THREE ESSAYS ON GLOBAL SUPPLY CHAINS

by

Mengxiao Liu

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Department of Economics
University of Toronto

© Copyright 2017 by Mengxiao Liu
Abstract

Three Essays on Global Supply Chains

Mengxiao Liu
Doctor of Philosophy
Graduate Department of Department of Economics
University of Toronto
2017

Firm boundary decision – the decision of firms to integrate or outsource to one another, have important implications on the distribution of value added along global supply chains. My dissertation contributes to the booming literature on global supply chains in three aspects.

Chapter 1 develops a theoretical framework that incorporates choice of scope into a firm’s boundary decision. A firm with a larger scope can produce a product with higher functionality, but also faces a more severe holdup problem. Outsourcing reduces the holdup on the firm’s side, while integration reduces the holdup on its suppliers’ side. In industries where consumers value functionality, high-productivity firms maintain large, integrated production networks. In industries where consumers value cheaper price, high-productivity firms maintain small, outsourced production networks.

In Chapter 2, I compile an empirical database that shows (for the first time) the coexistence of backward and forward integration in global supply chains. This observation contradicts the prevailing assumption that integration is unilateral. To explain this coexistence, I modify the classic model from Antràs and Helpman (2004) to allow for bilateral integration. I show that my model explains the empirical coexistence of backward and forward integration, and that it performs much better than all existing models.

In Chapter 3, I explore the impact of institutional quality on firms’ boundary decisions. A country with better legal institutions provides better protection to firms by better securing their investments in the event of a legal appeal. Effectively, judicial quality acts as a substitute to integration: with better legal protection, firms are less likely to resort to integration as means to mitigate contractual frictions. I provide theoretical and empirical evidence for this hypothesis.
To my parents, my husband, and huanhuan.
Acknowledgements

I am indebted to my supervisor Daniel Trefler for his invaluable guidance and supervision for the past few years. I am also immensely grateful to my committee members Peter Morrow and Kunal Dasgupta for their insightful comments and suggestions. I would like to extend my special thanks to Professor Davin Chor for his thorough and insightful appraisal as my external committee member. I thank the following people for their helpful inputs and encouragements at various stages of this dissertation: Joanne Oxley, Nicholas Li, Andrey Stoyanov, Heski Bar-Issac, Loren Brandt, Robert McMillan, Xiaodong Zhu, Chaoran Chen, Arthur Blouin, Laura Turner, Peter Cziraki, Leandro Freylejer, Garth Frazer, Ignatius Horstmann, Bernardo Blum, Gustavo Bobonis, Gilles Duranton. I would also like to thank the seminar participants at Toronto, Syracuse, U Arkansas at Fayetteville, UC Merced, U Florida, along with conference participants at the Canadian Economics Association, the Midwest Economics Theory and International Trade Meetings, the International Atlantic Economic Society, and the Rimini Conference in Economics and Finance, for many helpful comments.
# Contents

1 Scope and Quality in Production Networks .......................... 1
   1.1 Introduction ................................................. 2
   1.2 Setup ....................................................... 4
      1.2.1 Preferences and Production .......................... 4
      1.2.2 Timing .................................................. 6
      1.2.3 Hold-up ................................................ 6
   1.3 Equilibrium .................................................. 7
      1.3.1 SSPE ..................................................... 7
      1.3.2 Optimal Choice of Idea Inputs \((h_j, m_j)\) ................. 8
      1.3.3 Optimal Choice of Organization Size \(N\) and Form \(k \in \{O, V\}\) ............... 9
   1.4 Two Types of Industries ....................................... 10
   1.5 Ideas-Oriented Industry ....................................... 11
   1.6 Cost-Oriented Industries ..................................... 13
   1.7 Conclusion .................................................... 14

Appendix 1.A Firm’s Shapley values under various functional forms ............ 17
   1.A.1 Leontief production function .......................... 17
   1.A.2 CES function .............................................. 17
   1.A.3 O-ring function .......................................... 19

Appendix 1.B Proof of the Existence of a SSPE .......................... 20

Appendix 1.C Firm and suppliers’ levels of investments ....................... 21

Appendix 1.D Marginal firms in the ideas-oriented industry .................. 21
   1.D.1 Strict concavity of the profit function .................... 22
   1.D.2 The marginal firm’s organizational behavior ................ 22
   1.D.3 Uniqueness of the threshold productivity level ............. 31
1.D.4 Monotonicity of the threshold productivity .......................... 32
1.D.5 Existence of threshold productivity and threshold industries .......... 33

Appendix 1.E Firms’ scope decisions in the ideas-oriented industry .......... 35
1.E.1 Firm’s scope decision in (fp2) is limited ............................... 36
1.E.2 Firm’s scope decision is limited ...................................... 36

Appendix 1.F Proofs for the cost-oriented industry .......................... 37

2 Global Supply Chains without Principals 38

2.1 Introduction ................................................................. 40
2.2 Data and Descriptive Statistics ......................................... 44
  2.2.1 Data ................................................................. 44
  2.2.2 Pros and Cons of this Database ..................................... 46
  2.2.3 R&D and Capital Intensity ......................................... 47
  2.2.4 Descriptive Statistics ............................................... 48
2.3 Model .......................................................... 49
  2.3.1 The Homogeneous Model ............................................. 50
  2.3.2 The Heterogeneous Model ......................................... 59
  2.3.3 Comparison to a Principal-Agent Framework ....................... 62
2.4 Empirical Results for the Homogeneous Model ......................... 63
  2.4.1 Multinomial Logit Model .......................................... 63
  2.4.2 Linear Probability Model ........................................... 66
  2.4.3 Robustness: Industry-level R&D and Capital Intensity ............ 70
  2.4.4 Robustness: Regression with Only Imputed Relationships ........ 72
2.5 Empirical Results for the Heterogeneous Model ....................... 74
  2.5.1 Empirical Specification ............................................ 74
  2.5.2 Regression Results for Pair-level R&D Intensities ................ 75
  2.5.3 Regression Results for Pair-level R&D Intensities ............... 77
2.6 Conclusion ................................................................. 77

Appendix 2.A Imputing Seller-Buyer Relationships ......................... 79
Appendix 2.B Proof of Theorem 5 ........................................... 79
Appendix 2.C Properties of $\mu^k$ ........................................... 80
Appendix 2.D Proof of Theorem 5 ........................................... 81
3 Judicial Quality, Relative Efficiency, and Firm Boundary Decisions

3.1 Introduction .............................................................. 93
3.2 Model ................................................................. 96
  3.2.1 Setup ............................................................. 96
  3.2.2 Industry Equilibrium ............................................ 98
3.3 Empirical Results ...................................................... 107
  3.3.1 Organizational Form \( k^*_i \) and Relative Efficiency \( \tilde{\theta}_i \) ........................................ 107
  3.3.2 Organizational Form \( k^*_i \) and Judicial Quality \( \delta_b, \delta_s \) ....................................... 110
3.4 Conclusion ............................................................ 115

Appendix 3.A Proof of Lemma 5 ........................................... 117
Appendix 3.B Proof of Theorem 7 ......................................... 118
Appendix 3.C Proof of Theorem 8 ........................................ 120
Appendix 3.D The Distribution of \( m^*_b \) ............................... 121

Bibliography ................................................................. 121
List of Tables

1.1 Relationship between $(N^O, N^V)$ and $(n^O, n^V)$ .................................................. 23

2.1 Effects of a 10% Increase in Relative Buyer’s R&D Intensity in Different Types of PRT Models ............................................................... 43

2.2 Composition of the Different Types of Relationships .............................................. 46

2.3 Firm Characteristics by Organizational Forms ......................................................... 49

2.4 Buyer and Seller’s Revenue Shares under Different Ownership Structures .... 54

2.5 Multinomial Logistic Regression Results ............................................................... 65

2.6 Average Marginal Effects of the Multinomial Logit Model .................................. 66

2.7 Linear Probability Model – Pair-level R&D/Capital Intensities .............................. 68

2.8 Linear Probability Model - Industry-level R&D/Capital Intensities ...................... 71

2.9 Robustness Check: Regressions using Only Imputed Relationships .................... 73

2.10 Regression Results with Tertiles of Pair-specific $RD$ and $(1 – RD)$ ............... 76

2.11 Regression Results with Tertiles of industry-specific $RD$ and $(1 – RD)$ .......... 77

2.12 Logit Regression Results ...................................................................................... 85

2.13 Probit Regression Results ..................................................................................... 86

2.14 Regression Results with Quintiles of Pair-specific $RD$ and $(1 – RD)$ ............ 88

2.15 Regression Results with 7 Quantiles of Pair-specific $RD$ and $(1 – RD)$ ........ 89

2.16 Regression Results with Quintiles of industry-specific $RD$ and $(1 – RD)$ ....... 89

2.17 Regression Results with 7 Quantiles of industry-specific $RD$ and $(1 – RD)$ .... 90

3.1 Buyer and Seller’s Revenue Shares under Different Ownership Structures .... 99

3.2 Marginal Effects from Column (4), Table 3.3 ...................................................... 109

3.3 Multinomial Logit Regression Results .................................................................. 110

3.4 Linear Regression Results ..................................................................................... 111
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5 Linear Regression Results</td>
<td>113</td>
</tr>
<tr>
<td>3.6 Multinomial Logit Regression Results</td>
<td>114</td>
</tr>
<tr>
<td>3.7 Marginal Effects from Column (4), Table 3.6</td>
<td>115</td>
</tr>
</tbody>
</table>
# List of Figures

1.1 Simulation of $\delta(N)$ .................................................. 23  
1.2 $N^V < n^O$ .......................................................... 24  
1.3 $N^O > n^V$ .......................................................... 25  
1.4 $N^O \leq N' < N^V$ .................................................. 26  
1.5 $N' < N^O < N^V$ .................................................. 28  
1.6 $N^O < n^O, N^V > n^V$ ....................................... 29  
1.7 $N^O < N' \leq N^V$ .................................................. 30  
1.8 $N^O < N^V < N'$ .................................................. 31  
1.9 $n^O < N^O < N^V < n^V$ ....................................... 32  
1.10 Profit functions in ideas-oriented industry ......................... 33  
1.11 Outsourcing Industry ............................................. 35  
1.12 Integration Industry ............................................. 36  

2.1 Histograms of Industry Pairs by Their Shares of Buyer Integration .... 42  
2.2 Ownership Decisions with Pair-specific Productivity ................... 60  
2.3 The Optimal Revenue Share as a Function of $\eta$ ......................... 80  

3.1 Relationship between Relative Efficiency and Organization Decision (and the Effect of an Increase in Relative Judicial Quality) .................. 94  
3.2 Distribution of Organizational Forms across Country Pairs ............... 95  
3.3 Optimal Revenue Share ($\beta_i^*$) as a function of relative efficiency ($\hat{\theta}_i$) .................. 103  
3.4 Optimal Revenue Share ($\beta_i^*$) as a function of relative efficiency ($\hat{\theta}_i$) ............... 112
Chapter 1

Scope and Quality in Production Networks
Abstract
This paper develops a framework for firms’ quality, scope and boundary decisions. In this framework, the firm faces a trade-off between product functionality and control over its production network. To produce a high-functionality product, the firm needs to maintain a complex production network that involves many suppliers. In a production network involving many firms, contractual incompleteness distorts each firm’s investment incentive, leading to lower product quality. A firm uses make-or-buy decision as a tool for mitigating incentive distortions. Integration provides a stronger incentive for the firm and weaker incentives for its suppliers, while outsourcing works in the opposite direction. In equilibrium, a firm’s organizational behavior depends on both firm and industry characteristics. In industries where consumers value functionality, high-productivity firms maintain large, integrated production networks. In industries where consumers value cost-efficiency, high-productivity firms maintain small, outsourced production networks.

1.1 Introduction
Recent empirical and theoretical work has documented the fact that firms organize their production networks differently.¹ Firms’ production networks vary in at least three dimensions: the quality of the product produced, the number of suppliers engaged in the network, and firms’ make-or-buy decisions. This paper examines the nexus between these three dimensions. It tries to understand why we see such variation across industries as well as within industries. We begin with the observation that a consumer’s valuation of products depends on two dimensions of quality. First, it depends on what we refer to as ‘functionality’. The archetypical example is a smartphone such as Apple’s iPhone 7, which embodies dozens of functions including phone, wifi, camera, etc. Less archetypically, but more common empirically, functionality deals with ‘product lines’: a firm produces many sizes of TVs and many types of refrigerators. The second

¹For example, Bernard et al. (2007) documents the extensive margin in exporting firms’ product scope heterogeneity, which is addressed in the theoretical work by Nocke and Yeaple (2006), Bernard et al. (2010a), and Eckel and Neary (2010). To the best of our knowledge, this paper is the first to incorporate firm-boundary decisions into the context of multi-product firms, although the interpretation of scope in our paper is not necessarily limited to product scope, as will be elaborated later.
dimension of consumer’s valuation is the quality of each function/product, here understood in a more conventional sense of reliability or better design. We model the firm’s choices of functions or, alternatively, the size of the product line as a blueprint for exactly what the function will look like. We assume that each function is developed in conjunction with a single supplier and that the development of a blueprint requires relationship-specific, non-contractible investments from both the firm and its supplier. The larger the investments, the higher the quality of the function.

We identify two determinants of functionality and quality of individual functions. First, the greater the functionality, i.e., the larger the number of suppliers—the greater the hold-up problem. This is modeled by assuming that the firm and all \( N \) suppliers engage in multilateral bargaining so that, roughly speaking, each receives a share \( 1/(N + 1) \) of revenues. This creates a tension: products with many functions and hence many suppliers generate more revenue, but also create a larger hold-up problem. For reasons familiar from the property rights theory of the firm, larger firms will vertically integrate their suppliers to overcome the hold-up problem.

The second determinant deals with a managerial tension: some managers are good at cutting costs, while others are good at identifying functions that are valued by consumers and building the supplier networks needed to deliver these functions. Specifically, in “ideas-oriented” industries more productive firms are firms with a high marginal return to more functions (productivity and functions are complements in the profit function). In these industries, more productive firms maintain larger, more integrated production networks. In “cost-oriented” industries more productive firms are firms with a higher marginal return to “leaner and meaner” production networks (Eckel and Neary, 2010). In these industries, more productive firms maintain smaller, more outsourced networks.

This paper springs from the extensive literature on property rights theory and firm-boundary decisions. It is most closely related to the line of papers by Antràs (2003), Antràs and Helpman (2004), and Acemoglu et al. (2007), but differs from them in important ways. Unlike Antràs (2003) and Antràs and Helpman (2004), heterogeneity in firms’ organizational choice does not depend on the assumption that the fixed cost of production is higher under vertical integration.\(^2\) In fact, there is no fixed cost of production in this paper. The trade-off between

\(^2\)See the seminal work by Grossman and Hart (1986) and Hart and Moore (1990).
\(^3\)See Nunn and Trefler (2008), Helpman (2011), and Antràs (2013, 2015) for reviews of this literature.
\(^4\)In Antràs (2003) and Antràs and Helpman (2004), vertical integration brings a higher variable profit relative to outsourcing, but also entails a higher fixed cost of production. The difference between the variable profits under vertical integration and under outsourcing is increasing in productivity. High-productivity firms choose
vertical integration and outsourcing stems from the tension between scope and incentive: a firm might want to have a large supplier network in order to generate a higher demand,\footnote{This could either mean that the firm produces a product with higher functionality, or that the firm has a wider product line.} but this exacerbates the hold-up problem. Acemoglu et al. (2007) also models the firm’s scope decision, but the authors assume that the firm does not make any non-contractible, relationship-specific investments, and thus always chooses outsourcing to incentivize its suppliers.

