermediation and the nature of trade costs: Theory and evidence. *University of Toronto, mimeograph*.

Boehm, J. (2016). The impact of contract enforcement costs on outsourcing and aggregate productivity. mimeo, Sciences Po.

Brambilla, I., Lederman, D., and Porto, G. (2012). Exports, export destinations, and skills. *American Economic Review*, 102(7):3406–38.

Brambilla, I., Lederman, D., and Porto, G. (forthcoming). Exporting firms and the demand for skilled tasks. *Canadian Journal of Economics*.

Brambilla, I. and Porto, G. (2016). High-income export destinations, quality and wages. *Journal of International Economics*.

Brandt, L., Van Biesebroeck, J., Wang, L., and Zhang, Y. (2017). WTO Accession and Performance of Chinese Manufacturing Firms. *American Economic Review*, 107(9):2784–2820.

Brandt, L., Van Biesebroeck, J., Wang, L., and Zhang, Y. (2019). WTO Accession and Performance of Chinese Manufacturing Firms: Corrigendum. *American Economic Review*, 109(4):1616–1621.

Bresnahan, T. F. and Levin, J. D. (2012). Vertical integration and Market Structure. In Gibbons, R. and Roberts, D., editors, *Handbook of Organizational Economics*.

Breza, E. and Liberman, A. (2017). Financial contracting and organizational form: Evidence from the regulation of trade credit. *Journal of Finance*.

Cajal-Grossi, J., Macchiavello, R., and Noguera, G. (2019). International buyers' sourcing and suppliers' markups in bangladeshi garments.

Callahan, M. P. (2007). *Political Authority in Burma's Ethnic Minority States: Devolution, Occupation and Coexistence*. East-West Center, Washington, DC.

Chekir, H. and Diwan, I. (2014). Crony Capitalism in Egypt. *Journal of Globalization and Development*, 5(2):1–34.

Chen, C.-C. and Tyler, C. W. (1999). Accurate approximation to the extreme order statistics of gaussian samples. *Communications in Statistics-Simulation and Computation*, 28(1):177–188.

Coase, R. (1937). The nature of the firm. *Economica*.

Cuesta, J. I. and Sepulveda, A. (2018). Price regulation in credit markets: A trade-off between consumer protection and credit access. *Available at SSRN 3282910*.

De La Puente, O., Sueiro, J. C., Heck, C., Soldi, G., and Puente, S. D. L. (2011). Evaluacion de los sistemas de gestion pesquera en el marco de la certificacion a cargo del marine. la pesqueria peruana de anchoveta. Stewardship Council, Cayetano Heredia University WP.

Demir, F. (2005). Militarization of the Market and Rent-Seeking Coalitions in Turkey. *Development and Change*, 36(4):667–690.

Eslava, M., Tybout, J., Jinkins, D., Krizan, C., Eaton, J., et al. (2015). A search and learning model of export dynamics. In *2015 Meeting Papers*, number 1535. Society for Economic Dynamics.

European Commission (2018). European commission: Fishing quotas. `https://ec.europa.eu/fisheries/cfp/fishing_rules/tacs_en`.

Faccio, M. (2006). Politically Connected Firms. *The American Economic Review*, 96(1):369–386.

FAO (1986). The production of fish meal and oil. FAO FISHERIES TECHNICAL PAPER: 142.

Ferraz, C., Finan, F., and Szerman, D. (2015). Procuring firm growth: the effects of government purchases on firm dynamics. Technical report, National Bureau of Economic Research.

Fisman, R. (2001). Estimating the Value of Political Connections. *American Economic Review*, 91(4):1095–1102.

Fisman, R., Schulz, F., and Vig, V. (2014). The Private Returns to Public Office. *Journal of Political Economy*, 122(4):806–862.

Forbes, S. J. and Lederman, M. (2009). Adaptation and vertical integration in the airline industry. *American Economic Review*.

Forbes, S. J. and Lederman, M. (2010). Does vertical integration affect firm performance? evidence from the airline industry. *The RAND Journal of Economics*, 41(4):pp. 765–790.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, 98(1):394–425.

Frías, J. A., Kaplan, D. S., and Verhoogen, E. (2009). Exports and wage premia: Evidence from mexican employer-employee data. mimeo Columbia University.

Gaulier, G. and Zignago, S. (2010). BACI: International Trade Database at the Product-Level. The 1994-2007 Version. Working Papers 2010-23, CEPII research center.

Gibbons, R. (2005a). Four formal(izable) theories of the firm? *Journal of Economic Behavior and Organization*, 58(2):200 – 245.

Gibbons, R. (2005b). Incentives between firms (and within). *Management Science*.

Gibbons, R. and Roberts, D. (2013). *Handbook of Organizational Economics*. Princeton University Press.

Gil, R. (2009). Revenue sharing distortions and vertical integration in the movie industry. *Journal of Law, Economics, and Organization*.

Gil, R. and Kim, M. (2016). Does competition increase quality? evidence from the us airline industry. mimeo, Carey School of Business.

Gil, R., Kim, M., and Zanarone, G. (2016). The value of relational adaptation in outsourcing: Evidence from the 2008 shock to the us airline industry. mimeo, Carey School of Business.

Giulietti, M., Waterson, M., and Wildenbeest, M. (2014). Estimation of search frictions in the b ritish electricity market. *The Journal of Industrial Economics*, 62(4):555–590.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., and Topalova, P. (2009). Trade Liberalization and New Imported Inputs. *American Economic Review: Papers & Proceedings*, 99(2):494–500.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., and Topalova, P. (2010). Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. *Quarterly Journal of Economics*, 125(4):1727–1767.

Goldberg, P. K. and Pavcnik, N. (2007). Distributional effects of globalization in developing countries. *Journal of Economic Literature*, 45:39–82.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2017). Bartik instruments: What, when, why, and how. Mimeo, Yale.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4):pp. 691–719.

Hallak, J. (2006). Product quality and the direction of trade. *Journal of International Economics*.

Hallak, J. C. and Schott, P. (2011). Estimating cross-country differences in product quality. *Quarterly Journal of Economics*.

Halpern, L., Koren, M., and Szeidl, A. (2015a). Imported inputs and productivity. *American Economic Review*.

Halpern, L., Koren, M., and Szeidl, A. (2015b). Imported Inputs and Productivity. *American Economic Review*, 105(12):3660–3703.

Hart, O. and Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy*, 98(6):pp. 1119–1158.

Hart, O., Shleifer, A., and Vishny, R. W. (1997). The proper scope of government: theory and an application to prisons. *The Quarterly Journal of Economics*, 112(4):1127–1161.

Holmstrom, B. (1999). The firm as a subeconomy. *Journal of Law, Economics, and Organization*.

Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, pages 24–52.

Holmstrom, B. and Milgrom, P. (1994). The firm as an incentive system. *American Economic Review*.

Holmstrom, B. and Tirole, J. (1991). Transfer pricing and organizational form. *Journal of Law, Economics, and Organization*.

Hong, H. and Shum, M. (2006). Using price distributions to estimate search costs. *The RAND Journal of Economics*, 37(2):257–275.

Hortacsu, A. and Syverson, C. (2004). Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds. *The Quarterly Journal of Economics*, 119(2):403–456.

Hortacsu, A. and Syverson, C. (2007). Cementing relationships: Vertical integration, foreclosure, productivity, and prices. *Journal of Political Economy*.

Hottman, C. J., Redding, S. J., and Weinstein, D. E. (2016). Quantifying the Sources of Firm Heterogeneity *. *The Quarterly Journal of Economics*, 131(3):1291–1364.

Hummels, D. and Klenow, P. (2005). The variety and quality of a nation's exports. *American Economic Review*.

IFFO (2014). Chile fishing quotas set and landing information certification. http://www.iffo.net/node/544.

IRF (2017). Catch quotas for 2017/2018 in accordance with scientific advice. https://www.responsiblefisheries.is/news-and-media/news/catch-quotas-for-2017-2018-in-accordance-with-scientific-advice/328/.

James, H. (2010). Resources, Rent-Seeking, and Reform in Thailand and Myanmar (Burma): The Economics-Politics Nexus. *Asian Survey*, 50(2):426–448.

Johnson, R. C. (2012). Trade and prices with heterogeneous firms. *Journal of International Economics*.

Jones, L. (2014). The Political Economy of Myanmar's Transition. *Journal of Contemporary Asia*, 44(1):144–170.

Kamal, F. and Sundaram, A. (2016). Buyer–seller relationships in international trade: Do your neighbors matter? *Journal of International Economics*, 102:128–140.

Kang, D. C. (2003). Transaction Costs and Crony Capitalism in East Asia. *Comparative Politics*, 35(4):439–458.

Khandelwal, A. (2010). The long and short (of) quality ladders. *The Review of Economic Studies*, 77(4):1450–1476.

Khandelwal, A. K., Schott, P. K., and Wei, S.-J. (2013). Trade liberalization and embedded institutional reform: evidence from chinese exporters. *American Economic Review*, 103(6):2169–95.

Kosová, R., Lafontaine, F., and Perrigot, R. (2013). Organizational form and performance: Evidence from the hotel industry. *Review of Economics and Statistics*.

Kranton, R. and Swamy, A. V. (2008). Contracts, hold-up, and exports: textiles and opium in colonial india. *American Economic Review*, 98(3):967–89.

Krishna, V. (2009). *Auction theory*. Academic press.

Kugler, M. and Verhoogen, E. (2009). Plants and Imported Inputs: New Facts and an Interpretation. *American Economic Review: Papers & Proceedings*, 99(2):501–507.

Kugler, M. and Verhoogen, E. (2012). Prices, Plant Size, and Product Quality. *Review of Economic Studies*, 79(1):307–339.

Kuksov, D. (2004). Buyer search costs and endogenous product design. *Marketing Science*, 23(4):490–499.

Lafontaine, F. and Slade, M. (2007). Vertical integration and firm boundaries: The evidence. *Journal of Economic Literature*, 45(3):629–685.