This paper also contributes to the booming literature on multi-product firms and international trade because an alternative interpretation for the scope of the supplier network in this paper is the scope of product. Theoretical papers on multi-product firms include Ottaviano and Thisse (1999), Allanson and Montagna (2005), Nocke and Yeaple (2006), Feenstra and Ma (2007), Eckel and Neary (2010), Dhingra (2013), and Mayer et al. (2014). See empirical papers by Bernard et al. (2007, 2009b, 2010b), Liu (2010), Liu and Rosell (2013), and Lopresti (2016). Specifically, in Bernard et al. (2010a), the authors observe that “some pairs of products are more likely to be coproduced within firms than others”. To the best of our knowledge, such an empirical pattern has not been theoretically modeled. Our paper provides a framework for analyzing such observations.

\section{Setup}

\subsection{Preferences and Production}

Consider a closed economy with a continuum of sectors. Each sector produces one final product. The representative consumer’s preference is:

\[ U = \left\{ \int_{\omega} \left[ \varphi(\omega)^{\nu} y(\omega) \right]^{(\sigma - 1)/\sigma} d\omega \right\}^{\sigma/(\sigma - 1)}, \]

where \( \omega \) is a product index, \( y(\omega) \) is a consumption level, \( \varphi(\omega)^{\nu} \) is a demand shifter (\( \nu \) is a parameter and \( \varphi \) is explained in detail below), and \( \sigma \) is the elasticity of substitution. We assume \( \nu(\sigma - 1) > 0 \) and \( \sigma > 1 \).

Production of a variety has three stages. The firm first decides on a level of functionality \( N \), that is, on the number of functions the product will have or the number of products in vertical integration because the higher variable profit under vertical integration compensates for the higher fixed cost. Without the assumption that the fixed cost of production is higher under vertical integration (relative to outsourcing), all firms would choose integration.
the product class. For example, an iPhone 4 has many functions (wifi, voice recognition, apps etc.) and a Mercedes has many products (sports car, sedan etc.). Second, the firm identifies \(N\) suppliers, each of which will help the firm develop one of the functions. This blueprint or ‘ideas’ stage involves non-contractable inputs from both the firm and the supplier. Third, in the ‘production’ stage the final good is produced in a complete-contracting environment. The ideas stage is the key stage and we discuss it in detail next.

In the ideas stage each function is developed using the shared inputs of the firm and the supplier. For simplicity, we assume that each function is developed by the firm with the help of a single supplier.\(^6\) A function can be of variable quality. For example, voice recognition is better in some cell phones than in others and compressors are better in some refrigerators than in others. Let \(q_j\) be the quality of function \(j = 1, \ldots, N\). It depends on the firm’s input \(h_j\) and the seller’s input \(m_j\):

\[
q_j = h_j^\eta m_j^{1-\eta}/\hat{\eta}
\]

where \(\hat{\eta} \equiv \eta^n(1 - \eta)^{1-\eta}\). Quality \(q_j\) and inputs \((h_j, m_j)\) are non-contractible.

Consumer valuation of functionality and function quality are captured by the demand shifter

\[
\varphi = D(N, \theta) \min\{q_1, q_2, ..., q_N\}
\]  

where \(\theta \in [0, 1]\) is a firm index that replaces \(\omega\); it plays no role yet, but we will later interpret it as the firm’s productivity as in Melitz (2003).

The particular functional form in equation (1.1) is not all that important to our argument. We obtain similar results with either a production function that is CES in functions (of which equation (1.1) is the special case of perfect complements) or with O-Ring function.\(^7\) What is very important is that the buyer and all suppliers are essential in a Shapley-value sense. That is, \(\varphi = 0\) if any player is not part of the team. This will ensure that the buyer’s Shapley value is decreasing in the size of the team \((N)\). Restated, essentiality rather than the functional form of equation (1.1) is what provides our key modelling assumption, namely, that more functionality comes at the cost of greater hold-up.\(^8\)

\(^6\)It is possible to allow for multiple suppliers of a single function.

\(^7\)For CES, \(\left\{\sum_{j=1}^N q_j^\beta\right\}^{1/\beta}\) which, under the symmetry that we impose below, becomes \(N^{1/\beta}q\). For O-Ring, \(B(N)\prod_{j=1}^N q_j\), which when symmetry is imposed and logs taken becomes \(\ln B(N) + N \ln q\). These different specifications affect the functional form of the optimal inputs \((h_j, m_j)\), but otherwise do not matter.

\(^8\)In contrast, AAH does not assume that players are essential. In their setup, the Shapley value is independent of \(N\).
The marginal cost of input \( j \in \{ h, m \} \) is \( C_j(N, \theta) \). For simplicity, we assume that \( C_j(N, \theta) = w_j C(N, \theta) \), where the constant \( w_j \) captures the price of inputs and other things that are log-separable from \( N \) and \( \theta \). Note that both \( D \) and \( C \) depend on \( \theta \). Not surprisingly, we will find (roughly) that only \( D/C \) matters. This is the usual point that demand shifters and productivity are isomorphic.

Demand for the final product \( y \) is

\[
y = A \varphi^\alpha p^{-\sigma} \quad \text{where} \quad \alpha \equiv \nu(\sigma - 1) \in (0, 1).
\]

The firm is a monopolistic competitor and sets price equal to \( [\sigma/(\sigma - 1)] c \). This generates revenues

\[
R = \hat{A} \varphi^\alpha = \hat{A} [D(N, \theta) \min\{q_1, q_2, ..., q_N\}]^\alpha \tag{1.2}
\]

where \( \hat{A} \equiv \sigma - \sigma[(\sigma - 1)/\alpha]^{\sigma-1} A \).

### 1.2.2 Timing

1. The firm and all the suppliers observe \( \theta \).
2. The firm chooses organizational form \( k \in \{ O, V \} \) (\( O \) is outsourcing and \( V \) is vertical integration), adopts technology \( N \), and offers contract \( \{ \tau_j \}_{j=1}^N \), where \( \tau_j \) is an upfront payment to supplier \( j \).
3. A continuum of potential suppliers apply for the contract and the firm chooses \( N \) suppliers from them.
4. The firm and the suppliers simultaneously choose their investment levels \( \{(h_j, m_j)\}_{j=1}^N \).
5. The firm and the suppliers bargain over the division of future revenue. At this stage, the firm and the suppliers can decide to withdraw their investments.
6. Ideas are generated (\( \varphi \) is determined). Output is produced and sold. Revenue is divided according to the bargaining agreement.

### 1.2.3 Hold-up

We assume that in the negotiation stage, if supplier \( j \) decides to withdraw from the production process, the firm cannot use the input as efficiently. We model this by assuming that the quality of the input drops from \( q_j \) to \( \Delta^k q_j \), where \( k \in \{ O, V \} \) and \( \Delta^O < \Delta^V \). On the other hand, if
the firm withdraws its investment for function $j$, $q_j$ drops to 0 regardless of the organizational form $k$.

## 1.3 Equilibrium

### 1.3.1 SSPE

We define the symmetric sub-game perfect equilibrium (henceforth SSPE) as a tuple $\{N, \tau, h, m\}$, where $N$ is the firm’s choice of functionality. In SSPE, $\tau$ is the firm’s upfront payment to every supplier, that is, $\tau_j = \tau$ for $j = 1, \ldots, N$. Similarly, $h$ is the firm’s investment for each function, and $m$ is each supplier’s investment. That is $(h_j, m_j) = (h, m)$, for $j = 1, \ldots, N$.

SSPE can be characterized by backward induction as in AAH. Since this is familiar (and notationally difficult) territory, we jump immediately to the revenue in any SSPE. This is given by

$$R = \hat{A} \{D(N, \theta) \hat{h}^N m^1 - \eta/\hat{\eta}\}^\alpha.$$

**Lemma 1.** In every SSPE, the firm’s Shapley value under organizational form $k \in \{O, V\}$ is $\gamma^k(N)R$, where

$$\gamma^k(N) = \frac{\delta^k N + 1}{N + 1},$$

and $\delta^k \equiv (\Delta^k)^\alpha$. Each seller’s Shapley value is $(1 - \gamma^k(N))R/N$.\(^9\)

In AAH, the firm’s share of revenue $\gamma^k$ is independent of $N$. Here, organizations with more suppliers face larger hold-up problems. This is reflected in the fact that $\gamma^k$ is decreasing in $N$. This has an important implication. If in our model $\gamma^k$ were independent of $N$, then the choice of number of suppliers and choice of organizational form would not interact. Specifically, the choice of organizational form would be determined as in Antràs (2003) or as in Antràs and Helpman (2004) with $f_V = f_O$, i.e., if $\eta$ is large then all firms integrate and if $\eta$ is small then all firms outsource. Here, a productive firm may want to have a large $N$ that will lead to a smaller share of revenue (a small $\gamma^k$); the firm may find it optimal to offset this loss of revenue by moving from the $O$ form to the $V$ form, which has the effect of increasing the firm’s revenue share from $\gamma^O$ to $\gamma^V$. **In essence, productive firms will want to vertically integrate to offset the endogenously greater hold-up problem that comes with having more suppliers.**

\(^9\)Appendix 1.B derives the SSPE using backward induction. It also proves the existence of SSPE.

\(^{10}\)Appendix 1.A provides a derivation of $\gamma^k(N)$ under several production functions.
1.3.2 Optimal Choice of Idea Inputs \((h_j, m_j)\)

The firm and the suppliers’ problems are familiar from Antràs (2003) and Antràs and Helpman (2004). They simultaneously choose their investment levels taking the others’ investment levels as given.

The firm’s problem can be written as:

\[
\max_{(h_1, h_2, \ldots, h_N)} \gamma^k(N) \frac{A}{\eta^\alpha} \left[ D(N, \theta) \min_{1 \leq j \leq N} \{h_j^{\eta}m_j^{1-\eta}\} \right]^{\alpha} - w_h C(N, \theta) \sum_{j=1}^{N} h_j \quad (FP1)
\]

A supplier’s problem can be written as:

\[
\max_{m_j} \frac{1 - \gamma^k(N)}{N} \frac{A}{\eta^\alpha} \left[ D(N, \theta) \min_{1 \leq j \leq N} \{h_j^{\eta}m_j^{1-\eta}\} \right]^{\alpha} - w_m C(N, \theta)m_j \quad (IC1)
\]

We assume that \(\alpha < 1\) so that the firm’s problem (FP1) and the supplier’s problem (IC1) are concave.

**Assumption 1.** \(0 < \alpha < 1\)

**Lemma 2.** In any SSPE, the unique solution \((h_j, m_j)\) to the firm’s problem is:

\[
m^k(N, \theta, \eta) = \kappa A \frac{1 - \eta}{w_m} \left\{ \left[ \gamma^k(N) \right]^{\alpha \eta} \left[ 1 - \gamma^k(N) \right]^{1+\alpha \eta} \right\}^{\frac{1}{1+\alpha - \sigma}} D(N, \theta)^\alpha \gamma^{\frac{1}{1+\alpha - \sigma}} \quad (1.3)
\]

\[
h^k(N, \theta, \eta) = \frac{\gamma^k(N)}{1 - \gamma^k(N)} \frac{w_m/(1 - \eta)}{w_h/\eta} m^k(N, \theta, \eta). \quad (1.4)
\]

where \(\kappa \equiv \left\{ \alpha (\sigma - 1)^{\frac{\sigma - 1}{\sigma - \sigma}} \right\}^{\frac{1}{1+\alpha - \sigma}}\).

These are messy expressions, but ones that are not fundamentally new. The only new insight comes from equations (1.3) and (1.4): \(h/m\) will vary within an industry not only because different firms choose different organizational forms \(k\), but also because they choose different-sized organizations which affects \(h^k/m^k\) via the effects of \(N\) on \(\gamma^k\). Thus, our framework offers a natural explanation of the enormous within-industry heterogeneity in relationship-specific investments that we see in the data.\(^{11}\)

\(^{11}\)There are two main (old) insights from equations (3) and (4). First and obviously, the optimal input levels are both less than the first-best (contractible) input levels. Second, \(h^k/m^k\) equals the first-best input ratio if and only if \(\gamma^k = 1/2\). This points to how the Grossman-Hart logic plays out in this model. When \(\eta\) is large so that the firm’s investment is most important, the firm wants to choose a form that will raise \(h^k/m^k\). This is the form with the larger \(\gamma^k\) and, since \(\gamma^V(N) > \gamma^O(N)\), vertical integration is preferred.
1.3.3 Optimal Choice of Organization Size $N$ and Form $k \in \{O, V\}$

Plugging in the Lemma 2 optimal inputs into the firm’s problem FP1, the firm’s problem simplifies to

$$\max_{k \in \{O, V\}, N \in [1, \infty)} \Pi^k(N, \theta, \eta) = \kappa AG(N, \theta) \Psi(\gamma^k(N), \eta),$$

where

$$G(N, \theta) \equiv \left[ \frac{D(N, \theta)}{NC(N, \theta)} \right]^{\frac{\alpha}{1-\alpha}},$$

and

$$\Psi(\gamma, \eta) \equiv \frac{1 - \alpha[\gamma \eta + (1 - \gamma)(1 - \eta)]}{[\gamma \eta (1 - \gamma)^{1-\eta}]^{\frac{\alpha}{1-\alpha}}}.$$

$k$ is a constant as defined in Lemma 2.

It is now apparent that only $G = D/(NC)$ matters, not $D$ or $NC$ separately.\footnote{Note that in the expressions for $h^k$ and $m^k$, what matters is $D^\alpha/NC$, so $D$ and $NC$ matter separately. However, they only matter for the levels of $h^k$, $m^k$ and hence for quality $q_j$. They do not matter separately for anything else whatsoever.} Note that up to this point we have not said anything about $\theta$. It is now clear that the appropriate assumption is that $G$ is increasing in $\theta$.

**Assumption 2.** $G(N, \theta)$ is strictly increasing in $\theta$.

This is a good spot to compare our model with that of Antràs and Helpman (2004), equation (10), where $N = 1$. Their model has an almost identical profit function: In our notation it is basically $\Pi^k(1, \theta, \eta) = \theta^{\sigma-1} \Psi(\gamma, \eta)$ where, as is standard in Melitz-like models, $G(1, \theta) = \theta^{\sigma-1}$. However, there are four differences to note.

1. $N \neq 1$ is a choice variable.
2. There are no fixed costs of organizations ($f_V$ and $f_O$ in their notation). Recall that in their model, when there are no fixed costs as is the case here (or even when there are fixed costs and $f_V = f_O$) then their model reduces to Antras (2003). That is, when $\eta$ is small all firms outsource and when $\eta$ is large all firms vertically integrate.
3. The most important difference is that $\Psi(\gamma^k(N), \eta)$ depends on $N$. In Antràs (or Antràs and Helpman with $f_V = f_O$), the firm chooses the organizational form $k$ that maximizes $\Psi(\gamma^k(1), \eta)$ where $\gamma^k(1)$ and $\eta$ are parameters. In our setting, there is an interaction between the choice of organization and the choice of functionality. The larger the organization ($N$), the smaller $\gamma^k$. This creates a tension: the firm might want to grow bigger in
order to have higher demand, but this exacerbates the hold-up problem. In short, we have heterogeneity of organizational forms without fixed costs because we have endogenized the extent of the hold-up problem.

This is also a good spot to compare our profit function to that in AAH. First, in their model only the supplier makes a relationship-specific investment ($\eta = 0$) so that the firm always outsources. Second, in their model the Shapley value is completely determined by exogenous parameters so that there is no trade-off between size and hold-up.

We now make assumptions that make it easier to solve for the optimal $N$. We will use first-order conditions and so ignore the integer constraint on $N$. The following assumption ensures that there is a unique optimal $N^k$ and that it is bounded away from 1 and infinity.

**Assumption 3.** $G(N, \theta)$ satisfies the following conditions:

1. $G(N, \theta)$ is strictly log-concave in $N$: $-\frac{\partial^2 \ln G(N, \theta)}{\partial N^2} < 0$.
2. $\lim_{N \to 1} \frac{\partial \ln G(N, \theta)}{\partial \ln N} > \frac{1}{2}$.
3. $\lim_{N \to \infty} \frac{\partial \ln G(N, \theta)}{\partial \ln N} < 0$.

Note that some of our main results rely on monotone comparative static arguments and thus do not require convexity or uniqueness.

### 1.4 Two Types of Industries

There are two types of industries, ideas-oriented and cost-oriented. In ideas-oriented industries, consumers highly value functionality $N$ so that $D_N > 0$ is salient. Further, high-productivity firms develop the best functions in the sense that each function generates a high marginal revenue. Mathematically, $D_N$ is increasing in $\theta$ or $D(N, \theta)$ is log supermodular in $(N, \theta)$. One can get at this same notion of ideas-oriented industries from the cost side by noting that in these industries, high-productivity firms are really good at managing the integration of complex designs. With complex designs, more functionality raises the marginal costs for each supplier because each firm-supplier pair must ensure its design is compatible with all the other suppliers’ designs. That is $C_N > 0$. Moreover, this problem is more salient for more productive firms. That is $C(N, \theta)$ is log supermodular in $(N, \theta)$. Thus in cost-oriented industries, more productive firms hire managers who are better at keeping cost low. Whether tackled from the demand side or the supply side, both imply the following:
Assumption 4. Ideas-oriented industries: $G(N, \theta)$ is log-supermodular in $(N, \theta)$.

In cost-oriented industries, the manager is good at keeping costs down ($C_\theta < 0$), but this focus on lower costs excludes a focus on good functionality. Porter’s (1996) “What is strategy?” is precisely about the tension between being cost-efficient and being able to build a product with many intertwined functions that support each other and cannot be disentangled (‘cream skimming’) by competitors. This means that in cost-oriented industries, high-productivity firms do not get a big bang for their functionality. Mathematically, $D_N$ is decreasing in $\theta$.

Assumption 5. Cost-oriented industries: $G(N, \theta)$ is log-submodular in $(N, \theta)$.

1.5 Ideas-Oriented Industry

Taking the log-transformation of the firm’s problem in (FP2) yields

$$\max_{k \in \{O, V\}, N \in [1, \infty)} \pi^k(N, \theta, \eta) = \tilde{a} + g(N, \theta) + \psi(\gamma^k(N), \eta), \quad (fp1)$$

where $\pi^k(N, \theta, \eta) \equiv \ln \Pi^k(N, \theta, \eta)$, $\tilde{a} \equiv 1/(1 - \alpha) \ln \hat{A}$, $g(N, \theta) \equiv \ln G(N, \theta)$, and $\psi(\gamma, \eta) \equiv \ln \Psi(\gamma, \eta)$.

By choosing $k \in \{O, V\}$, the firm is indirectly choosing the value of $\delta^k \in \{\delta^O, \delta^V\}$. To find the optimal $\delta^k$ we adopt the methodology used in AH (2004), where we begin by allowing the firm to treat $\delta$ as a continuous variable on the interval $(0, 1)$. Then (fp1) generalizes to

$$\max_{\delta \in (0,1), N \in [1,\infty]} \pi(N, \delta, \theta, \eta) = \tilde{a} + g(N, \theta) + \psi(\gamma(N, \delta), \eta), \quad (fp2)$$

where $\gamma(N, \delta) \equiv \frac{\delta \alpha N + 1}{N + 1}$. Since the transformation from (FP2) to (fp1) is monotone, the optimal $k$ and $N$ that solve (FP2) also solve (fp1). By Assumption 4, $G(N, \theta)$ is log-supermodular in $(N, \theta)$, so $g(N, \theta)$ is supermodular in $(N, \theta)$. In equation (fp1), $N$ and $\theta$ jointly appear in $g(N, \theta)$ only, so the log-profit function $\pi(N, \delta, \theta, \eta)$ is supermodular in $(N, \theta)$.

In ideas-oriented industries, the profit function is log-supermodular in $(N, \theta)$, meaning a more productive firm has a higher profit margin from a larger $N$. However, the firm’s revenue share $\gamma^k(N)$ is decreasing in $N$. Therefore, a more productive firm is more likely to choose $k = V$ because integration helps mitigate the firm’s loss in revenue share from a larger $N$. This tension only works in certain industries. In industries with extremely low $\eta$, suppliers’ inputs
are extremely important. The firm will always find it optimal to incentivize its suppliers by outsourcing. In industries with extremely high $\eta$, the firm’s inputs are much more important than the suppliers’. The firm always chooses $k = V$ to incentivize itself. Therefore, there should be two threshold values of $\eta$, $\underline{\eta}$ and $\bar{\eta}$, such that in industries with $\eta < \underline{\eta}$, firms always choose $k = O$ regardless of its productivity. In industries with $\eta > \bar{\eta}$, firms always choose $k = V$ regardless of its productivity. For those industries in between, a more productive firm has a higher $N$ and so chooses $k = V$ to compensate its lower revenue share. A less productive firm has a lower $N$ and so has a higher revenue share even with $k = O$. This intuition is formalized in Theorem 1 and Theorem 2.\footnote{The proofs for Theorem 1 and Theorem 2 are provided in Appendix 1.D.}

**Theorem 1.** Consider an ideas-oriented industry, i.e., $G(N, \theta)$ is supermodular in $(N, \theta)$. There exist two threshold values of $\eta$, $\eta^L_{io}$ and $\eta^H_{io}$, with $0 < \eta^L_{io} < \eta^H_{io} < 1$, such that:

1. For $\eta < \eta^L_{io}$ industries, all firms choose outsourcing;
2. For $\eta > \eta^H_{io}$ industries, all firms choose vertical integration;
3. For $\eta^L_{io} < \eta < \eta^H_{io}$ industries, there exists a $\tilde{\theta}_{io}(\eta)$, such that
   (a) firms with $\theta < \tilde{\theta}_{io}(\eta)$ choose outsourcing;
   (b) firms with $\theta > \tilde{\theta}_{io}(\eta)$ choose vertical integration;
   (c) $\tilde{\theta}_{io}(\eta)$ is strictly decreasing in $\eta$.

Compared to Antrás (2003), Antrás and Helpman (2004, 2008) and Acemoglu et al. (2007), we have heterogeneity of organizational form within industry that does not rely on assumptions about fixed organizational cost. In Antrás (2003), all firms outsource in small $\eta$ industries. In Antrás and Helpman (2004, 2008), productive firms integrate because of higher fixed costs of integration ($f_V > f_O$). In Acemoglu et al. (2007), firms never integrate because there is no firm relationship-specific investment.

For the remainder of this paper, we focus on those industries with heterogeneous organizational forms, meaning we focus on industries with $\eta \in (\underline{\eta}, \bar{\eta})$. As previously stated, we solve for the firm’s problem (fp1) in these industries by indirectly solving for equation (fp2).

The firm’s problem in (fp1) can be broken down into two steps. First, the firm chooses an optimal $N$ for each organizational form $k$. Denote this choice by $N^k(\theta, \eta), k \in \{O, V\}$, and the resulting profits by $\pi^k(N^k(\theta, \eta), \theta, \eta)$. The firm then compares its profit under $k = O, V$, and
chooses the \( N^k(\theta, \eta) \) that brings it a higher profit. Denote this optimal solution to (fp1) by \( N^* (\theta, \eta) \).

**Theorem 2.** In industries with \( \eta_{io}^L < \eta < \eta_{io}^H \), the following results are true:

1. \( N^O(\theta, \eta), N^V(\theta, \eta) \) and \( N^* (\theta, \eta) \) are strictly increasing in \( \theta \).
2. \( \gamma^O(N^O(\theta, \eta)) \) and \( \gamma^V(N^V(\theta, \eta)) \) are strictly decreasing in \( \theta \).
3. \( N^O(\tilde{\theta}_{io}(\eta), \eta) < N^V(\tilde{\theta}_{io}(\eta), \eta) \) and \( \gamma^O(N^O(\tilde{\theta}_{io}(\eta), \eta)) < \gamma^V(N^V(\tilde{\theta}_{io}(\eta), \eta)) \).

Parts 1 and 2 of Theorem 2 capture the key tradeoff of the paper: a more productive firm has a larger supplier network (larger \( N \)), but also a larger hold-up problem (a smaller Shapley value or share of revenue \( \gamma \)). Part 3 deals with a firm that is just indifferent between the two organizational forms. By Theorem 1, this firm has productivity \( \theta = \tilde{\theta}(\eta) \) As the firm moves from \( O \) to \( V \), two offsetting things happen to its share of revenue. The direct effect is the improved outside option \( (\delta^O < \delta^V) \), which raises its share of revenue. The indirect effect is that the firm increases its supplier network \( (N^O < N^V) \) which lowers the firm’s share of revenue. Part 3 states that the direct effect dominates, meaning the revenue share is higher under \( V \). Part 3 is ancillary to part 1 and 2.

### 1.6 Cost-Oriented Industries

In the cost-oriented industries, a firm’s profit function is log-submodular in \((N, \theta)\), meaning that for a more productive firm, higher functionality lowers profit. Therefore, more productive firms have a higher incentive to choose low functionality. Lower functionality generates a higher revenue share for the firm, and thus a lower revenue share for the suppliers. This exacerbates the incentive problem because a lower revenue share hurts the suppliers’ investment incentives. The firm would then choose outsourcing to mitigate these incentive distortions.

However, similar to the ideas-oriented industries, in industries with extremely low \( \eta (\eta < \tilde{\eta}) \), the firm’s investment is not as important. The firm would always find it optimal to incentivize the suppliers by choosing \( k = O \). In industries with extremely high \( \eta (\eta > \tilde{\eta}) \), firms always find it optimal to choose \( k = V \) to incentivize itself. In the industries in-between \((\eta \in (\tilde{\eta}, \eta))\), more productive firms choose lower \( N \) and choose \( k = O \) to mitigate the incentive distortion, whilst less productive firms choose higher \( N \) and choose \( k = V \) to mitigate the incentive distortion. This intuition is formally stated in Theorem 3 and Theorem 4.
Theorem 3. There exist two threshold values of $\eta$, $\underline{\eta}$ and $\bar{\eta}$, with $0 < \underline{\eta} < \bar{\eta} < 1$, such that:

1. For $\eta < \underline{\eta}$ industries, all firms choose outsourcing;
2. For $\eta > \bar{\eta}$ industries, all firms choose vertical integration;
3. For $\underline{\eta} < \eta < \bar{\eta}$ industries, there exists a $\tilde{\theta}(\eta)$, such that
   (a) firms with $\theta > \tilde{\theta}(\eta)$ choose outsourcing;
   (b) firms with $\theta < \tilde{\theta}(\eta)$ choose vertical integration;
   (c) $\theta(\eta)$ is strictly increasing in $\eta$.

Theorem 4. In industries with $\underline{\eta} < \eta < \bar{\eta}$, the following results are true:

1. $N^O(\theta, \eta)$, $N^V(\theta, \eta)$ and $N^*(\theta, \eta)$ are strictly decreasing in $\theta$.
2. $\gamma^O(N^O(\theta, \eta))$ and $\gamma^V(N^V(\theta, \eta))$ are strictly increasing in $\theta$.
3. $N^O(\tilde{\theta}(\eta), \eta) > N^V(\tilde{\theta}(\eta), \eta)$ and $\gamma^O(N^O(\tilde{\theta}(\eta), \eta)) < \gamma^V(N^V(\tilde{\theta}(\eta), \eta))$.

Parts 1 and 2 of Theorem 4 summarize the firm’s trade-off in cost-oriented industries. A more productive firm maintains a smaller network, but suffers from smaller hold-up problems. Part 3 describes the behavior of the threshold firm: the firm with productivity $\tilde{\theta}(\eta)$ is indifferent between outsourcing and integration. As the firm switches from outsourcing to integration, the direct effect is the increase in its outside option, which raises the firm’s revenue share $\gamma$. The indirect effect is a decrease in its production network ($N^O > N^V$), which lowers the firm’s revenue.

The most interesting implication from the cost-oriented industries (and in the ideas-oriented industries) is not the correlation between productivity $\theta$ and organizational decision $k$, but the correlation between the scope of production networks $N$ and organizational decisions $k$. In the ideas-oriented industries, larger firms (firms with more suppliers) tend to be vertically integrated. In the cost-oriented industries, larger firms tend to outsource. This will be further discussed in Section 1.7.

1.7 Conclusion

This paper develops a theoretical framework that connects firms’ quality, scope, and boundary decisions. A firm’s product quality depends on its functionality: a higher functionality requires a firm to manage a more complex production network with more inputs (and thus more suppliers).
With more suppliers, the haggling problem that is typical in holdup problems (Whinston, 2001) is more severe. I solve for the Shapley value for the firm and its suppliers under Nash bargaining. The Shapley values indicate that the firm’s revenue share is decreasing in the size of its production network (number of suppliers). The firm thus faces a trade-off between a higher-quality product and a higher revenue share: a higher-quality product shifts up consumer demand and generates higher profit for the firm; however, a higher-quality product also requires the firm to manage a more complex production network. The haggling within this production network lowers the firm’s profit.

In ideas-oriented industries, a higher productivity firm is better at managing complex production networks. The profit margin for a high functionality product is high, meaning the firm’s profit function is supermodular in functionality $N$ and productivity $\theta$. A more productive firm would choose higher functionality, since higher functionality lowers the firm’s revenue share. A more productive firm then chooses integration to mitigate its loss in revenue share. In these industries, a more productive firm chooses integration, while less productive firms choose outsourcing.