Larkin, S. (2015). Myanmar's Tycoons: Vested Interests Resisting Reform or Agents of Change? *ISEAS Perspective*, 39:1–14.

Legros, P. and Newman, A. F. (2013). A price theory of vertical and lateral integration. *Quarterly Journal of Economics*.

Macchiavello, R. (2011). Financial development and vertical integration: theory and evidence. *Journal of the European Economic Association*.

Macchiavello, R. and Miquel-Florensa, J. (2016). Vertical integration and relational contracts: Evidence from costa rica's coffee chain. Unpub. paper.

Macchiavello, R. and Morjaria, A. (2015). The value of relationships: Evidence from a supply shock to kenya rose exports. *American Economic Review*.

Manova, K. and Zhang, Z. (2012). Export prices across firms and destinations. *The Quarterly Journal of Economics*, 127(1):379–436.

Martinez-Carrasco, J. (2017). Coordination and gains from relational contracts: Evidence from the peruvian anchovy fishery. Mimeo, Duke University.

McMillan, J. and Woodruff, C. (1999). Interfirm relationships and informal credit in vietnam. *Quarterly Journal of Economics*.

Medina, P. (2017). Import competition, quality upgrading and exporting: Evidence from the peruvian apparel industry. Mimeo, University of Toronto.

Meehan, P. (2011). Drugs, insurgency and state-building in Burma: Why the drugs trade is central to Burma's changing political order. *Journal of Southeast Asian Studies*, 42(3):376–404.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Mullainathan, S. and Sharfstein, D. (2001). Do firm boundaries matter? *American Economic Review*.

Myoe, M. A. (2009). *Building the Tatmadaw: Myanmar Armed Forces Since 1948*. Institute of Southeast Asian Studies, Singapore.

Natividad, G. (2014). Integration and productivity: Satellite-tracked evidence. *Management Science*, 60(7):1698–1718.

Novak, S. and Stern, S. (2008). How does outsourcing affect performance dynamics? evidence from the automobile industry. *Management Science*, 54(12):pp. 1963–1979.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The Quarterly Journal of Economics*, 122(2):569–600.

Ordover, J., Saloner, G., and Salop, S. (1990). Equilibrium vertical foreclosure. *American Economic Review*.

Paredes, E. C. and Gutierrez, M. E. (2008). The peruvian anchovy sector: Costs and benefits. an analysis of recent behavior and future challenges. IIFET 2008 Vietnam Proceedings.

Park, A., Yang, D., Shi, X., and Jiang, Y. (2009). Exporting and Firm Performance: Chinese Exporters and the Asian Financial Crisis. NBER Working Papers 14632.

Pavcnik, N. (2002). Trade Liberalization , Exit , and Productivity Improvements : Evidence from Chilean Plants Nina Pavcnik. *Review of Economic Studies*, 69(1):245–276.

Pigou, A. C. (1912). *Wealth and Welfare*. London: Macmillan.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of international Economics*, 48(1):7–35.

Redding, S. and Weinstein, D. (2018). Accounting for trade patterns. *Princeton University, mimeograph*.

Rijkers, B., Freund, C., and Nucifora, A. (2017). All in the family: State capture in Tunisia. *Journal of Development Economics*, 124:41–59.

Ruhl, K. J. (2015). How well is u.s. intrafirm trade measured? *American Economic Review: Papers and Proceedings*.

Salz, T. (2018). Intermediation and competition in search markets: An empirical case study.

Schott, P. (2004). Across-product versus within-product specialization in international trade. *Quarterly Journal of Economics*.

Startz, M. (2016). The value of face-to-face: Search and contracting problems in nigerian trade. *Available at SSRN 3096685*.

Stigler, G. J. (1961). The economics of information. *Journal of political economy*, 69(3):213–225.

Stiglitz, J. E. (1987). Competition and the number of firms in a market: Are duopolies more competitive than atomistic markets? *Journal of political Economy*, 95(5):1041–1061.

Sutton, J. (1991). *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. MIT Press.

Sutton, J. (1998). *Technology and Market Structure: Theory and History*. MIT Press.

Syverson, C. (2004). Product Substitutability and Productivity Dispersion. *The Review of Economics and Statistics*, 86(2):534–550.

Tanoue, W. (2015). Japan's total allowable catch systems in fishery resource management. `https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/34012/Tanoue_washington_0250O_14735.pdf?sequence=1&isAllowed=y`.

Tintelnot, F., Mogstad, K. K. M., and Dhyne, E. (2017). Trade and domestic production networks. Mimeo, Chicago University.

Topalova, P. and Khandelwal, A. (2011). Trade Liberalization and Firm Productivity: The Case of India. *Review of Economics and Statistics*, 93(3):995–1009.

Trefler, D. (2004). The Long and Short of the Canada-U.S. Free Trade Agreement. *American Economic Review*, 94(4):870–895.

Tveteras, S., Paredes, C. E., and Torres, J. P. (2011). Individual vessel quotas in peru: Stopping the race for anchovies. *Marine Resource Economics*, 26:225–232.

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the mexican manufacturing sector. *The Quarterly Journal of Economics*, 123(2):489–530.

Williamson, O. (1971). The vertical integration of production: Market failure considerations. *The American Economic Review*, 61(2):pp. 112–123.

Woodruff, C. (2002). Non-contractible investments and vertical integration in the mexican footwear industry. *International Journal of Industrial Organization*.

Woods, K. (2011). Ceasefire capitalism: military-private partnerships, resource concessions and military-state building in the Burma-China borderlands. *Journal of Peasant Studies*, 38(4):747–770.

Yoichi, S., Kensuke, T., and Enrique, S. (2016). Assortative matching of exporters and importers. Technical report, Institute of Developing Economies, Japan External Trade Organization (JETRO).

Appendix A

Appendix to Chapter 1

A.1 Estimation procedure in details

This appendix describes in more details how the marginal cost parameters are estimated. As explained in the text, the challenge of the estimation lies in translating the parameters $mc_{i,s}$ and v that define the distribution $\mathcal{F}(.|mc_{i,s})$ into the parameters that define $\mathcal{H}_{ijs}(.)$ and $\widetilde{\mathcal{H}}_{ijs}(.)$ as defined in equation 4.

First, it is worth noting that the optimal bidding strategy $\beta_{j,s}(mc) = \mathbb{E}[Y_j|Y_j > mc]$ is the expectation of the truncated distribution of the first-order statistic of $N_{j,s} - 1$ draws of the normal distribution. While the distribution of the minimum of random draws of the normal distribution is not a well-known function form, I use the results from Chen and Tyler (1999) which provide a very good approximation. If $X \sim \mathcal{N}(\mu, \sigma^2)$, the first order statistic can be approximated by a normal with mean $\mu(N) = \mu - \Phi^{-1}(0.5264^{1/N}) \cdot \sigma$ and standard deviation $\sigma(N) = 1/2 \cdot \sigma \cdot \left(\Phi^{-1}(0.8832^{1/N}) - \Phi^{-1}(0.2142^{1/N})\right)$, where Φ^{-1} is the inverse of the standard normal distribution. Thus, $\beta_{j,s}$ is just the truncated mean of a normal distribution and so can be approximated by the following expression:

$$\beta_{j,s}(mc) = m(N_{j,s} - 1) + \frac{\phi\left(\frac{mc - \mu(N_{j,s}-1)}{\sigma(N_{j,s}-1)}\right)}{1 - \Phi\left(\frac{mc - \mu(N_{j,s}-1)}{\sigma(N_{j,s}-1)}\right)} \cdot \sigma(N_{j,s} - 1)$$

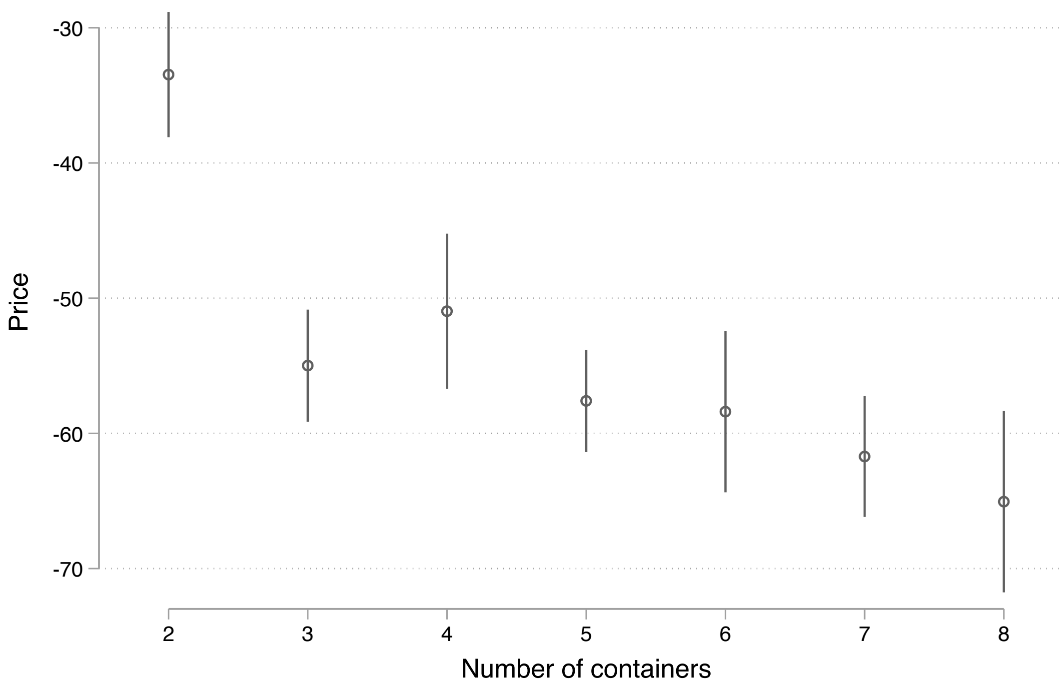
where ϕ and Φ are respectively the pdf and cdf of the standard normal distribution.