In cost-oriented industries, a higher productivity firm is better at lowering costs. Keeping functionality $N$ constant, a more productive firm can lower the marginal cost for each member in its production network. A product with higher functionality is also more costly to produce. When functionality increases costs by too much, firms may prefer to design products with less functionality, but with higher quality for each function. A higher-productivity firm is better at lowering cost, but is not suited for managing a large team. In these industries, more productive firms choose lower functionality, but produce each function at a better quality. Lower functionality induces a higher revenue share for the firm, and lower revenue shares for its suppliers. This hurts the incentives of the suppliers, more so for more productive firms. In these industries, more productive firms will end up choosing outsourcing to mitigate the incentive distortions, so more productive firms choose outsourcing, while less productive firms choose integration.

The model in this paper provides a framework of analysis in several directions. First, the distinction between idea- and cost-oriented industries allows one to interpret two types of empirical patterns. In some industries, larger firms largely use outsourcing (e.g., Apple). In others, larger firms are integrated (automobile). If one relaxes the parameter $\eta$ in this model to be firm-specific, then the model allows for mixed patterns within industry. For example, Xbox
is outsourced by Microsoft, while PlayStation is produced in-house by Sony. This connects with the literature on multi-product firms and international trade. See the literature review from Section 1.1.

Second, since non-contractible, relationship-specific investments play a crucial role in this paper, this paper provides a natural framework for examining the impact of contractibility and legal institutions on firms’ scope, quality, and make-or-buy decisions. Simply put, if a firm sources from suppliers from a developed country, then the firm’s outside option under vertical integration is high because a developed country provides better protection for the firm’s non-contractible, relationship-specific investment, which increases the firm’s outside option, and hence its revenue share under integration. The firm is then less likely to integrate its suppliers because integration exacerbates the incentive distortion on the suppliers’ part. The opposite is true if the suppliers were from a developing country. With poorer protection on its non-contractible, relationship-specific investments, the firm is more likely to integrate its suppliers. This extends the line of research on the relationship between institution quality and international trade.¹⁴

Last but not least, this framework constitutes as part of an active research area on global supply chains, as in Antrás and Chor (2013), Costinot et al. (2013), and Alfaro et al. (2015). This literature focuses on the geography and boundary decisions in a global value chain. But so far there is no specific discussion of scope decisions.

In conclusion, this paper develops a theoretical tool for analysis in several directions. It provides interesting references for my future research.

Appendix

1.A Firm’s Shapley values under various functional forms

1.A.1 Leontief production function

Each player’s Shapley value is the average of her contributions to all coalitions that consist of players ordered below her in all permutations of the order. A coalition generates one of three possible values.

1. In a coalition without the firm, the value is 0.
2. In a coalition with the firm and all the suppliers, the value is revenue \( R = \hat{AD}(N, \theta)^{\alpha} q^{\alpha} \), where \( q = h^{\eta} m^{1-\eta} / \hat{\eta} \) as in the statement of the Lemma.
3. In a coalition with the firm, but not all the suppliers, the minimum quality is \( \Delta^k q \) so that the value is revenue \( \delta^k R = \hat{AD}(N, \theta)^{\alpha} (\Delta^k q)^{\alpha} \), where \( \delta^k \equiv (\Delta^k)^{\alpha} \).

Consider the firm’s contribution. Pick a permutation (a ranking of each player from 0 to \( N \)) and let \( g(B) \) be the firm’s rank in this permutation. If \( g(B) < N \) then there is at least one supplier not in the coalition and the firm’s contribution is \( \delta^k R \) i.e., case 3 less case 1. If \( g(B) = N \) then all suppliers are in the coalition and the firm’s contribution is \( R \) i.e., case 2 less case 1. The share of permutations with \( g(B) = N \) is \( 1/(N+1) \). The share of permutations with \( g(B) < N \) is \( N/(N+1) \). Therefore, the firm’s Shapley value is

\[
R \frac{1}{N+1} + \delta^k R \frac{N}{N+1} = \frac{\delta^k N + 1}{N+1} \frac{R}{N+1}
\]

The value generated by a coalition of all players is \( R \) (case 2). Since the Shapley value is efficient, suppliers must receive

\[
R - \frac{\delta^k N + 1}{N+1} R = \frac{1 - \delta^k}{N+1} NR
\]

The Shapley value is symmetric so that all suppliers have the same Shapley value. Dividing the above expression by the \( N \) suppliers gives each supplier’s Shapley value: \( [(1-\delta^k)/(N+1)]R \).

1.A.2 CES function

Suppose the demand-shifter is:

\[
\varphi = D(N, \theta) N^{-1/\beta} Q,
\]
where \( Q = \left( \sum_{j=1}^{N} q_j^\beta \right)^{1/\beta} \) is the overall quality of the firm’s product.

In a symmetric equilibrium, \( q_j = q \) for all \( j \). The product quality is

\[
Q = (Nq^\beta)^{1/\beta} = N^{1/\beta}q.
\]

The demand-shifter is

\[
\varphi = D(N, \theta)N^{-1/\beta}Q = D(N, \theta)N^{-1/\beta}N^{1/\beta}q = D(N, \theta)q.
\]

The revenue is

\[
R = \hat{A}\varphi^\alpha = \hat{A}D(N, \theta)^\alpha q^\alpha.
\]

The scale effect from CES is killed by \( N^{-1/\beta} \). Similar to the previous section, a coalition generates one of three possible values:

1. In a coalition without the firm, the value is 0.
2. In a coalition with the firm and all the suppliers, the value is revenue \( R = \hat{A}D(N, \theta)^\alpha q^\alpha \).
3. In a coalition with the firm, but not all the suppliers, the overall quality is

\[
\left[ nq^\beta + (N - n)(\Delta^k q)^\beta \right]^{1/\beta} = [n + (N - n)(\Delta^k)^{1/\beta}]^{1/\beta}q,
\]

where \( n \) is the number of suppliers who are in the coalition. The demand-shifter generated by this coalition is

\[
D(N, \theta)N^{-1/\beta}Q = D(N, \theta)[n/N + (1 - n/N)(\Delta^k)^{1/\beta}]q.
\]

The value of this coalition is

\[
\hat{A}\{D(N, \theta)[n/N + (1 - n/N)(\Delta^k)^{1/\beta}]q\}^\alpha = [n/N + (1 - n/N)(\Delta^k)^{1/\beta}]^{\alpha/\beta}R.
\]

The firm’s Shapley value can be calculated as:

\[
\frac{1}{N+1} \sum_{n=0}^{N} \{[1 - (\Delta^k)^{\beta}]n/N + (\Delta^k)^{\beta}\}^{\alpha/\beta}R \equiv \gamma(N, \Delta^k)R.
\]
where $\gamma(N, \Delta^k)$ is the firm’s revenue share, and

$$\gamma(N, \Delta^k) \equiv \sum_{n=0}^{N} \left\{ [1 - (\Delta^k)^\beta] n/N + (\Delta^k)^\beta \right\}^{\alpha/\beta} / N + 1.$$  

It can be easily shown that $\gamma(N, \Delta^k)$ is strictly decreasing in $N$ and strictly increasing in $\Delta^k$. It has the same properties as the Leontief production function. Moreover, when $\beta \to -\infty$, $\gamma(N, \Delta^k) \to \delta^k N + 1$, which is the firm’s revenue share under the min function.

1. A. 3 O-ring function

Suppose the overall quality $Q$ is an o-ring function of the individual functions:

$$Q = \prod_{j=1}^{N} q_j.$$  

The firm’s demand shifter is $\varphi = D(N, \theta)Q$, and its revenue is $R = \hat{A}D(N, \theta)^\alpha Q^\alpha$. In a symmetric equilibrium, $q_j = q$ for all $j$. The firm’s quality is $Q = q^N$, and its revenue is $R = \hat{A}D(N, \theta)^\alpha q^{\alpha N}$. Again, there are three values that can be generated by a coalition:

1. In a coalition without the firm, the value is 0.

2. In a coalition with the firm and all the suppliers, the value is revenue $R = \hat{A}D(N, \theta)^\alpha q^{\alpha N}$.

3. In a coalition with the firm, but not all the suppliers, the overall quality is $Q = q^n(\Delta^k q)^{N-n} = (\Delta^k)^N q^N$, where $n$ is the number of suppliers who are in the coalition. The corresponding value is $\hat{A}D(N, \theta)^\alpha Q^\alpha = (\Delta^k)^{\alpha(N-n)} R$. Note again, case 2 is a special case of case 3 when $n = N$.

The firm’s Shapley value is

$$\frac{1}{N+1} \sum_{n=0}^{N} (\delta^k)^{N-n} R = \gamma(N, \delta^k) R,$$

where $\delta^k \equiv (\Delta^k)^\alpha$, and $\gamma(N, \delta^k) \equiv 1 - (\delta^k)^{N+1} / (1-\delta^k)(N+1)$ is increasing in $\delta^k$ and decreasing in $N$. 
1.B Proof of the Existence of a SSPE

First, consider the firm’s problem:

\[
\max_{(h_1, h_2, \ldots, h_N)} \gamma^k(N) \frac{\hat{A}}{\eta^\alpha} D(N, \theta)^\alpha \min\{\alpha_1 m_1^{\alpha(1-\eta)}, \ldots, \alpha_N m_N^{\alpha(1-\eta)}\} - w h C(N, \theta) \sum_{j=1}^{N} h_j
\]

Suppose all suppliers stick to their equilibrium strategies. The firm’s problem can be simplified to

\[
\max_{(h_1, h_2, \ldots, h_N)} \gamma^k(N) \frac{\hat{A}}{\eta^\alpha} D(N, \theta)^\alpha \min\{\alpha_1 h_1^\alpha_1, \ldots, \alpha_N h_N^\alpha_N\} m^{\alpha(1-\eta)} - w h C(N, \theta) \sum_{j=1}^{N} h_j
\]

If the firm deviates by choosing \((h_1, h_2, \ldots, h_N) \neq (h, h, \ldots, h)\):

1. The firm always chooses \((h_1, h_2, \ldots, h_N)\) such that \(h_1 = h_2 = \ldots = h_N\). If not, then the firm can do strictly better by lowering the levels of all \(h_i > \min_{j=1,2,\ldots,N}\{h_j\}\) to \(h_i = \min_{j=1,2,\ldots,N}\{h_j\}\), so firm’s problem can be further simplified to:

\[
\max_{h'} \gamma^k(N) \frac{\hat{A}}{\eta^\alpha} D(N, \theta)^\alpha (h')^{\alpha N} m^{\alpha(1-\eta)} - w h C(N, \theta) N h'.
\]

2. It is never optimal for the firm to choose \(h' \neq h\) because the objective function is strictly concave in \(h'\), so \(h' = h\) is, by definition, the unique maximizer of the objective function.

Now consider supplier \(j\)’s problem:

\[
\max_{m_j} \frac{1 - \gamma^k(N)}{N} \frac{\hat{A}}{\eta^\alpha} D(N, \theta)^\alpha \min\{\alpha_1 m_j^{\alpha(1-\eta)}, \ldots, \alpha_N m_N^{\alpha(1-\eta)}\} - w_m C(N, \theta) m_j
\]

Suppose the firm and all the other players stick to the equilibrium strategy. Supplier \(j\)’s problem can be written as:

\[
\max_{m_j} \frac{1 - \gamma^k(N)}{N} \frac{\hat{A}}{\eta^\alpha} D(N, \theta)^\alpha m^{\alpha(1-\eta)} \min\{m_j^{\alpha(1-\eta)}, \ldots, m_N^{\alpha(1-\eta)}\} - w_m C(N, \theta) m_j
\]

If supplier \(j\) deviates by choosing \(m_j \neq m\), supplier \(j\) will be strictly worse off because supplier’s objective function is strictly concave in \(m_j\), which means \(m_j = m\) is the unique maximizer of the supplier’s objective function.
1.C Firm and suppliers’ levels of investments

Substituting in \( h_j = h \) and \( m_j = m \), \( \forall j \) to the firm and the supplier’s problems in the above lemma and solving for \( h \) and \( m \) gives the following expressions:

\[
h^k(N, \theta, \eta) = \left\{ \frac{\alpha A D(N, \theta)^\alpha}{\eta NC(N, \theta)} \left[ \frac{\eta \gamma^k(N)}{w_h} \right]^{1-\alpha+\alpha \eta} \left[ \frac{(1-\eta)(1-\gamma^k(N))}{w_m} \right]^{\alpha-\alpha \eta} \right\}^{1/(1-\alpha)}
\]

\[
m^k(N, \theta, \eta) = \left\{ \frac{\alpha A D(N, \theta)^\alpha}{\eta NC(N, \theta)} \left[ \frac{\eta \gamma^k(N)}{w_h} \right]^{\alpha \eta} \left[ \frac{(1-\eta)(1-\gamma^k(N))}{w_m} \right]^{1-\alpha \eta} \right\}^{1/(1-\alpha)}
\]

Substituting them into the definition of \( q, \varphi \) and \( R \), we can get the following expressions:

\[
q^k(N, \theta, \eta) = \left\{ \frac{\alpha A D(N, \theta)^\alpha}{\eta NC(N, \theta)} \left[ \frac{\eta \gamma^k(N)}{w_h} \right]^{\eta \left( 1-\gamma^k(N) \right)} \left( \frac{1-\gamma^k(N)}{w_m} \right)^{1-\eta} \right\}^{1/(1-\alpha)}
\]

\[
\varphi^k(N, \theta, \eta) = \left\{ \frac{\alpha A D(N, \theta)^\alpha}{\eta NC(N, \theta)} \left[ \frac{\eta \gamma^k(N)}{w_h} \right]^{\eta \left( 1-\gamma^k(N) \right)} \left( \frac{1-\gamma^k(N)}{w_m} \right)^{1-\eta} \right\}^{1/(1-\alpha)}
\]

\[
R^k(N, \theta, \eta) = \left\{ \frac{\alpha A^{1/\alpha} D(N, \theta)^\alpha}{\eta NC(N, \theta)} \left[ \frac{\eta \gamma^k(N)}{w_h} \right]^{\eta \left( 1-\gamma^k(N) \right)} \left( \frac{1-\gamma^k(N)}{w_m} \right)^{1-\eta} \right\}^{\alpha/(1-\alpha)}
\]

1.D Marginal firms in the ideas-oriented industry

Recall from section 1.5 that the firm’s ‘actual’ problem is

\[
\max_{k \in \{O,V\}, N \in [1,\infty)} \pi^k(N, \theta, \eta) = \tilde{a} + g(N, \theta) + \psi^k(N, \eta).
\]  \hspace{1cm} (fp1)

We indirectly solve for this problem by solving the following problem

\[
\max_{\delta \in (0,1), N \in [1,\infty]} \pi(N, \delta, \theta, \eta) = \tilde{a} + g(N, \theta) + \psi(N, \delta, \eta).
\]  \hspace{1cm} (fp2)

Firm’s problem in (fp1) can be broken down into two steps. First, the firm chooses an optimal \( N \) for each organizational form \( k \). Denote these choices by \( N^k(\theta, \eta), k \in \{O, V\} \), and the resulting profits by \( \pi^k(N^k(\theta, \eta), \theta, \eta) \). The firm then compares \( \pi^O(N^O(\theta, \eta), \theta, \eta) \) and \( \pi^V(N^V(\theta, \eta), \theta, \eta) \), and chooses the \( k \) that brings a higher profit. If there is a firm that is indifferent between \( k = O \)
and $k = V$, denote this firm’s productivity by $\tilde{\theta}(\eta)$.

### 1.D.1 Strict concavity of the profit function

$
\pi(N, \delta, \theta, \eta)$ is strictly concave in $(N, \delta)$ if and only if its Hessian matrix is negative definite. Its Hessian matrix, using first order conditions to simplify, can be written as

$$
\begin{pmatrix}
\pi_{NN} & \pi_{N\delta} \\
\pi_{\delta N} & \pi_{\delta\delta}
\end{pmatrix} =
\begin{pmatrix}
g_{NN} + \psi_{\gamma\gamma} \gamma_N^2, & \psi_{\gamma\gamma} \gamma_N \gamma_\delta \\
\psi_{\gamma\gamma} \gamma_N \gamma_\delta, & \psi_{\gamma\gamma} \gamma_\delta^2
\end{pmatrix},
$$

(1.5)

The above matrix is negative definite if and only if $g_{NN}$ and $\psi_{\gamma\gamma}$ are both negative.\(^{15}\)

$$
\psi_{\gamma\gamma} = -\left\{ \frac{\alpha(2\eta - 1)}{1 - \alpha[\gamma\eta + (1 - \gamma)(1 - \eta)]} \right\}^2 - \frac{\alpha}{1 - \alpha} \left[ \frac{\eta}{\gamma^2} + \frac{1 - \eta}{(1 - \gamma)^2} \right] < 0,
$$

so $\pi(N, \delta, \theta, \eta)$ is strictly concave if and only if

$$
g_{NN} = \frac{\alpha}{1 - \alpha} \left\{ \frac{\partial^2 \ln D(N, \theta)}{\partial \ln N} - \frac{\partial^2 \ln C(N, \theta, \eta)}{\partial \ln N} + 1 \right\} < 0,
$$

or

$$
\frac{\partial^2 \ln C(N, \theta, \eta)}{\partial \ln N} > \frac{\partial^2 \ln D(N, \theta)}{\partial \ln N} + 1.
$$

By Assumption 3, $G(N, \theta)$ is strictly log-concave in $N$, so $g_{NN} < 0$. The Hessian matrix is thus negative definite, meaning $\pi(N, \delta, \theta, \eta)$ is strictly concave in $(N, \delta)$. The optimal $(N, \delta)$ that solve (fp2) are complements. Denote the relationship between $\delta$ and $N$ by $\delta(N)$, Figure 1.1 shows a simulation for $\delta(N)$.

Note that since we do not assume the specific functional form of $G(N, \theta)$, we cannot assume the range of $\delta(N)$.

### 1.D.2 The marginal firm’s organizational behavior

The marginal firm’s behavior helps us pin down the other firms’ behaviors within the same industry. Solving for the marginal firm’s problem in (fp2) also renders a function $\delta(N)$, as shown in Figure 1.1.

\(^{15}\)A $2 \times 2$ matrix is negative definite if and only if its first determinant is negative and its second determinant is positive. In this case, these conditions translate into $g_{NN} + \psi_{\gamma\gamma} \gamma_N^2 < 0$ and $g_{NN} \psi_{\gamma\gamma} \gamma_\delta^2 > 0$. These two inequalities translate into the condition that $g_{NN} < 0$ and $\psi_{\gamma\gamma} < 0$. 
Figure 1.1: Simulation of $\delta(N)$

Now relate to the firm’s actual problem in equation (fp1). The firm cannot choose any combination of $(N, \delta)$ on $\delta(N)$. Instead, the firm can only choose $N$ from two horizontal lines $\delta = \delta^O$ and $\delta = \delta^V$. Let us assume that $\delta(N)$ crosses both lines for now. We’ll deal with the other cases later.

Suppose $\delta(N)$ crosses $\delta = \delta^O$ at $(n^O, \delta^O)$ and $\delta = \delta^V$ at $(n^V, \delta^V)$. Since $\delta(N)$ is increasing in $N$ and $\delta^O < \delta^V$, it must be that $n^O < n^V$. Denote the marginal firm’s choice under $k = O$ and $k = V$ by $N^O$ and $N^V$.\(^{16}\) Depending on the values of $N^O$ and $N^V$ relative to the interval $(n^O, n^V)$, there are 9 cases, as shown in Table 1.1 below.

<table>
<thead>
<tr>
<th>$N^O &lt; n^O$</th>
<th>$n^O &lt; N^V &lt; n^V$</th>
<th>$n^V &lt; N^V$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>$n^O &lt; N^O &lt; n^V$</td>
<td>$N^O &lt; N^V$, $\gamma^O &lt; \gamma^V$</td>
<td>N/A</td>
</tr>
<tr>
<td>$n^V &lt; N^O$</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

We show that for $\tilde{\theta}(\eta)$ to exist, only one out of the nine scenarios in Table 1.1 is possible: when $N^O, N^V \in (n^O, n^V)$. We also show that in this scenario, the marginal firm’s choice

\(^{16}\)They are the shortened forms for $N^O(\tilde{\theta}(\eta), \eta)$ and $N^V(\tilde{\theta}(\eta), \eta)$.
satisfies $N^O < N^V$ and $\gamma^O < \gamma^V$.\footnote{Note that this is item 3 in Theorem 2}

$N^V < n^O$ contradicts the definition of $\tilde{\theta}(\eta)$

Refer to Figure 1.2. Recall that we assume $\delta(N)$ crosses $\delta = \delta^O$ at $(n^O, \delta^O)$ and $\delta = \delta^V$ at $(N^V, \delta^V)$. If $N^V < n^O$, moving from $(N^V, \delta^V)$ to $(N^V, \delta^O)$ increases firm’s profit $\pi(N, \delta, \theta, \eta)$, because keeping $N = N^V$ constant, we are approaching the optimal $\delta$ at $\delta(N^V)$.

The profit at $(N^V, \delta^V)$ is $\pi(N^V, \delta^V, \tilde{\theta}, \eta)$. The profit at $N^V, \delta^O$ is $\pi(N^V, \delta^O, \tilde{\theta}, \eta)$. The former is smaller than the latter meaning

$$\pi(N^V, \delta^V, \tilde{\theta}, \eta) < \pi(N^V, \delta^O, \tilde{\theta}, \eta).$$

$N^V$ may or may not be the optimal $N$ that maximizes firm’s profit at $\delta = \delta^O$, so

$$\pi(N^V, \delta^O, \tilde{\theta}, \eta) \leq \max_{N \in (1, \infty)} \pi(N, \delta^O, \tilde{\theta}, \eta).$$
Combining the above two equations gives

\[
\pi(N^V, \delta^V, \tilde{\theta}, \eta) < \pi(N^V, \delta^O, \tilde{\theta}, \eta) \leq \max_{N \in (1, \infty)} \pi(N, \delta^O, \tilde{\theta}, \eta). \tag{1.6}
\]

The left end of equation (1.6) is the firm’s optimal profit at \( k = V \). The right end of equation (1.6) is the firm’s optimal profit at \( k = O \). Equation (1.6) contradicts the definition of the marginal firm, because it implies that the marginal firm’s profit under integration is less than its profit under outsourcing.

\( N^O > n^V \) **contradicts the definition of** \( \tilde{\theta}(\eta) \)

Refer to Figure 1.3. If \( N^O > n^V \), moving from \((N^O, \delta^O)\) to \((N^O, \delta^V)\) increases the firm’s profit because keeping \( N = N^O \) constant, we are approaching the optimal \( \delta(N^O) \). The firm’s profit at \((N^O, \delta^O)\) is \( \pi(N^O, \delta^O, \tilde{\theta}, \eta) \). The firm’s profit at \((N^O, \delta^V)\) is \( \pi(N^O, \delta^V, \tilde{\theta}, \eta) \). The former is less than the later, meaning

\[
\pi(N^O, \delta^O, \tilde{\theta}, \eta) < \pi(N^O, \delta^V, \tilde{\theta}, \eta).
\]
\(N^O\) may or may not be the optimal \(N\) that maximizes the marginal firm’s profit at \(k = V\) so

\[
\pi(N^O, \delta^V, \tilde{\theta}, \eta) \leq \max_{N \in [1, \infty)} \pi(N, \delta^V, \tilde{\theta}, \eta).
\]

Combining the above two inequalities gives

\[
\pi(N^O, \delta^O, \tilde{\theta}, \eta) < \pi(N^O, \delta^V, \tilde{\theta}, \eta) \leq \max_{N \in [1, \infty)} \pi(N, \delta^V, \tilde{\theta}, \eta) \tag{1.7}
\]

The left end of equation (1.7) is the firm’s optimal profit under \(k = O\). The right end of equation (1.7) is the firm’s optimal profit under \(k = V\). Equation (1.7) implies that the marginal firm’s profit under outsourcing is lower than its profit under integration. This also contradicts its definition.

\(N^O < n^O < N^V < n^V\) contradicts the definition of \(\tilde{\theta}(\eta)\)

To see this, draw an iso-\(\gamma\) line through \((N^V, \delta^V)\). Since \(\gamma \equiv \frac{\delta N + 1}{N + 1}\) is increasing in \(\delta\) and decreasing in \(N\), the iso-\(\gamma\) lines is upward sloping. Suppose this iso-\(\gamma\) line crosses \(\delta = \delta^O\) at \(N'\). Based on the value of \(N^O\), there are two cases: \(N^O \leq N' < N^V\) and \(N' < N^O < N^V\).

1. \(N^O \leq N' < N^V\)

See Figure 1.4. In this case, both \(N^O\) and \(N^V\) are above \(\delta(N)\), it must be that \(\psi_\gamma(\gamma^O, \eta) < 0\)
and $\psi_\gamma(\gamma^V, \eta) < 0$. By definition, $N^k (k \in \{O, V\})$ satisfies

$$\pi^k_N(N^k, \tilde{\theta}, \eta) = g_N(N^k, \tilde{\theta}) + \psi_\gamma(\gamma^k, \eta)\gamma^k_N = 0.$$ 

$\gamma^k_N = \frac{\delta^k}{(N^k + 1)^2} < 0$ implies that $g_N(N^k, \tilde{\theta})$ and $\psi_\gamma(\gamma^k, \eta)$ are of the same sign. Since $\psi_\gamma(\gamma^k, \eta)$ is negative for $k \in \{O, V\}$, it must be that $g_N(N^k, \tilde{\theta}) < 0$ as well.

By Assumption 3, $G(N, \theta)$ is log-concave in $N$, so $g_{NN}(N, \theta) < 0$, $\forall N$. Since $N^V > N' \geq N^O$, $g_N(N^V, \tilde{\theta}) < g_N(N', \tilde{\theta}) \leq g_N(N^O, \tilde{\theta}) < 0$. If we move from $(\delta^V, N^V)$ to $(\delta^O, N')$ along the iso-$\gamma$ line (the arrow line), $\psi(\gamma, \eta)$ remains constant. But $g(N, \theta)$ increases because $g_N(N, \theta)$ remains negative as we decrease the value of $N$. It follows that profit increases from $(N^V, \delta^V)$ to $(N', \delta^O)$. If we then move from $(N', \delta^O)$ to $(N^O, \delta^O)$, profit continues to increase. Because $g_N(N, \tilde{\theta})$ and $\psi_\gamma(\gamma, \eta)$ remain negative as we decrease $N$ and $\gamma$, so both $g(N, \tilde{\theta}, \eta)$ and $\psi(\gamma)$ increase. This argument implies the following inequalities:

$$\pi(N^V, \delta^V, \tilde{\theta}, \eta) < \pi(N', \delta^O, \tilde{\theta}, \eta) < \pi(N^O, \delta^O, \tilde{\theta}, \eta) \quad (1.8)$$

The left end of equation (1.8) is the marginal firm’s optimal profit under $k = V$, and the right end is the marginal firm’s optimal profit under $k = O$. This inequality contradicts the definition of $\tilde{\theta}(\eta)$.

2. $N' < N^O < N^V$

See Figure 1.5. In this case, $g_N(N^V, \tilde{\theta}) < g_N(N^O, \tilde{\theta}) < 0$ still holds. From $(N^V, \delta^V)$ to $(N^O, \delta')$ along the iso-$\gamma$ line (the arrow line), $\psi(\gamma, \eta)$ remains constant, $g(N, \tilde{\theta})$ increases because $g_N(N, \tilde{\theta}) < 0$ and $N$ decreases, so profit increases. From $(N^O, \delta')$ to $(N^O, \delta^O)$, $g(N, \tilde{\theta})$ remains constant because $N$ does not change. $\psi(\gamma, \eta)$ increases because $\gamma$ decreases and $\psi_\gamma(\gamma, \eta) < 0$, so profit increases along this route (the arrow line). This argument implies the following inequalities:

$$\pi(N^V, \delta^V, \tilde{\theta}, \eta) < \pi(N^O, \delta', \tilde{\theta}, \eta) < \pi(N^O, \delta^O, \tilde{\theta}, \eta) \quad (1.9)$$

The left end of the above inequality is the marginal firm’s optimal profit under $k = V$. The right end is the marginal firm’s optimal profit under $k = O$. This inequality also contradicts the definition of $\tilde{\theta}$.

---

18The fact that $(N^V, \delta^V)$ is above the “ridge” of profit function implies that $\delta$ is “too big”. As $\delta$ decreases, $\gamma$ also decreases, but profit increases, so $\pi_\gamma(\gamma^V, \eta) = \psi_\gamma(\gamma^V, \eta) < 0$. To generalize, whenever $(N, \delta)$ is above $\delta(N)$, $\pi_\gamma = \psi_\gamma > 0$. Whenever $(N, \delta)$ is below $\delta(N)$, $\pi_\gamma = \psi_\gamma < 0$. 
Figure 1.5: \( N' < N^O < N^V \)

\( N^O < n^O < n^V < N^V \) is impossible.

See Figure 1.6. In this case, \( N^O \) is above \( \delta(N) \) and \( N^V \) is below \( \delta(N) \), so \( \psi_\gamma(\gamma^O, \eta) < 0 < \psi_\gamma(\gamma^V, \eta) \). Recall from the previous section that \( g_N(N^k, \tilde{\theta}) \) is of the same sign as \( \psi_\gamma(\gamma, \eta) \), so \( g_N(N^O, \theta) < 0 < g_N(N^V, \theta) \). This implies \( N^O > N^V \) because \( g_{NN}(N, \theta) < 0 \), which contradicts the assumption that \( N^O < N^V \).

\( n^O < N^O < n^V < N^V \) contradicts the definition of \( \tilde{\theta}(\eta) \).