Using the Taylor expansions for the moments of functions of random variables in Benaroya et al. (2005), $\mathcal{H}_{ijs}(.)$ can then be approximated by a normal $\mathcal{N}(\mu_{ijs}, \sigma_{ijs})$ with $\mu_{ijs} = \beta_{j,s}(mc_{i,s})$ and $\sigma_{ijs} = \beta'_{j,s}(mc_{i,s}) \cdot v$.

Finally, I need to translate the parameters that describe $\mathcal{H}_{ij}(.)$ the distribution of the actual bids p_{ijt} into the distribution $\widetilde{\mathcal{H}}_{ij}(.)$ of the bids with errors in the data $\widetilde{p}_{ijt} = p_{ijt} + \epsilon_{ijt}$. As $[1 - \mathcal{H}_{kj}(p_{ijt})]$ corresponds the probability that $p_{kjt} > p_{ijt} = \widetilde{p}_{ijt} - \epsilon_{ijt}$, then $[1 - \mathcal{H}_{kj}(p_{ijt})] = [1 - \widetilde{\mathcal{H}}_{kj}(\widetilde{p}_{ijt})]$ where $\widetilde{\mathcal{H}}_{kj}$ is normal with mean μ_{ijs} and variance $\sigma^2_{ijs} + \sigma^2_e$.

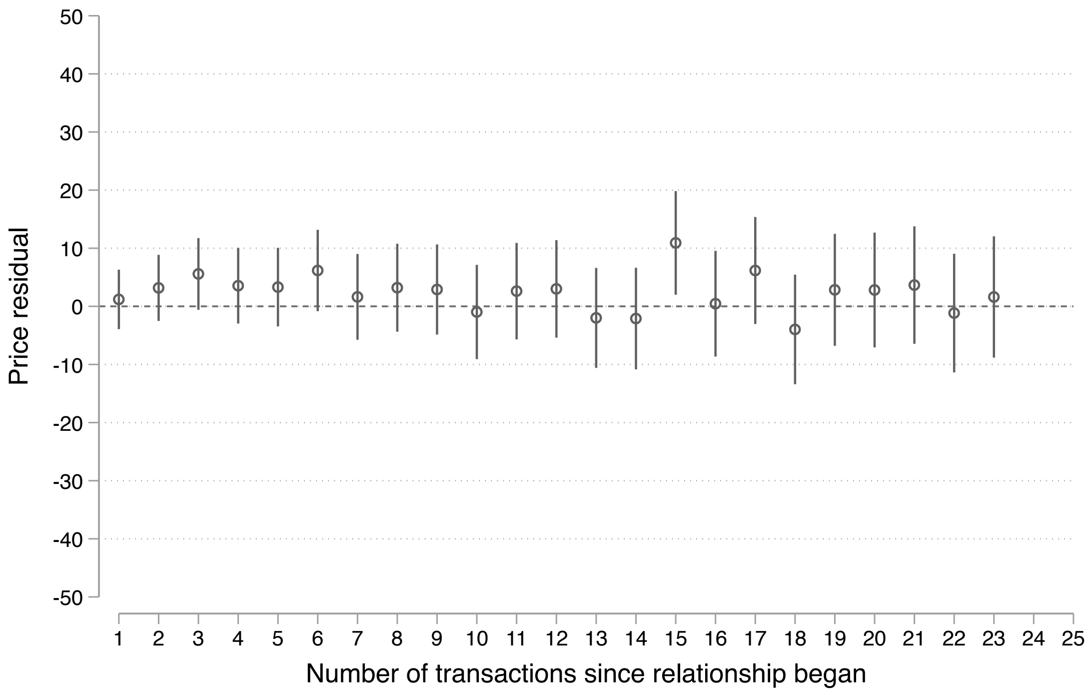
A.2 Additional Figures

Figure A.1: Quantity bins fixed effects



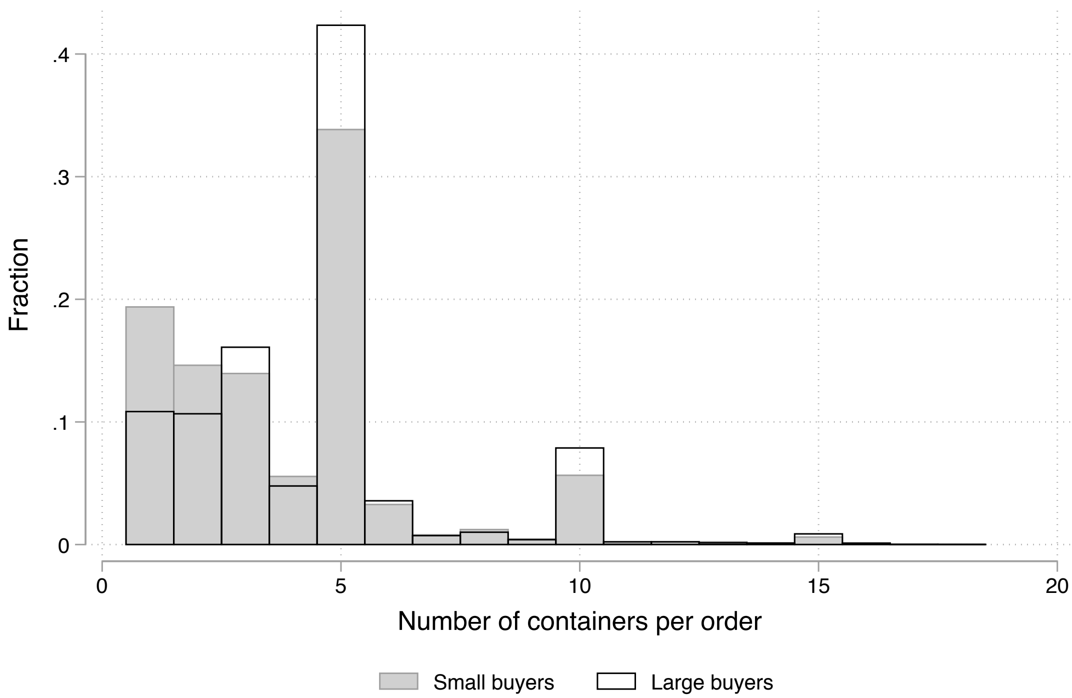
Notes: This figure shows the price discounts estimates obtained from the regression in 1. The x-axis variable represents the size of the shipment in containers. The omitted bin is a size of one container.

Figure A.2: Prices over time within a buyer-exporter relationship



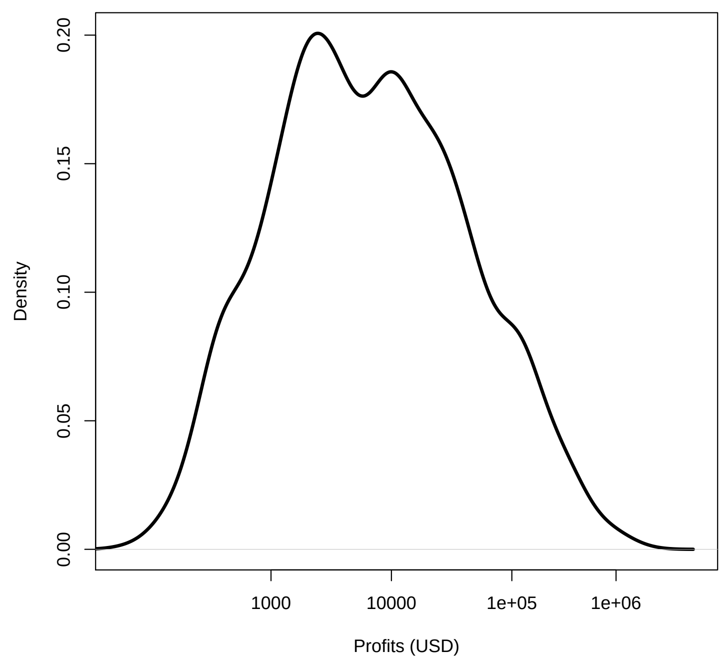
Notes: This figure shows transaction prices as a function of the length of a buyer-exporter relationship. If switching costs were important in explaining the observed price dispersion documented in section 1.3.1, prices should increase with the length of the relationship.

Figure A.3: Distribution of order sizes for small and large buyers



Notes: This figure shows the distribution of order size classifying buyers by their overall size. Small buyers are defined as buyers who make less than 25 shipments per year.

Figure A.4: Distribution of profits



Notes: This figure shows the distribution of profits (as defined as the sum of the expected difference between transaction price and marginal costs).

Appendix B

Appendix to Chapter 2

B.1 Additional Tables and Figures

Table B.1: Main importers of Peruvian fishmeal and Average Quality Imported

	Total Weight (1000 metric tons)	Average Protein content	Sd(Protein content)
CHINA	4266	66.06	1.60
GERMANY	972	65.42	1.62
JAPAN	545	66.12	1.69
CHILE	305	66.60	1.51
VIETNAM	277	65.91	1.59
TAIWAN	248	66.02	1.71
UNITED KINGDOM	147	65.26	1.62
TURKEY	128	64.91	1.52
INDONESIA	94	66.16	1.64
SPAIN	90	65.44	1.61
AUSTRALIA	85	66.06	1.80
CANADA	66	65.76	1.52
FRANCE	55	65.59	1.72
SOUTH KOREA	24	66.56	1.46
ITALY	21	64.97	1.52
BULGARIA	15	65.42	1.75
VENEZUELA	13	66.67	1.64
PHILIPPINES	12	64.92	1.47
BELGIUM	11	65.08	1.69
INDIA	10	65.17	2.03

Notes: This table reports the top 20 importers of Peruvian fishmeal, the total quantity imported over the whole period of our sample, the average quality imported and the standard deviation of the quality imported across all transactions.

Table B.2: Vertical integration and the Quality price premium – First Stage

Dep. var:	LQP (t-1) × Quality Premium	USP (t-1) × Quality Premium	LQP (t-1) × Quality Premium	USP (t-1) × Quality Premium
	(1)	(2)	(3)	(4)
LQ producer (t-1) × [Log(Qty exp. by Chile)] (t-1)	-0.081			
	(0.097)			
LQP (t-1) × [Log(Qty exp. by Denmark)] (t-1)	0.142^{***}			
	(0.036)			
LQP (t-1) × [Log(Qty exp. by Iceland)] (t-1)	0.351^{***}			
	(0.046)			
LQP (t-1) × [Log(Qty exp. by Thailand)] (t-1)	0.133^{***}			
	(0.012)			
USP (t-1) × [Log(Qty exp. by Chile)] (t-1)		-0.410		
		(0.245)		
USP (t-1) × [Log(Qty exp. by Denmark)] (t-1)		0.249^{**}		
		(0.116)		
USP (t-1) × [Log(Qty exp. by Iceland)] (t-1)		0.433^{**}		
		(0.165)		
USP (t-1) × [Log(Qty exp. by Thailand)] (t-1		0.104^{***}		
		(0.034)		
LQP (t-1) × [Log(Chile Price)]			1.175^{***}	
			(0.044)	
LQP (t-1) ×[Log(Denmark Price)]			-2.158^{***}	
			(0.534)	
LQP (t-1) × [Log(Iceland Price)]			0.667	
			(0.486)	
USP (t-1) × [Log(Chile Price)]				1.081^{***}
				(0.135)
USP (t-1) ×[Log(Denmark Price)]				-2.950^{***}
				(0.740)
USP (t-1) × [Log(Iceland Price)]				1.469^{**}
				(0.558)
Low quality producer (t-1)	Yes	No	Yes	No
Upgradable share of production (t-1)	No	Yes	No	Yes
Season FEs	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes

Notes: Results from the first stage of IV specifications reported in Table 2.4. One observation is a firm during a production season. *Quality premium* is equal to $Log(\text{High Quality}) - Log(\text{Low Quality})$ where High and Low Quality are the average price of "Prime" and "FAQ" fishmeal in the month preceding the current fishing season. LQP$(t-1)$ is equal to 1 if a firm's share of low quality output in the preceding season was at least 1 percent. USP$(t-1)$ is the share of a firm's production that was of low quality in the previous season. A firm that produces almost only low quality output has more potential to upgrade than a firm already producing almost only high quality output. In column 1 and 3, the instruments are LQP$(t-1)$ and USP$(t-1)$ interacted with average export prices in other high quality fishmeal exporting countries for which the data was available. In Column 2 and 4, the instruments are LQP$(t-1)$ and USP$(t-1)$ interacted with the quantity exported by the top fishmeal exporting countries (excluding Peru) for which there is a national fishing quota. Standard errors clustered at the firm level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: Vertical integration and the Quality price premium – Robustness Checks

Dep. var:	Share of inputs from VI suppliers								
	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Quality premium	0.832**		0.333						
	(0.303)		(0.199)						
Quality premium is high		0.375***							
		(0.061)							
Quality premium is low		0.294***							
		(0.061)							
Low quality producer (t-1)				0.218		0.279		0.208	
× Quality premium				(0.277)		(0.298)		(0.284)	
Upgradable share of production (t-1)					0.423		0.406		0.457
× Quality premium					(0.411)		(0.470)		(0.416)
Low quality producer (t-1)	No	No	No	Yes	No	Yes	No	Yes	No
Upgradable share of production (t-1)	No	No	No	No	Yes	No	Yes	No	Yes
Season FEs	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
N	203	203	203	203	203	203	203	203	203

Notes: One observation is a firm during a production season. Δ *Share VI* is *Share of inputs from VI suppliers (t) - Share of inputs from VI suppliers* $(t-1)$ as defined in 2.4. *Quality premium* is equal to $Log(\text{High Quality}) - Log(\text{Low Quality})$ where High and Low Quality are the average price of "Prime" and "FAQ" fishmeal in the month preceding the current fishing season. We choose to take the month preceding the fishing season rather than the fishing season itself as integration decisions are typically decided in the month preceding the season and integration within a season is extremely rare in the data. Low Quality Producer$(t-1)$ is equal to 1 if a firm's share of low quality output in the preceding season was at least 1 percent. Upgradable share of production$(t-1)$ is the share of a firm's production that was of low quality in the previous season. A firm that produces almost only low quality output has more potential to upgrade than a firm already producing almost only high quality output. Standard errors clustered at the firm level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4: Supplier characteristics

	Offload weight per trip (metric tons)	Cooling system	Capacity (m3)	Power engine (hp)	Max. Dist. from the plant's port (kms)
Wooden	41.00	0.00	65.73	215.40	56.10
	(16.24)	(0.06)	(27.34)	(94.78)	(7.74)
Steel - Independent	104.03	0.09	219.30	412.31	81.15
	(40.77)	(0.28)	(84.35)	(189.82)	(13.43)
Steel - Switchers	148.88	0.25	301.18	616.30	92.25
	(0.43)	(0.444)	(129.92)	(328.51)	(15.37)
Steel - VI	181.62	0.34	382.00	769.96	97.29
	(68.13)	(0.47)	(137.11)	(352.52)	(12.62)

Notes: *Offload weight* is the amount fished on a trip. *Maximum distance from port* is the maximum distance at which a boat is from the port on a fishing trip. Steel boats are generally bigger, better suited for industrial fishing, and are subject to different regulations. Wooden boats cannot be owned by fishmeal firms. *Independent* boats are owned by an individual or a company that is not a fishmeal company. *Switchers* are boats that move from VI to Independent or from Independent to VI at some point in our data. *VI* are boats that remain vertically integrated during the whole sample of our data.

Table B.5: Output Quality and Share of Inputs from Vertically Integrated Suppliers - Robustness checks using a major firm in the Peruvian fishmeal industry.

	Impact of Share of VI on Quality			
Dep. var:	Protein Content			
	OLS: Ind. Boats		IV: Ind. Boats	
	(1)	(2)	(3)	(4)
Share of inputs from VI suppliers	1.369**	1.338**	1.469*	1.390
	(0.654)	(0.656)	(0.807)	(0.918)
Has high technology	No	Yes	No	Yes
Month FEs	Yes	Yes	Yes	Yes
Plant FEs	Yes	Yes	Yes	Yes
N	2487	2487	2647	2647
	First Stage			
Dep. var:	Share of Inputs from VI Suppliers			
	(1)	(2)	(3)	(4)
Number of Independent Boats in Port			−0.000	−0.000
			(0.000)	(0.000)
Share of Independent Boats in Port			−0.412**	−0.398*
			(0.200)	(0.207)
Kleibergen-Paap LM p-value (Under-id)			0.005	0.006
Kleibergen-Paap Wald F statistic (Weak inst)			3.61	3.06
Anderson-Rubin Wald test p-value			0.24	0.31

Notes: This table presents robustness checks for Table 2.6 using a major firm's internal data that allows us to link export sales to a specific plant. *Share of inputs from VI suppliers* is the share of a firmÕs inputs that come from VI suppliers during a season. *Has high technology* is a dummy variable equal to one if the plant in question has any steam drying technologies installed. *Share of inputs from VI suppliers* is instrumented by (a) the number of independent boats present in the plant's port in the season in question, excluding those that interact directly with the plant itself, and (b) the ratio of the number of boats in (a) to the total number of boats in the plant's port in that season that do not interact with the plant itself. Robust standard errors are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: Output Quality and Share of Inputs from Vertically Integrated Suppliers - First Stage

	First Stage			
Dep. var:	Share of Inputs From VI Suppliers			
	IV: Ind. Boats		IV: Wooden Boats	
	(1)	(2)	(3)	(4)
Number of Independent Boats in Port (Leave-Out)	-0.001^{***}	-0.001^{***}		
	(0.000)	(0.000)		
Number of Wooden Boats in Port (Leave-Out)			-0.002^{***}	-0.002^{***}
			(0.000)	(0.000)
Kleibergen-Paap LM p-value (Under-id)	0.011	0.011	0.010	0.010
Kleibergen-Paap Wald F statistic (Weak inst)	23.55	23.73	27.69	27.69
Anderson-Rubin Wald test p-value	0.01	0.01	0.01	0.01
Has high technology	No	Yes	No	Yes
Month FEs	Yes	Yes	Yes	Yes
Plant FEs	Yes	Yes	Yes	Yes
N	2487	2487	2647	2647

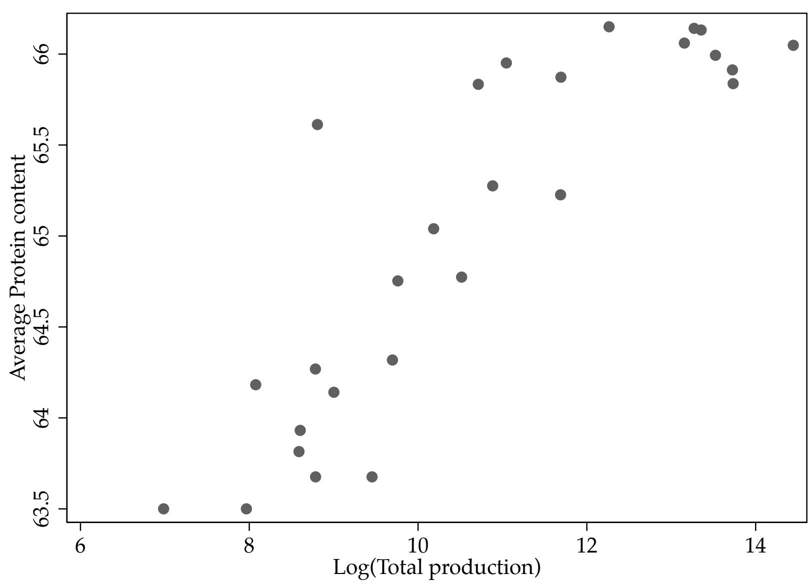
Notes: Results from the first stage of IV specifications reported in Table 2.6. One observation is a plant in a particular month. *Share of inputs from VI suppliers* is the share of a firmÕs inputs that come from VI suppliers during a season. *Has high technology* is a dummy variable equal to one if the plant in question has any steam drying technologies installed. Columns 1 and 2 instrument for *Share of inputs from VI suppliers* with the number of independent boats present locally (in the plant's port) in the season in question, excluding those that interact directly with the plant itself. Columns 3 and 4 instrument for *Share of inputs from VI suppliers* with the number of wooden boats present locally (in the plant's port) in the season in question, excluding those that interact directly with the plant itself. Standard errors are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Countries' output quality and vertical integration in export manufacturing