In this case, both \((N^O, \delta^O)\) and \((N^V, \delta^V)\) are below \( \delta(N) \), so \( \psi_\gamma(\gamma^k, \eta) > 0 \) for \( k \in \{O, V\} \). Draw an iso-\( \gamma \) line through \((N^O, \delta^O)\). Suppose it crosses \( \delta = \delta^V \) at \((N', \delta^V)\). There are two possible cases: \( N^O < N' \leq N^V \) and \( N^O < N^V < N' \).

1. \( N^O < N' \leq N^V \)

   See Figure 1.7. Since \( g_{NN}(N, \theta) < 0 \), \( N^O < N' < N^V \), it is implied that \( g_N(N^O, \tilde{\theta}) > g_N(N', \tilde{\theta}) > g_N(N^V, \tilde{\theta}) > 0 \).

   From \((N^O, \delta^O)\) to \((N', \delta^V)\) along the iso-\( \gamma \) lines, \( \gamma \) remains constant while \( N \) increases. \( \psi(\gamma, \eta) \) remains constant and \( g(N, \theta) \) increases, so profit increases.

   From \((N', \delta^V)\) to \((N^V, \delta^V)\), \( \delta \) remains constant and \( N \) increases. \( \psi(\gamma, \eta) \) also increases because \( \gamma \) increases and \( \psi_\gamma(\gamma, \eta) > 0 \). \( g(N, \theta) \) increases because \( N \) increases and \( g_N(N, \theta) > 0 \), so profit

\[19\]Recall from previous analysis that \( g_N(N^k, \tilde{\theta}) \) is of the same sign as \( \psi_\gamma(\gamma^k, \eta) \). We know that \( \psi_\gamma(\gamma^k, \eta) > 0 \) because \((N^k, \delta^k)\) is below \( \delta(N) \) for \( k \in \{O, V\} \), so \( g_N(N^k, \tilde{\theta}) \) is also positive for \( k \in \{O, V\} \).
increases along the arrow route, which implies the following inequalities:

$$\pi(N^O, \delta^O, \tilde{\theta}, \eta) < \pi(N', \delta', \tilde{\theta}, \eta) \leq \pi(N^V, \delta^V, \tilde{\theta}, \eta).$$

(1.10)

The left end of the inequality is the marginal firm’s optimal profit under $k = O$. The right end is the marginal firm’s optimal profit under $k = V$. This inequality contradicts the definition of $\tilde{\theta}$.

2. $N^O < N^V < N'$

See Figure 1.8. In this case, $g_N(N^O, \tilde{\theta}) > g_N(N^V, \tilde{\theta}) > 0$ still holds. From $(N^O, \delta^O)$ to $(N^V, \delta')$ along the iso-$\gamma$ line, $\gamma$ remains constant while $N$ increases, so $\psi(\gamma, \eta)$ remains constant while $g(N, \tilde{\theta})$ increases and profit increases. From $(N^V, \delta')$ to $(N^V, \delta^V)$, $\delta$ increases while $N^V$ remains constant. $g(N, \tilde{\theta})$ remains constant. $\psi(\gamma, \eta)$ increases because $\psi_{\gamma} > 0$ and $\gamma$ increases. Profit increases along the arrow route. This argument implies that

$$\pi(N^O, \delta^O, \tilde{\theta}, \eta) < \pi(N^V, \delta^V, \tilde{\theta}, \eta).$$

This contradicts the definition of $\tilde{\theta}(\eta)$. 

Figure 1.6: $N^O < n^O, N^V > n^V$
Figure 1.7: \( N^O < N' \leq N^V \)

\( N^O, N^V \in (n^O, n^V) \)

We have excluded all the other possibilities in Table 1.1, so if \( \tilde{\theta}(\eta) \) does exist, it must be that \( N^O, N^V \in (n^O, n^V) \). We now prove \( N^O < N^V \) and \( \gamma^O < \gamma^V \).

1. \( N^O < N^V \)

We have shown that \( \pi(N, \delta, \theta, \eta) \) is strictly concave in \( (N, \delta) \). Since \( \delta^V > \delta^O \), it must be that \( N^V > N^O \).

2. \( \gamma^O < \gamma^V \)

See Figure 1.9. From \( (N^O, \delta^O) \) to \( (N^O, \delta') \), \( N \) is constant while \( \delta \) increases, so \( \gamma \) increases because \( \gamma \equiv \frac{\delta N + 1}{N + 1} \) is increasing in \( \delta \). From \( (N^O, \delta') \) to \( (n^V, \delta^V) \) along the iso-\( \gamma \) line, \( \gamma \) remains constant. From \( (n^V, \delta^V) \) to \( (N^V, \delta^V) \), \( \delta \) is constant while \( N \) decreases. \( \gamma \) increases because it is decreasing in \( N \). Along the arrow route, \( \gamma \) increases. \( \gamma = \gamma^O \) at the origin, and \( \gamma = \gamma^V \) at the end, so \( \gamma^O < \gamma^V \).

Now that we have proved that if \( \delta(N) \) crosses both \( \delta = \delta^O \) and \( \delta = \delta^V \), and if the marginal firm exists, then the marginal firm’s choice satisfies \( N^O < N^V \) and \( \gamma^O < \gamma^V \).

We now show that it is impossible for \( \delta(N) \) not to cross \( \delta = \delta^O \) or \( \delta = \delta^V \). To show this, we first need to show that the firm’s choice of \( N \) is finite.
1.D.3 Uniqueness of the threshold productivity level

By Envelope Theorem, \( \pi^k(\theta, \eta) = g_\theta(N^k(\theta, \eta), \theta) > 0 \) if and only if

\[
g_\theta(N^k(\theta, \eta), \theta) = \frac{\alpha}{1 - \alpha} \left\{ \frac{\partial \ln D(N, \theta)}{\partial \theta} - \frac{\partial \ln C(N, \theta, \eta)}{\partial \theta} \right\} > 0,
\]

or

\[
\frac{\partial \ln D(N, \theta)}{\partial \theta} > \frac{\partial \ln C(N, \theta, \eta)}{\partial \theta}.
\]

We have shown in Appendix 1.D.2 that \( N^V(\bar{\theta}, \eta) > N^O(\bar{\theta}, \eta) \). By Assumption 4, \( g(N, \theta) \) is supermodular in \( (N, \theta) \) meaning \( g_\theta(N, \theta) \) is increasing in \( N \). Thus

\[
g_\theta(N^V(\bar{\theta}, \eta), \bar{\theta}) > g_\theta(N^O(\bar{\theta}, \eta), \bar{\theta}),
\]

which is equivalent to

\[
\pi^V(\bar{\theta}, \eta) > \pi^O(\bar{\theta}, \eta).
\]

As \( \theta \) increases, the difference between \( \pi^V \) and \( \pi^O \) increases. This means that if \( \pi^V \) and \( \pi^O \) cross, they can cross only once. This crossing point is \( \bar{\theta}(\eta) \). Therefore, if \( \bar{\theta}(\eta) \) exists, it is unique. See Figure 1.10 for an illustration.
In Figure 1.10, $\pi^V$ and $\pi^O$ cross at $\theta^*$, which is what we define as $\tilde{\theta}(\eta)$. It is obvious from Figure 1.10 that firms with $\theta > \tilde{\theta}(\eta)$ choose integration, while firms with $\theta < \tilde{\theta}(\eta)$ choose outsourcing.

### 1.D.4 Monotonicity of the threshold productivity

By the definition of $\tilde{\theta}(\eta)$,

$$\pi^V(N^V, \tilde{\theta}(\eta), \eta) = \pi^O(N^O, \tilde{\theta}(\eta), \eta).$$

By Implicit Function Theorem,

$$\frac{d\tilde{\theta}(\eta)}{d\eta} = -\frac{\pi^V_{\eta}(N^V, \tilde{\theta}(\eta), \eta) - \pi^O_{\eta}(N^O, \tilde{\theta}(\eta), \eta)}{\pi^V_{\delta}(N^V, \tilde{\theta}(\eta), \eta) - \pi^O_{\delta}(N^O, \tilde{\theta}(\eta), \eta)} = -\frac{\psi_{\eta}(\gamma^V, \eta) - \psi_{\eta}(\gamma^O, \eta)}{g_{\theta}(N^V, \tilde{\theta}) - g_{\theta}(N^O, \tilde{\theta})}$$

$\psi(\gamma, \eta)$ is supermodular in $(\gamma, \eta)$, and by Assumption 4, $g(N, \theta)$ is supermodular in $(N, \theta)$. Therefore, $\gamma^V > \gamma^O$ and $N^V > N^O$ imply $\psi_{\eta}(\gamma^V, \eta) > \psi_{\eta}(\gamma^O, \eta)$ and $g_{\theta}(N^V, \tilde{\theta}) > g_{\theta}(N^O, \tilde{\theta})$. If it exists, $\tilde{\theta}(\eta)$ is decreasing in $\eta$. 

Figure 1.9: $n^O < N^O < N^V < n^V$
1. D. 5 Existence of threshold productivity and threshold industries

In this section we prove that \( \hat{\theta}(\eta) \) exists for \( 0 < \eta < 1 \), and that \( 0 < \underline{\eta} < \bar{\eta} < 1 \), where \( \underline{\eta} \) and \( \bar{\eta} \) are as defined in Theorem 1.

Firms’ choice of scope is limited.

Denote the firm’s choice of \( N \) in (fp2) by \( N(\theta, \eta) \). A sufficient condition for \( 1 < N(\theta, \eta) < \infty \) is \( \lim_{N \to 1} \pi_N(N, \delta, \theta, \eta) > 0 \) and \( \lim_{N \to \infty} \pi_N(N, \delta, \theta, \eta) < 0 \).

\[
\pi_N(N, \delta, \theta, \eta) = \frac{\alpha}{1 - \alpha} \left\{ \partial \left[ \ln \frac{D(N, \theta)}{NC(N, \theta)} \right] / \partial N \right\} - \frac{1 - \delta}{(N + 1)^2} \cdot \psi_r(\gamma(N, \delta), \eta),
\]
\[ \lim_{N \to 1} \pi_N(N, \delta, \theta, \eta) = \frac{\alpha}{1 - \alpha} \cdot g_N(N, \theta) - \frac{1 - \delta}{4} \cdot \psi_\gamma(\frac{\delta + 1}{2}, \eta) \]

\[ > \frac{\alpha}{1 - \alpha} \cdot g_N(N, \theta) - \frac{1 - \delta}{4} \cdot \psi_\gamma(\frac{\delta + 1}{2}, 1) \]

\[ = \frac{\alpha}{1 - \alpha} \left\{ \lim_{N \to 1} g_N(N, \theta) - \frac{(1 - \delta)^2}{2(1 + \delta)(2 - \alpha - \alpha \delta)} \right\} \]

\[ \geq \frac{\alpha}{1 - \alpha} \left\{ \lim_{\delta \to 0} g_N(N, \theta) - \frac{1}{2(1 + \delta)(2 - \alpha - \alpha \delta)} \right\} \]

\[ = \frac{\alpha}{1 - \alpha} \left\{ \lim_{N \to 1} g_N(N, \theta) - \frac{1}{4 - 2\alpha} \right\} \]

\[ > \frac{\alpha}{1 - \alpha} \left\{ \lim_{N \to 1} g_N(N, \theta) - \frac{1}{2} \right\} . \]

By Assumption 3, \( g_N(N, \theta) > 1/2 \), so \( \lim_{N \to 1} \pi_N(N, \delta, \theta, \eta) > 0 \)

\[ \lim_{N \to \infty} \pi_N(N, \delta, \theta, \eta) = \frac{\alpha}{1 - \alpha} \cdot \lim_{N \to \infty} g_N(N, \theta) - \frac{1}{4} \cdot \psi_\gamma(\delta) \]

\[ = \frac{\alpha}{1 - \alpha} \cdot \lim_{N \to \infty} g_N(N, \theta). \]

By Assumption 3, \( \lim_{N \to \infty} g_N(N, \theta) < 0 \), so \( \lim_{N \to \infty} \pi_N(N, \delta, \theta, \eta) < 0 \).

Since \( \lim_{N \to 1} \pi_N(N, \delta, \theta, \eta) > 0 \) and \( \lim_{N \to \infty} \pi_N(N, \delta, \theta, \eta) < 0 \), \( N = 1 \) and \( N = \infty \) can never be optimal. Therefore, \( 1 < N(\theta, \eta) < \infty \), \( \forall \delta, \eta \in (0, 1) \). This implies that \( 1 < N^k(\theta, \eta) < \infty \) for \( k \in \{O, V\} \) because \( \gamma^k \) is a special case of \( \gamma(N, \delta) \).

**Existence of the threshold industries**

We have shown in the previous section that \( 1 < N^k(\theta, \eta) < \infty \), so there must be an inferior and a superior bound for the continuous function \( N(\theta, \eta) \). Define \( \underline{N} \) and \( \bar{N} \) as \(^{20}\):

\[ \underline{N} = \inf_{\theta \in [0,1], \eta \in (0,1)} N(\theta, \eta) \]

and

\[ \bar{N} = \sup_{\theta \in [0,1], \eta \in (0,1)} N(\theta, \eta) \]

Recall that we defined \( (n^O, \delta^O) \) and \( (n^V, \delta^V) \) as the crossing points of \( \delta(N) \) with \( \delta = \delta^O \) and \( \delta = \delta^V \). Using the function \( \gamma = \frac{\delta + 1}{N + 1} \), \( n^O \) and \( n^V \) can be derived as \( n^O = \frac{1 - \gamma(\eta)}{\gamma(\eta) - \delta^O} \) and

\(^{20}\)\( \underline{N} \) and \( \bar{N} \) exist because of the Completeness Axiom.
\[ n^V = \frac{1 - \gamma(\eta)}{\gamma(\eta) - \delta^V}. \]

If \( n^O = \frac{1 - \gamma(\eta)}{\gamma(\eta) - \delta^O} > \bar{N}, \) or if \( \eta < \gamma^{-1}\left(\frac{\delta^O N + 1}{N + 1}\right), \) \( \delta(N) \) is always below \( \delta = \delta^O. \) See Figure 1.11 for illustration. In this case, all firms choose outsourcing.

If \( n^V = \frac{1 - \gamma(\eta)}{\gamma(\eta) - \delta^V} < \bar{N}, \) or if \( \eta > \gamma^{-1}\left(\frac{\delta^V N + 1}{N + 1}\right), \) \( \delta(N) \) is always above \( \delta = \delta^V. \) See Figure 1.12 for illustration. In this case, call firms choose integration.

By continuity of the profit functions, for each value of \( \theta, \) there must be at least one \( \eta, \) such that \( \pi^V(\theta, \eta) = \pi^O(\theta, \eta). \) And there cannot be more than one \( \eta \) that satisfies this condition because this would violate the monotonicity of \( \theta(\eta), \) as proved in Appendix 1.D.3. Therefore, there is a one-to-one mapping from \( \theta \) to \( \eta. \) Since \( \theta(\eta) \) is strictly decreasing, \( \eta(\theta) \) is also strictly decreasing over the interval \( \theta \in [0, 1], \) so \( \eta(1) \leq \eta(\theta) \leq \eta(0). \)

Define \( \underline{\eta} \equiv \eta(1) \) and \( \bar{\eta} \equiv \eta(0). \) For \( \eta < \underline{\eta}, \pi^V(\theta, \eta) - \pi^O(\theta, \eta) < 0 \) for all \( \theta, \) all firms choose outsourcing; For \( \eta > \bar{\eta}, \pi^V(\theta, \eta) - \pi^O(\theta, \eta) > 0, \) all firms choose vertical integration. In other words, \( \underline{\eta} \) and \( \bar{\eta} \) exist and \( 0 < \underline{\eta} < \bar{\eta} < 1. \)

1.E Firms’ scope decisions in the ideas-oriented industry

We have proved that the marginal firm’s behavior satisfies \( N^V > N^O \) and \( \gamma^V > \gamma^O. \) We now prove the first two statements in Theorem 2.
1.E.1 Firm’s scope decision in (fp2) is limited

Since $\pi(N, \delta, \theta, \eta)$ is strictly concave in $(N, \delta)$, the optimal $(N, \delta)$ is determined by the two first order conditions, $\pi_N = 0$ and $\pi_\delta = 0$. Differentiating these two equations with respect to $\theta$ and rearranging,

$$
\begin{pmatrix}
\frac{dN}{d\theta} \\
\frac{d\delta}{d\theta}
\end{pmatrix} = \frac{1}{\text{det}} \begin{pmatrix}
\pi_{\delta\delta} & -\pi_{N\delta} \\
-\pi_{N\delta} & \pi_{NN}
\end{pmatrix} \begin{pmatrix}
\pi_N \\
\pi_\delta
\end{pmatrix}
$$

where $\text{det}$ is the determinant of the Hessian matrix. The above equation can be simplified to

$$
\begin{pmatrix}
\frac{dN}{d\theta} \\
\frac{d\delta}{d\theta}
\end{pmatrix} = \frac{g_{N\theta} \psi_{\gamma\gamma}}{\text{det}} \begin{pmatrix}
\gamma_\delta^2 \\
-\gamma_\delta \gamma_N
\end{pmatrix}
$$

It can be easily shown that $\psi_{\gamma\gamma} < 0$, $\text{det} > 0$. $\gamma_\delta = \frac{N}{N+1} > 0$ and $\gamma_N = \frac{\delta - 1}{(N+1)^2} < 0$. By Assumption 2, $g_{N\theta} > 0$, so $dN/d\theta > 0$, and $d\delta/d\theta > 0$.

1.E.2 Firm’s scope decision is limited.

We know that $\pi^k(N, \theta, \eta) = \bar{a} + g(N, \theta, \eta) + \psi(\gamma^k(N), \eta)$, and $\pi^k_{N\theta} = g_{N\theta}$, so $g_{N\theta} > 0$ implies $\pi^k_{N\theta} > 0$, and that $N^k(\theta, \eta)$ is strictly increasing in $\theta$. 
1.F Proofs for the cost-oriented industry

The proof of Theorems 3 and 4 follows the same logic as the proof of Theorems 1 and 2.
Chapter 2

Global Supply Chains without Principals
Abstract

In a world dominated by global production chains and multinational firms, it is increasingly important to understand why some activities are integrated within firms and others are outsourced. Using a new database of 85,783 buyer-seller relationships in 122 countries and 64 industries, I examine the extent to which these relationships feature (a) buyer integration of the seller, (b) seller integration of the buyer or (c) neither i.e., outsourcing. An important feature of my 85,783 buyer-seller relationships is the coexistence of all three organizational forms. Noting that each relationship involves a buyer industry and a seller industry, the three forms coexist in 28% of all industry pairs and 42% of all industry pairs that have at least some integration. This coexistence has never before been documented. Surprisingly, coexistence is ruled out both empirically and theoretically in the voluminous international trade literature inspired by Antrás and in the smaller literature testing integration theories using large-scale databases (e.g., Acemoglu et al.). In these literatures, a principal decides whether or not to integrate an agent and, since the agent cannot integrate the principal, all integration is one-directional i.e., the coexistence of buyer and seller integration (a and b) is assumed away. To explain coexistence I develop a model without principals: The buyer and seller meet and choose one of options (a), (b) or (c). If the buyer’s relationship-specific, non-contractible investment is ‘relatively important’ (in the Grossman-Hart sense) then the buyer integrates the seller, if the seller’s investment is relatively important then the seller integrates the buyer, and in the intermediate case there is outsourcing. Empirically, I measure the relative importance of these investments using data on buyer and seller R&D. I show that my model explains the above empirical patterns of coexistence and performs much better than all existing models with principals.
2.1 Introduction

Over the past few decades, the development of information and communication technology (ICT) greatly facilitated the global disintegration of production processes. This phenomenon has kindled a large literature studying the global sourcing decisions of multinational firms. As Antràs (2013) argues, property-rights theory (PRT) is a particularly relevant framework for studying the sourcing decisions of multinational firms. This is because international transactions are more subject to contracting disputes and thus more sensitive to contracting environments, and PRT specifically models how imperfect contract enforcement affects firms’ integration decisions. Much empirical evidence has supported the validity of PRT in explaining multinational firms’ sourcing decisions.¹

Many firm boundary studies with a PRT framework tend to make an assumption that in a global supply chain, there is a principal firm making the make-or-buy decisions. This principal firm integrates an agent firm when the principal firm’s non-contractible investment is more important for the relationship, and outsources the agent firm when the agent firm’s non-contractible investment is more important for the relationship. However, the identity of a principal firm is not empirically defined, and is largely customary in previous literature.² The ambiguity of the principal status creates an identification problem: in an integrated buyer-seller relationship, it could be that the buyer is the principal, and the buyer’s non-contractible investment is more important, or that the seller is the principal, and the seller’s non-contractible investment is more important. This means that without assigning the principal status to one side of the relationship, PRT is not testable (Whinston, 2003). In other words, PRT is not falsifiable, because by switching the identity of the principal firm, one can support any significant relationship between integration decision, and the relative importance of buyer/seller firm’s non-contractible investment. Moreover, wrongly assigned principal status would result in estimation bias. This paper attempts to address these issues.

In this paper I contribute to the global sourcing literature in three ways. First, I construct

---

¹For general surveys of the literature on the global sourcing decisions of multinational firms, see survey papers by Antràs (2013) and Nunn (2007). The books by Antràs (2015) and Helpman (2011) contain extensive theoretical and empirical description on the global sourcing literature.

²In an integrated relationship, one may think of the parent company as the principal firm. In an outsourced relationship, it is far less clear who is the principal firm. For example, when Apple outsources the production of the displays for iPad to Samsung, who is the principal firm in this relationship? Not to mention that the lack of ownership information in most databases does not even allow one to identify the parent company in integrated relationships. The literature norm is to assume the principal firms to be either the downstream firms, or the upstream firms. This norm is not just limited to the PRT literature.
a PRT model without principals. The essence of a PRT model is that firms allocate property rights to minimize efficiency loss from contractual frictions. This does not require the existence of a principal firm. Second, I test the predictions of my model using a unique relationship-level database that I compile from S&P Capital IQ. This database contains 85,783 buyer-seller relationships between 25,919 firms from 122 countries. For each relationship, the data identify the buyer, the seller, and whether one side owns the other. This is the first database that allows one to identify the owner in an integrated relationship. My model finds strong and robust support in this database. Finally, to assess the effect of a principal assumption, I test the predictions of a PRT model with principals. This model is supported when buyers are assumed to be the principals, and rejected when sellers are assumed to be the principals. Even when it is supported, the magnitudes of coefficients are much smaller than when using my model without principals. These findings testify the non-falsifiability of PRT models with principals. They also suggest that previous empirical tests of a PRT model with principals may suffer from underestimation.

The key difference between my model and a PRT model with principals lies in the determination of organizational form. In my model, both firms (buyer and seller) in a relationship bargain over three organizational forms: buyer integration (buyer owns seller), seller integration (seller owns buyer), and outsourcing (neither firm owns the other). This means fixing the seller’s industry and the buyer’s industry, there can be two types of integrations in this industry pair. In a PRT model with a principal firm, a predetermined principal firm decides to integrate or outsource the agent firm. This means that given an industry pair, there is only one type of integration – buyer integration (when buyers are assumed to be the principals) or seller integration (when sellers are assumed to be the principals).

My database shows that in many industry pairs, buyer integration and seller integration coexist. There are 2,754 industry pairs in my database, 1,760 of them have integrated relationships. The model with principals predicts that there should be only one type of integration in

---

3Despite the popularity of a principal firm in later PRT models, the original PRT model proposed by Grossman and Hart (1986) does not rely on a principal. Like my model, Grossman and Hart (1986) allow a pair of firms to choose from three organizational forms. The reason why a principal firm is introduced in later models is twofold. First, PRT models are applied to certain situations where there is clearly a principal firm (e.g., the employer-employee situation). Second, and more relevant to this paper, most databases do not contain ownership information. This means in an integrated relationship, one cannot tell who is the owner, and thus one cannot tell the direction of integration in the relationship. A principal assumption removes this problem by assuming the direction of integration.

4In fact, we can add one additional step to a PRT model with principals by allowing the two firms to bargain over who gets to be the principal firm. Then the model generates the same predictions as my model. This is because as long as bargaining is efficient, firms allocate property rights to maximize their joint surplus.
Figure 2.1: Histograms of Industry Pairs by Their Shares of Buyer Integration

Notes: This figure shows the histograms of industry pairs by their shares of buyer integration relative to all integrated relationships. The values on the horizontal axis are calculated as the number of buyer integration relationships divided by the number of integrated relationships in an industry pair. The vertical axis shows the percentage of industry pairs that fall into each bin. There are a total of 2,754 industry pairs. Panel (a) shows the unweighted histogram of industry pairs. Panel (b) shows the weighted histogram of industry pairs, where industry pairs are weighted by their total number of integrated relationships.

Each of these 1,760 industry pairs. In other words, each industry’s share of buyer integration should be 0 (when sellers are the principals) or 1 (when buyers are the principals). However, 774 out of these 1,760 industry pairs have both types of integration relationships. As panel (a) of Figure 2.1 shows, 29% of these industry pairs have only seller integration, 27% of these industry pairs have only buyer integration, and 42% of these industry pairs have both types of integration. When industry pairs are weighted by their numbers of integrated relationships, as panel (b) shows, only 12%(=6%+6%) of all industry pairs have only one type of integration. Note that this has not been verifiable in previous databases because they do not contain ownership information, and thus cannot tell buyer integration from seller integration.

In industry pairs with only one type of integration, the model with principals would not result in estimation bias because integrations are indeed unidirectional as assumed. But if an industry pair contains both types of integrations (which is quite prevalent, as shown in Figure 2.1), the principal status is wrongly assigned for some pairs regardless of which firms are assumed to be the principals. Recall that a PRT model with principals predicts a monotonic relationship between the likelihood of integration, and the relative importance of the principal/agent’s non-contractible investments. If PRT was correct, the pairs for which the principal status is wrongly assigned would create a downward bias in the estimation of this relationship. On the other hand,
my model allows integration to go in either direction. It captures both the industry pairs with only one type of integrations, and those with both types of integrations. This is consistent with what I observe in the data, as shown in Table 2.1.

Table 2.1 summarizes the effects of a 10% increase in the buyer’s relative R&D intensity in different types of PRT models, where the buyer’s relative R&D intensity is a proxy for the relative importance of the buyer firm’s non-contractible investment.\(^5\) In my model without principals, a 10% increase in the relative importance of the buyer’s relative R&D intensity increases a pair’s probability of choosing buyer integration by 2.4%, and decreases its probability of choosing seller integration by 1.5%. In a model with buyers as principals, a 10% increase in the buyer’s relative R&D intensity results in a 0.9% increase in the probability of integration, which fits the theory’s prediction. The magnitude of the coefficient, however, is substantially smaller than my model. In a model with sellers as principals, a 10% increase in the buyer’s relative R&D intensity results in a 0.9% decrease in integration. This is against the prediction of the PRT model with sellers as principals.

The paper is organized as follows. In Section 2.2, I introduce the database and the construction of some key variables. In Section 2.3.1, I construct two PRT models with bilateral integration decisions: one with homogeneous firm pairs and the other with heterogeneous firm pairs (in terms of their levels of productivity). In Section 2.4, I test the predictions of the ho-

---

\(^5\)The specific definition of buyer’s relative R&D intensity is explained in Section 2.2. In addition to R&D intensity, I also include multiple alternatives to proxy for a firm’s non-contractible investment.
mogeneous model, and compare its empirical results with those of a PRT model with unilateral integration decisions. In Section 2.5, I test the predictions of a heterogeneous model. Section 2.6 concludes the paper.

2.2 Data and Descriptive Statistics

2.2.1 Data

The data used in this paper is compiled from S&P Capital IQ.\(^6\) It has two components, a firm database and a relationship database.

The firm database contains industry, country and financial information of 460,499 public and private firms from 122 countries. The financial information covers a firm’s balance sheet data such as its total revenue, employment, R&D expenditures etc., over the period 2005-2012. This database also contains the ownership information of these firms including their (current and prior) parent, ultimate parent, investors, etc. I define a firm’s owner as its parent company. Of the 460,499 firms, 107,496 are subsidiaries, 88,041 are parents, 44,692 are both parent and subsidiary, and 309,654 are independent (have no parent or subsidiary).

The relationship database contains 782,742 buyer-seller relationships between 242,785 firms. Capital IQ collects a firm’s customer and supplier information from various sources including regulatory filings (e.g., 10K, 10Q, 10KSB), news aggregators (e.g., news articles, press release, corporate announcements, bankruptcy reports), and surveys from investment firms. A customer is defined as “a company that receives products or services and that gives business.” A supplier is defined as “a company that provides products or services”. Customer and supplier relationships are complementary in the sense that if firm A is firm B’s customer, then firm B is firm A’s supplier. I code the customer and supplier relationships into a common relationships database.\(^7\) These relationships are sampled during the period 2012-2014, which is predated by the span of their financial information.

I combine the relationship database with the firm database by matching firm names.\(^8\) Keeping those relationships with non-missing financial information results in 66,712 buyer-seller relationships.

---

\(^6\)S&P Capital IQ is a multinational financial information provider. It is a part of McGraw Hill Financial Inc who is also the parent company of Standard & Poor’s Ratings Services, Platts, J.D. Power and Associates, and is the majority owner of the S&P Dow Jones Indices joint venture.

\(^7\)There are many types of customer-supplier relationships besides the one described in the main text, such as borrower-lender, franchisor-franchisee, lessor-lessee, licensor-licensee, landlord-tenant, distributor, etc. The customer-supplier relationship described in the main text constitutes over 85% of all the buyer-seller relationships.
There are an additional 19,071 vertical parent-affiliate relationships that do not appear in the relationship database because there is no information on whether the parent is the customer or supplier. We add these 19,071 pairs to our database. This leaves us with 85,783 (= 66,712 + 19,071) buyer-seller relationships.