Dep. var:	Log(unit value) - Residuals from HS6×Year FEs and Country FEs
	(1)
Related party share of imports	0.038***
- Residuals from HS6×Year FEs and Country FEs	(0.007)
N	208 024

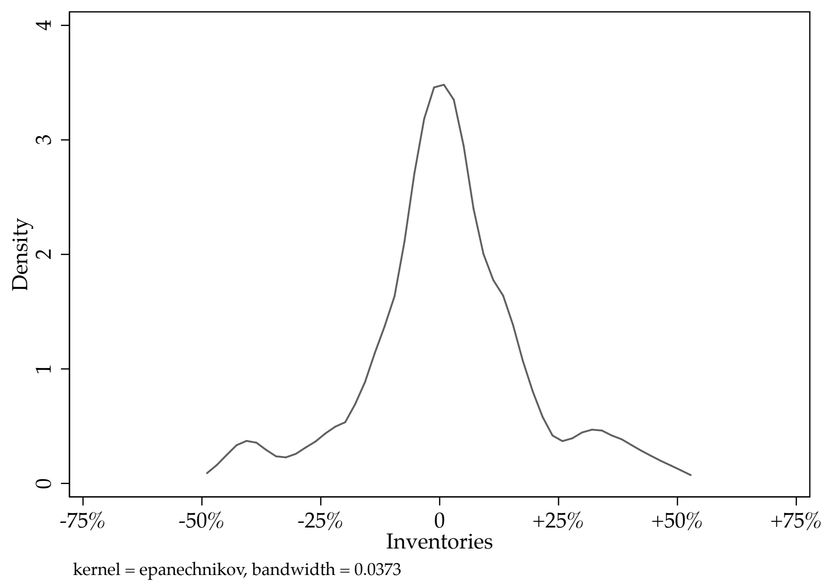
Notes: In this table, the dependent variable is ε_{cpt} from the regression log(unit value)$_{cpt} = \alpha_{pt} + \gamma_c + \varepsilon_{cpt}$, where log(unit value)$_{cpt}$ is the average log unit value of products exported from country c, of HS6 code p, in year t to the U.S.; α_{pt} is a product×year fixed effect; and γ_c is an origin country fixed effect. This regression is estimated using COMTRADE data from BACI (See Gaulier and Zignago (2010) for a description of the data). The independent variable is v_{cpt} from the regression Related party share of U.S. imports$_{cpt} = \beta_{pt} + \delta_c + v_{cpt}$, where Related party share of U.S. imports$_{cpt}$ is the share of products exported from country c, of NAICS code p, in year t to the U.S. that are imported by related parties (usually other units of the same firm (Ruhl, 2015)); β_{pt} is a product×year fixed effect; and δ_c is an origin country fixed effect. Related party share of U.S. imports$_{cpt}$ is constructed using data from the U.S. Census Bureau. Because the product level c (HS6) for the unit value residual is different from the product level p (NAICS) from the share of related party imports residuals, we compute the value weighted unit value residual at the p (NAICS) level using a HS6-NAICS conversion table. This regression includes data from 2005 to 2014. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.1: Average output quality and firm size



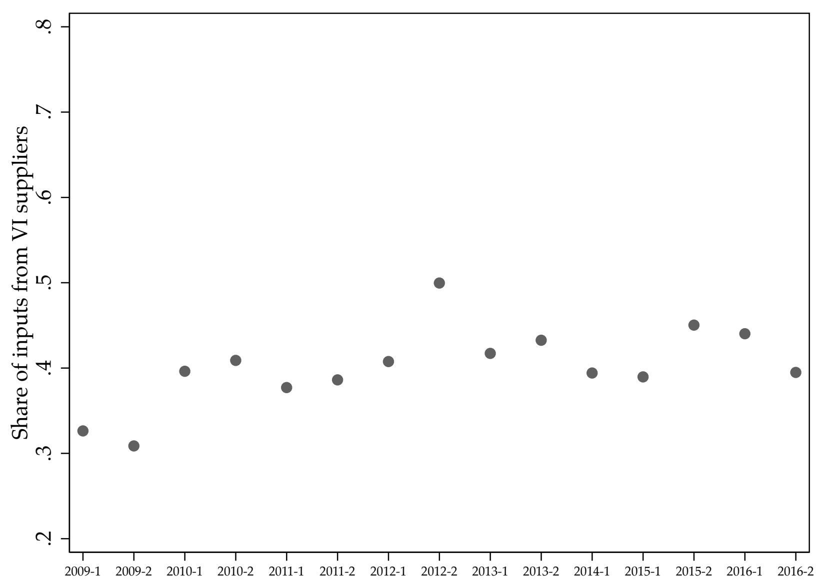
Notes: Each dot represents one fishmeal firm in our sample. Total production is the total weight of fishmeal the firm produced during our data period and average protein content is the quantity weighted average protein content of the firm's fishmeal exports.

Figure B.2: Density of inventories



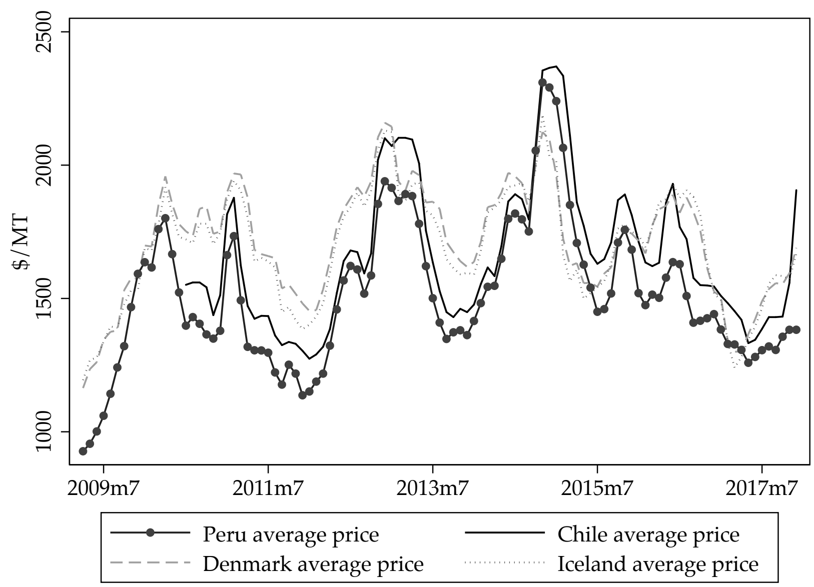
Notes: Kernel density of estimated inventories. Inventories are defined as the ratio of (Total Production - Total Exports) to Total Production, where Total Production is a firm's production during a given production season and Total Exports are the sum of exports that are shipped during the production season and the period directly following the relevant production season (before the next production season starts).

Figure B.3: Evolution of the Vertically Integrated share of inputs Industry-wide



Notes: This graph shoes the evolution of the Peruvian fishmeal industry's share of inputs from integrated suppliers by production season. For every year, -1 is the first production season in the calendar year, in general from April to July, and -2 is the second production season, in general from November to January.

Figure B.4: Fishmeal prices in Peru and other countries



Notes: This figure shows the evolution of the average fishmeal price in other fishmeal exporting countries. Denmark and Iceland export mostly “Super Prime” fishmeal grade while Peru exports mostly “Prime” grade fishmeal.

B.2 Dynamic Theoretical Framework and Relational Contracts

Dynamic theoretical framework

The model presented in the main body of the paper assumes that all transactions are done on the spot market. This stylized version of the model results in the upstream party not taking any action when integrated and the absence of incentives to take a quality-increasing action ($a_2 = 0$). In this version of the model, we follow closely Baker et al. (2001, 2002) in allowing the downstream party to use relational contracts to incentivize the quality action.

We make the same assumptions for Q as before, but add a shock to the alternative use P:

$$P = a_1 + \epsilon$$
$$Q = Q_0 - \gamma a_1 + \delta a_2$$

where ϵ is orthogonal to any action taken by the upstream party[1]. We assume that $\epsilon = \bar{\epsilon}$ with probability $\frac{1}{2}$ and $\epsilon = -\bar{\epsilon}$ with probability $\frac{1}{2}$ and that ϵ is known by the upstream party at the time of delivery of the inputs. [2]

As in the main text model, we assume that both P and Q are not contractible. P - the quantity focused alternative use- is perfectly observable at the time of delivery of the inputs, but Q -the quality surplus- is only observed to the downstream party with

[1] We could also assume uncertainty over the realization of the Q surplus, but it would not change the intuition of the result below.

[2] As in the main text model, we assume that $0 \leq \delta \leq 1$ and $0 \leq \gamma \leq 1 - \alpha$. Also, note again that P could itself be the result of a bargaining process between the boat and a quantity focused firm.

some delay (e.g. once the inputs are processed).[3] To incentivize the quality-increasing action, the downstream party can offer a payment contingent on the realization of the surplus Q to the upstream party. However, since this payment can only be made after the inputs are delivered, the downstream party can only credibly promise to make this delayed payment through repeated interactions with the upstream party.[4] Note again that at the time of delivery of the inputs, since all parties know the value of Q_0, and because $P = a_1 + \epsilon$ is observable, Q has an observable portion (in expectation) at the time of delivery of the inputs: $\tilde{Q} = Q_0 - \gamma \mathbb{E}(a_1|P) = Q_0 - \gamma P$. Hence, a payment on the spot, proportional to $\tilde{Q}$ is still feasible.

As in Baker et al. (2002), we consider four possible organizational structures:

1. Spot Outsourcing (Nonintegrated Asset Ownership, Spot Governance Environment)

2. Relational Outsourcing (Nonintegrated Asset Ownership, Relational Governance Environment)

3. Spot Employment (Integrated Asset Ownership, Spot Governance Environment)

4. Relational Employment (Integrated Asset Ownership, Relational Governance Environment)

We write the relational compensation contract as $\{b(Q)\}$, where $b(Q)$ is a payment contingent on the observation of Q[5].

---

[3]In our context, fish quality can hardly be assessed when the fish is offloaded at the factory. However, once the fish is processed in the factory, fishmeal quality can be measured.

[4]In the model, we suppose that this delay is shorter than a full time period, so the surplus Q is observed before the next period starts and the next transaction occurs. Thus, the downstream party does not discount the payment.

[5]Alternatively, we could consider a more general relational compensation contract of the form $\{s, b(Q)\}$ as in Baker et al. (2002), where salary s is paid by downstream to upstream at the beginning of each period and $b(Q)$ is a payment contingent on the realization of Q. Such an assumption would not change our results below.

First Best

The first-best actions $\{a_1^*, a_2^*\}$ maximize the expected value of Q minus the cost of actions $c(a_1, a_2) = \frac{1}{2}a_1^2 + \frac{1}{2}a_2^2$. This gives $a_1^* = 0$ and $a_2^* = \delta$ and total surplus:

$$S^* = Q(a_1^*, a_2^*) - c(a_1^*, a_2^*) = Q_0 + \frac{1}{2}\delta^2$$

Spot Market

On the Spot Market, the supplier does not take the first best actions. In particular, under both Spot Employment and Spot Outsourcing $a_2 = 0$, because the downstream firm cannot credibly commit to rewarding the supplier's quality-focused actions.

Relational Contracts

Whether the upstream party is integrated with the downstream party or not, if she accepts the relational contract, she will choose actions a_1 and a_2 to solve:

$$\max_{a_1, a_2} = b\big(Q(a_1, a_2)\big) - c(a_1, a_2)$$

It is straightforward to see that the first best can only be achieved if the contract is of the form $b\big(Q(a_1, a_2)\big) = Q(a_1, a_2) - t$, where t is a transfer independant of the surplus Q. In the remainder of this section, we assume that the relational contract is written in such a way and that under relational employment (when the downstream party owns the supplier) or under relational outsourcing (when the supplier is independent), the suppliers take the first best actions $\{a_1^*, a_2^*\}$ [6].

This relational contract is self-enforcing if both parties choose to honor it for all pos-

[6] In particular, t must be such that $t \leq Q(a_1^*, a_2^*) - c(a_1^*, a_2^*) = Q_0 + \frac{1}{2}\delta^2$ so that the downstream party would accept the contract

sible realizations of P. We next explore the feasibility of the first best contract under employment and outsourcing and show that if the shock to the alternative use P is high enough, the first best contract is only self-enforceable under Relational Employment. We use superscripts $\{RE, SE, RO, SO\}$ to indicate Relational Employment, Spot Employment, Relational Outsourcing and Spot Outsourcing and $\{U, D, S\}$ to denote the upstream party, downstream party and overall surplus respectively.

Relational Employment

Since $S^{SE} > S^{SO}$, [7] if one of the two party reneges, the downstream party will retain ownership and earn D^{SE} in perpetuity, while the upstream party will earn U^{SE} in perpetuity. The upstream party reneges if she refuses to accept the promised payment $b(Q)$. Thus, the upstream party does not renege as long as:

$$b(Q) + \frac{1}{r}U^{RE} \geq \frac{1}{r}U^{SE} \tag{B.1}$$

Similary, the downstream party reneges if she takes the inputs and refuses to pay the bonus to the upstream party. The downstream party honors the contract as long as:

$$\frac{1}{r}D^{RE} \geq b(Q) + \frac{1}{r}D^{SE} \tag{B.2}$$

Summing (B.1) and (B.2), and noting that $S^X = U^X + D^X$, we get the following necessary condition:

$$S^{RE} \geq S^{SE} \tag{B.3}$$

(B.3) is actually sufficient as well as necessary, because a transfer t can always be chosen so that when (B.3) is statisfied, (B.1) and (B.2) are also satisfied [8].

[7] See the proof in the main text model.

[8] For both (B.1) and (B.2) to be satisfied and the supplier to accept the contract, it must be that $Q_0 + \frac{1}{2}\frac{r}{1+r}\delta^2 \leq t \leq Q_0 + \frac{1}{2}\delta^2$

As $S^{RE} = S^* = Q_0 + \frac{1}{2}\delta^2$ and $S^{SE} = S^* = Q_0$, (B.3) is satisfied, and so **the first best can always be enforced under Relational Employment**.

Relational Outsourcing

Since $S^{SE} > S^{SO}$, if one of the two party reneges, the upstream party will purchase the ownership right from the downstream party for some price π, after which the upstream and downstream parties will earn U^{SE} and D^{SE}, respectively, in perpetuity. If the upstream party reneges on the relational-outsourcing contract, she negotiates to sell the good for the spot-outsourcing price of $(1-\alpha)P + \alpha\tilde{Q}$, where α is the supplier's bargaining coefficient and $\tilde{Q}$ is the observable portion of the surplus Q as in the main text model. Thus, the upstream party honors the contract as long as:

$$b(Q) + \frac{1}{r}U^{RO} \geq (1-\alpha)P + \alpha\tilde{Q} + \frac{1}{r}U^{SE} + \pi \tag{B.4}$$

The timing of reneging is slightly different for the downstream party. She has no incentives to renege at the time of delivery of the inputs as Q is unobservable. Instead, the downstream party reneges if she takes the inputs and refuses to pay the bonus to the upstream party. The downstream party does not renege as long as:

$$\frac{1}{r}D^{RO} \geq b(Q) + \frac{1}{r}D^{SE} - \pi \tag{B.5}$$

If (B.4) holds for all P and $\tilde{Q}$, then it must hold for the maximum value of $(1-\alpha)P + \alpha\tilde{Q}$. Summing (B.4) and (B.5) we get the following necessary condition:

$$\frac{1}{r}S^{RO} \geq \frac{1}{r}S^{SE} + \max\{(1-\alpha)P + \alpha\tilde{Q}\} \tag{B.6}$$

Evaluated at $\{a_1^*, a_2^*\}$, (B.6) is equivalent to:

$$(1-\alpha\gamma-\alpha)\bar{\epsilon} \leq \frac{1}{2r}\delta^2 - \alpha Q_0 \tag{B.7}$$

Thus, if $\bar{\epsilon}$ is high enough, **the first best contract cannot be enforced under Relational Outsourcing**.

The intuition for why quality-oriented downstream firms may need to own upstream productive assets and hire the suppliers operating the assets as employees is as follows. Under any sort of outsourcing, suppliers are free to allocate the inputs produced to their alternative use. As a result, when the value of the input is high in its alternative use (e.g. if the supplier happens to get more fish or if there is less competition on a specific day in the quantity-focused sector), quality-oriented firms may be unable to prevent the suppliers they interact with from breaking their relationship and selling the goods for its alternative use. In contrast, under Relational Employment, the downstream firm has control over the inputs, and will choose to allocate them efficiently regardless of the value of the inputs in their alternative use.

A key testable prediction of this model in our context is that (1) independent suppliers under a relational contract should not adopt a behavior consistent with delivering higher quality inputs and (2) downstream firms should not produce higher quality output when they source more of their inputs from non-integrated suppliers with whom they have a relational contract.

Empirical evidence on relational contracts in the Peruvian fishmeal industry

We now test these predictions. We show results for two different, frequency-of-interacting based observable proxies for a supplier being engaged in a relational outsourcing contract with a downstream firm: specifically, (i) that the supplier delivers more than 80 percent of its fish to the same fishmeal firm (approx. the 75th percentile of the underlying distribution) for two consecutive production seasons, and (ii) that the supplier delivers to the same firm more than 10 times (approx. the 25th percentile of the underlying distribution) in a given production season and does so for three seasons in a row. We "turn on" the inferred contract at the start of the relevant period, not when the "cut-off" used in the proxy is reached.

In Appendix Table B.8, which is analogous to Table 2.5, we show that relational outsourcing contracts appear not to be used to incentivize supplier quality-increasing actions in the Peruvian fishmeal industry, consistent with the dynamic version of our theoretical framework above. The results show that a supplier supplying a given plant does not deliver fresher fish when engaged in repeated interactions with the firm in question, relative to more isolated instances of supplying the same plant.

In Appendix Table B.9, which is analogous to Table 2.6, we relate output quality not only to the share of inputs coming from integrated suppliers, but also to the share coming from suppliers under relational outsourcing contracts (as defined by the proxies described above). The estimated coefficients on the share of inputs coming from integrated suppliers remain positive and highly significant, while the estimated coefficients on the share coming from suppliers under relational outsourcing contracts are very small and insignificant. These results indicate that repeated interactions are not used to incentivize the delivery of high quality inputs in the Peruvian fishmeal sector, as the model above predicts.

In combination with the results in the body of the paper, the findings in tables 2.5 and 2.6 provide support for the idea that vertical integration enables downstream firms to incentivize specific supplier behaviors—and consequently the types of output associated with those behaviors—that other organizational structures do not.

Organizational structure and supplier behavioral response to plant input quality needs

The dynamic model with relational contracts presented above also predicts the following result. When the return on the quality surplus Q of the quality-increasing action is higher (when δ increases), integrated suppliers will choose a higher level of the that action ($a_2^* = \delta$ increases). We test this prediction below.

A change in the need for input quality arises when the plant aims to produce fishmeal of the high quality type (for example because of a change in demand). As in Section 2.6, we compare periods when the supplier is integrated with the plant supplied and periods when the supplier is independent from but supplies the same plant, but now differentially when the downstream plant produces a low or high quality output.

We first estimate the following equation:

$$\begin{aligned} \mathrm{B}_{ijt} = \alpha &+ \beta_1 \,\mathrm{I}[\text{VI} \times \text{supplies owner firm}]_{ijt} \times \mathrm{I}[\text{Low Quality}]_{jt} \\ &+ \beta_2 \,\mathrm{I}[\text{VI} \times \text{supplies owner firm}]_{ijt} \times \mathrm{I}[\text{High Quality}]_{jt} \\ &+ \gamma_{ij} \times \mathrm{I}[\text{High Quality}]_{jt} + \gamma_{ij} \times \mathrm{I}[\text{Low Quality}]_{jt} + \delta_t + \varepsilon_{ijt} \end{aligned} \tag{B.8}$$

where $\mathrm{I}[\text{Low Quality}]_{jt}$ is a dummy equal to 1 when plant j—i.e. the plant supplier i supplies at t—produces comparatively low quality fishmeal in the month date t falls within (and conversely for $\mathrm{I}[\text{High Quality}]_{jt}$).[9] We include Supplier × Plant × Quality level fixed effects (that is, $\gamma_{ij} \times \mathrm{I}[\text{High Quality}]_{jt}$ and $\gamma_{ij} \times \mathrm{I}[\text{Low Quality}]_{jt}$) to focus on the supplier's *differential* response to the plant's input needs when integrated. The other variables are as defined in equation (2.5).

The marginal impact of the behavioral response of a single supplier on the output quality of the plant as a whole is likely to be limited. We thus interpret the coefficient of interest as the supplier's response to the plant's *intention* to produce higher quality output.

The results in Appendix Table B.10 suggest that suppliers differentially adapt their quality behavior to the current needs of the downstream plant they supply when integrated. Column 1 shows that boats tend to deliver a lower quantity per trip when integrated with the plant supplied, regardless of whether the plant produces low or

[9]We define this dummy variable using our directly observed measure of quality at plant level. The dummy is equal to 1 if the share of the plant's production that is of high quality type is higher than the median in our sample.

high quality at the time.[10] However, columns 2 and 3 show that, when integrated, boats adjust their behavior so as to deliver fresher fish when the plant supplied is producing high quality output. When integrated, boats fish about seven percent closer to port and spend about six percent less time at sea, when the plant supplied is producing fishmeal of the high quality type Overall, the evidence confirms the prediction from the relational model that integrated suppliers will provide more of the quality focused action when its return to the quality surplus is higher.

[10]The estimated decrease in quantity per trip when integrating with the plant being supplied is bigger when the plant is producing low quality fishmeal. This is surprising in light of our results in sections 2.7 and 2.5. A possible explanation is that independent suppliers face strong incentives to deliver high input quantities when the plant being supplied is attempting to produce high output quantities (and prioritizing output quality less) and that integrated suppliers do not.

Table B.8: Supplier Behavior and relational outsourcing

Panel A: Relational outsourcing = 80% of offloads to the same firm for 2 consecutive production seasons			
Dep. var:	Log(Quantity supplied)	Log(Max. distance the plant's port)	Log(Total time spent at sea)
	(1)	(2)	(3)
I[Relational × supplies relational firm]	0.010	0.016*	−0.000
	(0.007)	(0.009)	(0.006)
Date FEs	Yes	Yes	Yes
Supplier × Plant FEs	Yes	Yes	Yes
N	315,442	137,278	159,724

Panel B: Relational Outsourcing = more than 10 interactions with the same firm for at least 3 consecutive production seasons			
Dep. var:	Log(Quantity supplied)	Log(Max. distance the plant's port)	Log(Total time spent at sea)
	(1)	(2)	(3)
I[Relational × supplies relational firm]	−0.009	0.026	0.002
	(0.020)	(0.022)	(0.015)
Date FEs	Yes	Yes	Yes
Supplier × Plant FEs	Yes	Yes	Yes
N	315,442	137,278	159,724

Notes: One observation is a boat during a fishing trip. *Quantity supplied* is the amount of fish the boat delivers to the plant per trip. *Max. distance from the plant's port* is maximum distance a specific boat is observed away from port. *Max. distance from the plant's port* can only be measured if the boat leaves from and arrives at the same port. *Total time at sea* is the amount of time the boat is away from port per trip. The number of observations varies from one column to the next as GPS variables for a given trip are sometimes missing. We define I[Relational ×supplies relational firm] to be equal to one if the supplier is (i) currently under a relational contract (ii) currently delivering to the firm it is under a relational contract with. In Panel A, we define an independent boat as being under a relational contract if the boat delivers more than 80% of its offloads (75th percentile) to the same fishmeal firm for 2 consecutive fishing seasons. In Panel B, we define an independent boat as being under a relational contract if the boat interacts more than 10 times (25th percentile) with the same firm during a fishing season and so, for at least 3 consecutive fishing seasons. Because use Boat × Plant FEs, I[Relational ×supplies relational firm] is identified from boats moving in and out of a relational contract. Standard errors clustered at the boat level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.9: Output quality and share of inputs from vertically integrated suppliers and suppliers under a relational outsourcing contract

Panel A: First definition of relational contracts				
Dep. var:	High Quality share of prod.		Protein content	
	(1)	(2)	(3)	(4)
Share of inputs from VI suppliers	0.343**	0.372**	0.950**	1.039***
	(0.153)	(0.161)	(0.348)	(0.333)
Share of inputs from relational suppliers	−0.100	−0.009	−0.159	−0.003
	(0.093)	(0.065)	(0.520)	(0.430)
High technology share of capacity	No	Yes	No	Yes
Season FEs	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.75	0.75	65.7	65.7
N	275	275	208	208

Panel B: Second definition of relational contracts				
Dep. var:	High Quality share of prod.		Protein content	
	(1)	(2)	(3)	(4)
Share of inputs from VI suppliers	0.395**	0.395**	0.949***	0.997***
	(0.164)	(0.162)	(0.265)	(0.250)
Share of inputs from relational suppliers	−0.424***	−0.481***	0.809	1.076
	(0.141)	(0.131)	(2.015)	(1.812)
High technology share of capacity	No	Yes	No	Yes
Season FEs	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.75	0.75	65.7	65.7
N	275	275	208	208

Notes: One observation is a firm during a production season. *Protein content* is the quantity weighted average of a measure of quality inferred from a database that provides weekly prices by quality. *Log(unit price)* is the log of the quantity weighted average unit price of exports during a season. *Share of inputs from VI suppliers* is the share of a firm's (or plant's) inputs that come from VI suppliers during a season. Steel boats tend to be bigger, better suited for industrial fishing, and are subject to different regulations. High capacity boats are boats whose hold capacity is in the upper quartile of the distribution. Boats without integrated cooling system use ice to keep fish fresh. *High technology share of capacity* controls for the share of the firm's total processing capacity (measured in metric tons per hour and averaged across all active plants within the firm) that uses steam drying technology. Standard errors clustered at the firm level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.10: Supplier Behavior, Vertical Integration and Output Quality

Dep. var:	Log(Quantity supplied)	Log(Max. distance the plant's port)	Log(Total time spent at sea)
	(1)	(2)	(3)
I[VI × supplies owner firm]	-0.133^{***}	0.017	-0.013
×I[Plant producing low quality]	(0.043)	(0.047)	(0.031)
I[VI × supplies owner firm]	-0.066^{**}	-0.067^{***}	-0.042^{**}
×I[Plant producing high quality]	(0.029)	(0.026)	(0.019)
Date FEs	Yes	Yes	Yes
Supplier × Plant × High Quality FEs	Yes	Yes	Yes
N	314,383	136,538	158,918
p-val - Test: two coefficients equal	0.00	0.03	0.04

Notes: One observation is a supplier during a fishing trip. This table is similar to Table 2.5, but with I[VI × supplies owner firm] interacted with the quality produced by the downstream plant. *Quantity supplied* is the amount of fish the boat delivers to the plant per trip. *Max. distance from the plant's port* is maximum distance a specific boat is observed away from port. *Max. distance from the plant's port* can only be measured if the boat leaves from and arrives at the same port. *Total time at sea* is the amount of time the boat is away from port per trip. I[Plant producing high quality] is a dummy equal to one if the plant the supplier delivers to produces only high quality fishmeal. The number of observations varies from one column to the next as GPS variables for on given trip are sometimes missing. Standard errors clustered at the boat level are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3 How Demand for Output Quality Affects Vertical Integration

In this section, we develop an alternative strategy to the one presented in section 2.5 to show that firms choose to integrate their suppliers when they face increased demand for high quality fishmeal. To do so, we develop an IV strategy that exploits quality-differentiated firm-specific demand shocks. We find that these shocks cause firms to increase their Share VI.

B.3.1 Empirical strategy

The logic behind our instruments for the quality grade of a firm's exports at a given point in time relies on two important facts about the Peruvian fishmeal sector. First, there is an exceptionally tight link between quality grade and export destination. This is apparent in the export transactions data, where some destination countries (e.g. Chile and Japan) consistently buy higher unit price and protein content fishmeal than other countries.[11] Sales records provided by a large firm drive home this connection. Country names are frequently used as a shorthand to represent different qualities—the quality column for exports is often simply filled in with the name of a country (e.g. "Thailand quality"). An increase in demand from high quality importers should thus increase the quality content of Peruvian fishmeal exports.

The second important fact about the Peruvian fishmeal sector is that the timing of sales contracts relative to production is typically such that a firm can integrate or sell suppliers in a given production season in response to high or low demand from particular importer countries. An industry association informed us that almost all contracts for a given season's production are negotiated either before the season starts, or early in

[11]See Appendix Table B.1 for a list of the main importers of Peruvian fishmeal and the average quality imported. Note that, as for humans, quantity and quality of feed (the latter here defined by protein content) are highly imperfect substitutes for the animals that consume fishmeal.

the season.

In the second stage, we estimate how acquisitions/sales of suppliers and firms' input mix respond to the quality grade produced:

$$\text{VI}_{it} = \alpha + \beta_1 \text{Quality}_{it} + \gamma_i + \delta_t + \varepsilon_{it} \tag{B.9}$$

We control for firm and production season fixed effects and cluster the standard errors at firm level as in Section 2.7.

To construct our demand shocks, we follow an approach similar to Bastos et al. (2018) (see also Park et al. (2009); Brambilla et al. (2012)). In the first stage, quality grade produced is instrumented by demand shocks from specific destinations as follows:

$$\text{Quality}_{it} = \gamma_i + \delta_t + \sum_j \beta_j \, (\text{I}^j_{i,\bar{t}} \, S^j_{-i,t}) + \varepsilon_{it} \tag{B.10}$$

where j is an export destination country, and $\text{I}^j_{i,\bar{t}} \, S^j_{-i,t}$ are our excluded instruments. $\text{I}^j_{i,\bar{t}}$ is a dummy variable equal to one if firm i exports to destination j at least once during our analysis period. $S^j_{-i,t}$ is the leave-firm-out share of Peru's fishmeal exports going to country j in season t, a proxy for the relative demand for firm i coming from destination j at a given point in time. Changes in j's demand should matter more for firms that previously exported to j, which we capture in the interaction between $S^j_{-i,t}$ and $\text{I}^j_{i,\bar{t}}$. A high β_j should represent a high quality importer country. We present the results of this specification in columns 3 and 4 of Appendix Table B.11).

In an alternative approach (presented in columns 5 and 5 of Appendix Table B.11), we replace $S^j_{-i,t}$ by $Quality^j_{-i,t}$, the leave-firm-out average quality of Peru's fishmeal exports going to country j in season t. In that case, a positive and high β_j represents a high willingness of Peruvian firms to respond to the higher demand for quality expressed by other countries. Conversely, a negative β_j represents a substitution effect: if an importer country starts buying higher quality output from other exporters, that same country would start buying low quality from other firms.

B.3.2 Results

We find that firms respond to positive shocks to demand for high quality fishmeal by sourcing a higher share of their inputs from suppliers that have been integrated.[12]

The OLS and the second stage IV results are reported in Appendix Table B.11. The estimates indicate that a one percentage point increase in the average protein content demanded—about 20 percent of approx. 63-68 percent range observed in Peru—induces the firm to source between 4 and 6 percent more of its inputs from integrated suppliers. As presented in Table 2.2, most of the variation in the share of inputs sourced from integrated suppliers comes from acquiring and selling suppliers.

Our interpretation of the results in Appendix Table B.11 is that firms vertically integrate *in order to be able to produce high quality output*. A potential alternative is that the liquidity that comes along with greater demand (rather than the demand for quality itself) may affect firms' ability to integrate. That is, if firms' seasonal revenues are expected to be higher when relative demand for quality is high, they may be better able to access the capital necessary to vertically integrate, but actually integrate for other reasons than to satisfy the demand for high quality. We address the concern by including controls for total seasonal sales. This has little effect on the estimated coefficients.

In the first stage, we use the 20 countries that import the most fishmeal from Peru (see Appendix Table B.1). Since China represents about 50 percent of total exports, we split China into 4 sub-countries, and we do so by using the destination port within China of each specific shipment. Our results are very similar if instead we use the 10 biggest importer countries or we use LASSO regressions to choose the importer countries whose demand fluctuations most affect quality grade exported.[13]

[12]The IV coefficients in columns 3 to 6 are bigger than the OLS coefficients in columns 1 and 2. We believe this is in part to be expected because the relationship between output quality and vertical integration *at firm level* estimated in Table 2.3 partly reflects a causal effect of organizational structure on output quality and partly other mechanisms. If the OLS estimates in that table are biased upwards, we would expect the OLS estimates here to be biased downwards, as we study the inverse relationship.

[13]LASSO (least absolute shrinkage and selection operator) is a regression analysis method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability

Since the existing literature that uses destination country demand shocks for identification often struggles with weak instruments, we compute the Kleibergen-Paap and Anderson-Rubin Wald test statistics. Comparing the statistics reported in Table 2.4 to the Stock-Yogo critical values[14], while we do not pass the Kleibergen-Paap under-identification test, we reject the null hypothesis that our instruments are weak (as the F-statistic surpasses the 10 percent critical value). We also reject the hypothesis that the coefficients on the excluded instruments are jointly zero when they are included in place of quality itself in the second stage regression using the Anderson-Rubin Wald test. It is additionally important to note that weak instruments would bias the IV coefficients *downward*, i.e., towards the OLS coefficients, rather than upward. See Bastos et al. (2018) for a lengthier discussion of this issue in the context of "demand pull" instruments.

The strategic changes in organizational structure in response to changes in the composition of demand are consistent with the integration$\rightarrow$quality relationship shown in Section 2.7 and confirm the results shown in section 2.5 that firms integrate when they face incentives to quality upgrade. We conclude that Peruvian manufacturing firms are aware of, and act on, their greater ability to produce high quality grade output when their suppliers have been integrated.

---

of the statistical model it produces, penalizing the model for including more regressors. LASSO selects eight importer countries.

[14]Though Stock-Yogo's critical values are computed for the homoskedastic case, it is standard practice to compare the Kleibergen-Paap Wald test statistics to these critical values even when one reports standard errors that allow for heteroskedasticity.

Table B.11: Vertically Integrated Share of Inputs and Output Quality - Instrumenting with Firm-specific Demand Shocks

Dep. var:	Share of inputs from VI suppliers					
	OLS	OLS	IV	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Protein content	0.028*	0.031**	0.060*	0.061*	0.039*	0.042*
	(0.014)	(0.013)	(0.034)	(0.033)	(0.023)	(0.024)
Log(Sales)		−0.013		−0.031		−0.020
		(0.014)		(0.019)		(0.016)
Season FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.30	0.30	0.30	0.30	0.30	0.30
N	192	192	192	192	192	192
Kleibergen-Paap LM p-value (Under-id)			0.53	0.46	0.53	0.45
Kleibergen-Paap Wald F statistic (Weak inst)			527.3	504.0	962.6	671.75
Anderson-Rubin Wald test p-value			0.00	0.00	0.00	0.00

Notes: One observation is a firm during a production season. *Share of inputs from VI suppliers* is the share of a firm's inputs that come from VI suppliers during a season. Protein content is the quantity weighted average of a measure of quality inferred with a database that provides weekly prices by quality. The instruments are interactions of indicators for at least one export in our analysis period to each of the top 20 destination countries with leave-firm-out share of Peru's fishmeal exports towards the destination in the relevant season (columns 3 and 4) or leave-firm-out average protein content exported towards the destinatiion in the relevant season (columns 5 and 6). Standard errors clustered at the firm level are included in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix C

Appendix to Chapter 3

C.1 Additional Tables and Figures

Table C.1: HS-Chapter Import Share by Connectedness Status

	Not Connected	Connected	Government	Total
Animal/Animal Products	1.1	0.4	0.0	0.9
Vegetable Products	5.8	8.5	0.2	5.1
Foodstuffs	5.8	1.9	0.2	4.4
Mineral Products	16.4	41.0	12.0	18.5
Chemicals/Allied Industries	8.0	4.1	1.9	6.5
Plastics/Rubbers	5.8	1.5	0.6	4.4
Raw Hides, Skins, Leather	0.3	0.3	0.0	0.2
Wood/Wood Products	1.7	1.7	0.9	1.5
Textiles	5.7	1.4	0.2	4.3
Footwear/Headgear	0.4	0.3	0.0	0.3
Stone / Glass	1.6	1.5	0.2	1.3
Metals	11.0	9.2	15.8	11.6
Machinery / Electrical	14.9	13.4	30.3	17.4
Transportation	18.4	8.1	27.7	18.8
Miscellaneous	3.1	6.7	9.9	4.7
Total	100.0	100.0	100.0	100.0

Notes: This table reports share of HS-chapter level imports for connected and not connected firms and the aggregate import share. Aggregate shares are not the sum of individual shares but a value-weighted average.

Table C.2: Robustness check - Import license liberalization and connected sectors

Dep. var:	Standardized growth rate			
	Connected sectors		Non connected sectors	
	(1)	(2)	(3)	(4)
$1\{\text{Lib}\} \times 1\{\text{Post}\}$	-0.157^{***}	-0.164	0.336^{***}	0.546^{***}
	(0.060)	(0.115)	(0.044)	(0.106)
Year FEs	Yes	No	Yes	No
HS2 X Year FEs	No	Yes	No	Yes
HS6 FEs	Yes	Yes	Yes	Yes
N	1188	1188	1174	1174

Notes: One observation is a 6 digit product code during fiscal year. 1{Lib} is a dummy equal to one if the product is in the group of products which are *de facto* liberalized. The sample only includes the set od products which are *de facto* and *de jure* liberalized or not liberalized. 1{Post} is a dummy equal to one if the fiscal year considered is greater than 2013, the first fiscal year when the liberalization was implemented. HS2 X Year FEs are 2 digit product codes - fiscal years pairs of fixed effects. Sectors are defined to be connected if the market share of connected firms in that sector in the pre-reform period is higher than 30 percent. Standard errors are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.3: Shipment size and Number of transactions

Dep. var:	Log(av. shipment) size)			Number shipments		
	All sectors	Conn. sectors	Non conn. sectors	All sectors	Conn. sectors	Non conn. sectors
	(1)	(2)	(3)	(4)	(5)	(6)
1{Lib} $\times$ 1{Post}	-0.032^{**}	-0.163^{***}	-0.021	7.953^{***}	6.328^{***}	8.831^{***}
	(0.016)	(0.048)	(0.017)	(1.793)	(1.270)	(2.327)
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Company X HS6 FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	156,246	25,351	130,895	156,294	25,362	130,932

Notes: On observation is a 6 digit product code on a given fiscal year. This table presents the result of a standard Difference in Difference model to estimate the effect of import liberalization on the size of shipments being made and the number of shipments made per year. Variables are defined as in Table 3.2. Standard errors are included in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$