A limitation of these 19,071 relationships is that we do not know who is the buyer and who is the seller. I impute this information using a now standard technique involving input-output tables e.g., Acemoglu et al. (2009), Acemoglu et al. (2010), Antràs and Chor (2013), and Alfaro et al. (2015). Specifically, let \( i(p) \) and \( i(a) \) be the industries in which the parent and affiliate operate, respectively. Using the 2002 U.S. input output table, let \( b_{i',i} \) be industry \( i \)'s share of intermediate inputs sourced from sector \( i' \). If \( b_{i',i}(p) > 0 \), then the parent’s industry buys from the affiliate’s industry and I create a buyer-seller observation in which the parent is the buyer. If \( b_{i',i}(a) > 0 \), then the affiliate’s industry buys from the parent’s industry and I create a buyer-seller observation in which the affiliate is the buyer. If \( b_{i',i}(p) = b_{i',i}(a) = 0 \), then the parent-affiliate relationship is not vertical and thus not included in the data. See Appendix A for a more detailed discussion. In the literature just cited, 100% the buyer-supplier information is imputed in this way. I use this imputation for just 22% of the relationships in my database (0.22 = 19,071/85,783). Throwing away these relationships does not change the qualitative results of this paper.

I turn next to data on integration. For each of the 85,783 relationships I know the parent company of both the buyer and the seller. I use this information to categorize each relationship into one of three types: buyer integration, seller integration and outsourcing. If the buyer is the seller’s parent, then the relationship is classified as buyer integration. If the seller is the buyer’s parent, then the relationship is classified as seller integration. If neither is the other’s parent, then the relationship is classified as outsourcing. I exclude those relationships where the buyer and the seller have a common parent because the theoretical framework in this paper does not apply to situations like this, as will be clear in Section 2.3.1.

Table 2.2 summarizes the composition of different types of relationships. The total amount of buyer integration and seller integration are very similar in numbers.

---

8As of now I am using perfect name matches. The algorithm for fuzzy string match is under development. After the algorithm is perfected, we should expect to see a larger relationship database.
Table 2.2: Composition of the Different Types of Relationships

<table>
<thead>
<tr>
<th>Ownership Structure</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buyer integration</td>
<td>11,814</td>
<td>14%</td>
</tr>
<tr>
<td>Seller integration</td>
<td>11,959</td>
<td>14%</td>
</tr>
<tr>
<td>Non-integration</td>
<td>62,010</td>
<td>72%</td>
</tr>
<tr>
<td>Total</td>
<td>85,783</td>
<td>100%</td>
</tr>
</tbody>
</table>

2.2.2 Pros and Cons of this Database

As mentioned in Section 2.1, this database has certain tempting features for conducting a firm boundary study e.g., it contains symmetric information on both sides of a transaction relationship, it contains the ownership information so that it allows one to distinguish between backward and forward integration. However, this database also has its shortcomings and is by no means superior to other databases in every respect. To provide some directions for future improvements on databases, I discuss in this section the pros and cons of this database compared to some of the most frequently used databases.

The databases used by firm boundary studies can be divided into two groups: firm/plant-level databases and transaction-level databases. The former includes the Bureau of Economic Analysis (BEA) annual survey of U.S. Direct Investment Abroad (Antràs et al., 2008), the Dun & Bradstreet WorldBase database (Acemoglu et al. 2009, Alfaro and Charlton 2009, Alfaro and Chen 2014, and Alfaro et al. 2015), UK Annual Respondentata Database (Acemoglu et al., 2010), and French firm-level data (Corcos et al. 2013, Defever and Toubal 2013). The latter includes the U.S. imports and exports database (Bernard et al. 2009a and Nunn and Trefler 2013), the import and export data from the Customs General Administration of China (Feenstra and Hanson 2005 and Li (2013)) and more recently, the U.S. Commodity Flow Survey used by Atalay et al. (2014).

Firm/plant-level databases have detailed information on the firm, but not on its partners. For example, the BEA database contains information on the multinational firm and its subsidiaries, but not on this multinational firm’s outsourced partners. The content of the Dun & Bradstreet database varies across papers, but it mostly contains a firm’s balance sheet data, the industries that a firm operates in, but does not contain information on the firm’s partners. The UK Census of Production data has information on the firm and its plants, but does not list the
firm’s outsourced partners either. However, the BEA data contains information on transaction values between the multinational firm and its subsidiaries, as does the UK Annual Respondents Database. However, the BEA database contains the transaction values between the multinational firm and its subsidiaries, and the UK Annual Respondents Database has information on the input costs and output level of each UK plant. Such information is not available in my data.

Transaction/shipment databases usually contain detailed information on one side of each transaction/shipment, but not on the other side. The U.S. (Chinese) Customs data contains information on the U.S. (Chinese) firm, but not on the other side of the transaction/shipment. The U.S. Commodity Flow Survey identifies the shipping company, but not the receiving company. However, they contain information about the characteristics of the transactions/shipments, such as the price and quantity of the good. Such information is not available in my data.

Last but not least, the above mentioned databases are more comprehensive in the sense that they cover the universe of all U.S. multinationals or all UK manufacturing plants (above a certain size), all shipments by sea or by ground, etc. The relationships database in this paper does not cover all the partners of a firm, nor does it cover all firms around the world. What it captures is a sample of the important relationships of the large companies around the world.

### 2.2.3 R&D and Capital Intensity

In this paper, I use the PRT framework to model firms’ integration decisions. The key prediction of this theory is that firms’ organizational forms depend on their non-contractible, relationship-specific investments. I follow the methodology in Antràs (2003), Nunn and Trefler (2008), Alfaro et al. (2015), etc., and use a firm’s R&D and capital intensity as proxies for the proportion of a firm’s investment that is non-contractible and relationship-specific, since these investments are more subject to hold-up problems that are essential in the PRT approach.

Let $RD_b$ denote the buyer firm’s R&D intensity. It is calculated as

$$RD_b = \frac{\text{Buyer's average R&D expenditure}}{\text{Buyer's average total revenue}},$$

where the buyer’s average R&D expenditure and total revenue are averaged over the period 2005-2012. $RD_s$ is calculated in the same way. A firm’s capital intensity is calculated by the
same equation, replacing R&D expenditure with total capital.

The two firms’ organizational choice depends on the relative importance of their non-contractible, relationship-specific investment. I use the following measure as an indicator of the relative importance of the buyer firm’s non-contractible, relationship-specific investment

$$rd_b = \frac{RD_b}{RD_b + RD_s}.$$ 

By construction, \(rd_s = 1 - rd_b\). I use \(rd_b\) as a sufficient statistic for the relative importance of each firm’s non-contractible, relationship-specific investment. Buyer and seller’s relative capital intensities are calculated in similar ways.\(^9\)

### 2.2.4 Descriptive Statistics

Table 2.3 provides summary statistics for the seller firm, the buyer firm and an average firm (irrespective of whether it is a seller or a buyer) across all three types of relationships: buyer integration, seller integration and non-integration/outsourced relationships.

In buyer integration relationships, seller firms are smaller, younger, and have lower R&D intensity; in seller integration relationships, sellers are larger, older, and have higher R&D intensity; in outsourced relationships, sellers and buyers are similar in size, age, and other measures. Oddly, the same pattern does not apply to capital intensity. This is likely driven by the huge fluctuation in capital measures (the standard deviation of total capital and capital intensity are very high).

In buyer integration relationships, seller firms outnumber buyer firms by twofold, indicating that an average buyer firm has 3.2 sellers. On the other hand, in seller integration relationships, an average seller firm has 3.15 buyers. In outsourced relationships, buyer firms outnumber seller firms, but the difference is not as significant as in integrated relationships. The fact that in each group, the number of firms is greater than the number of buyers or the number of sellers indicates that in each group, there are several buyer firms and seller firms with multiple partners.

---

\(^9\)A firm may have missing entries for its R&D expenditures. If a firm has missing R&D expenditures for the whole period of 2005-2012, this firm’s average R&D expenditure is treated as a missing value; if a firm’s R&D expenditure is missing for some years and positive for the others, then the missing years are treated as 0. For tax reasons, some firms may have negative R&D expenditures for certain years. These values are treated as missing.

\(^{10}\)Buyer and seller may have missing or 0 R&D intensities. If either firm in a pair has missing R&D intensities, this pair will be excluded from the regression sample. If both firms have 0 R&D intensities, the \(rd_b\) will be missing and the pair is also excluded from the regression sample. If one firm has 0 R&D intensity and the other has positive R&D intensity, then \(rd_b\) is either 0 or 1. The pair will be included in the regression sample.
Chapter 2. Global Supply Chains without Principals

Table 2.3: Firm Characteristics by Organizational Forms

<table>
<thead>
<tr>
<th></th>
<th>Buyer Integration Relationships</th>
<th>Seller Integration Relationships</th>
<th>Outsourced Relationships</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Seller</td>
<td>Buyer</td>
<td>Firm</td>
</tr>
</tbody>
</table>
| Total Revenue           | 1,855  | 4,986 | 2,023 | 4,916  | 2,037 | 2,097 | 2,664  | 2,499 | 2,061 | 2,011  | 11410 | (11410)
|                         | 16049  | 11296 |       | 12099  | 11467 |       | 15127  |       |       | (13484) |       |       |
| Employment              | 5,577  | 13,996| 6,104 | 13,896 | 6,104 | 6,367 | 8,089  | 8,107 | 6,881 | 8,011  | 34625 | (33739)
|                         | 40683  | 33739 |       | 35720  | 34090 |       | 33533  |       |       | (30463) |       |       |
| Age                     | 40     | 55    | 42    | 55     | 40    | 43    | 46     | 49    | 47    | 47     | 35    | (35)
|                         | 46     | 35    | (35)  | 46     | 35    | (35)  | 42     | 43    | (42)  |       |       |       |
| R&D expenditure         | 36     | 131   | 36    | 130    | 37    | 36    | 39     | 28    | 22    | 22     | 338   | (318)
|                         | 613    | 318   |       | 346    | 319   |       | 276    |       |       | (245)  |       |       |
| Total Capital           | 1,505  | 4,350 | 1,703 | 4,367  | 1,647 | 1,779 | 2,516  | 2,269 | 1,849 | 1,849  | 11664 | (12279)
|                         | 19437  | 12279 |       | 12195  | 12525 |       | 17547  |       |       | (15732) |       |       |
| R&D intensity           | 0.022  | 0.053 | 0.027 | 0.053  | 0.023 | 0.027 | 0.047  | 0.03  | 0.029 | 0.029  | 0.073 | (0.078)
|                         | 0.096  | 0.073 |       | 0.073  | 0.079 |       | 0.09   |       |       | (0.091) |       |       |
| Capital Intensity       | 4.32   | 1.04  | 3.76  | 1.1    | 4.37  | 3.81  | 3.6    | 18.19 | 15.67 | 15.67  | 354   | (322)
|                         | 2       | 322   |       | 4      | 356   |       | 1925   |       |       | (1712)  |       |       |
| Observations            | 9,504  | 2,969 | 11,516| 2,984  | 9,404 | 11,396| 10,838 | 15,301| 19,407| 19,407 | 5,004 | (1116)
|                         | 2,969  | 11,516|       | 9,404  | 11,396|       | 15,301 |       |       | (19,407) |       |       |

Notes: This table lists the firm characteristic across three types of organizational forms. The first three columns summarize the buyer, seller and average firm characteristics for all buyer integration relationships. The next three columns report the buyer, seller and average firm characteristics for all seller integration relationships. The last three columns report the buyer, seller and average firm characteristics for all outsourced relationships. Numbers in parentheses report the standard deviations of the corresponding variables. Total revenue, R&D expenditure and total capital are averaged over the period 2005-2012, measured in millions of U.S. dollars. Employment is the number of employees for the year 2014. Age is the firm’s age since it was founded until the year 2014. R&D and capital intensity are as defined in the previous section.

2.3 Model

In this section, I introduce two property rights models with bilateral integration decisions: one with homogeneous pair productivity and the other with heterogeneous pair productivity. I refer to the first model as the homogeneous model and the second one as the heterogeneous model. The homogeneous model provides a direct comparison between a bilateral integration model and a unilateral integration model. The heterogeneous model allows me to stretch my examination of a property rights model. It also lays the groundwork for my future research.

The models in this paper are based on Antrás (2003) and Antrás and Helpman (2004), with two important differences. First, in my models, a buyer firm and a seller firm collectively choose from buyer integration, seller integration and outsourcing. In Antrás (2003), Antrás and Helpman (2004), and many other papers on firm boundary decisions (Acemoglu et al., 2009, 2010), there is a principal firm who chooses between integration and outsourcing. As I will
later explain, the essential difference is not the existence of a principal firm, but the choice set (three options versus two options). Second, in Antrás and Helpman (2004), firms with different productivity levels make heterogeneous organizational choices because of an assumption on fixed costs. More specifically, they assume that the fixed cost of integration is higher than the fixed cost of outsourcing. In my heterogeneous model, firms draw i.i.d fixed costs from a type I extreme value distribution. I show that the heterogeneous organizational forms do not have to be driven by an ad hoc assumption on the ranking of fixed costs.

I do not extend the models in this paper to general equilibrium because a general equilibrium model requires restricting the interpretation of a key parameter in this model, which is $\eta$, the relative importance of the buyer firm’s non-contractible, relationship-specific investment. Previous literature assume that $\eta$ is an industry characteristic (Antrás, 2003; Antrás and Helpman, 2004; Antrás et al., 2008; Acemoglu et al., 2010). I allow $\eta$ to be either industry-specific or pair-specific, and construct separate empirical measures for each. In a general equilibrium one needs to assume that $\eta$ is either industry-specific or pair-specific. I reserve this for another paper.

### 2.3.1 The Homogeneous Model

#### Setup

A closed economy is resided by a unit measure of consumers with CES preference

$$U = q_0 + \frac{1}{\omega} \int_{j=1}^{J} Q_j^\omega dj, \quad 0 < \omega < 1,$$

where $1/(1 - \omega)$ is the representative consumer’s cross-industry elasticity of substitution, $q_0$ is the consumption of a homogeneous good, and $Q_j$ is a CES aggregate of all varieties in industry $j$\footnote{Since this is a closed economy where consumption always equals production, I use the same letters for consumption and production levels.}

$$Q_j = \left( \int_{i=1}^{N_j} q_j(i)^\alpha di \right)^{1/\alpha}, \quad 0 < \alpha < 1.$$

$N_j$ is the number of varieties in industry $j$, which is endogenous in a general equilibrium but is taken as exogenous in a partial equilibrium. $q_j(i)$ is the consumption level of variety $i$ in industry $j$, and $\sigma \equiv 1/(1 - \alpha) > 1$ is this consumer’s within-industry elasticity of substitution, which is a constant for all industries.
The producer of variety $i$ in industry $j$ faces an inverse demand function

$$p_j(i) = Q_j^\omega - \alpha q_j(i)^{\alpha - 1}.$$ 

To produce $q_j(i)$, the producer (henceforth the buyer) needs to collaborate with a supplier (henceforth the seller).\textsuperscript{12} The buyer firm is located in industry $j$ and the seller firm is located in a related industry which does not produce any final good.\textsuperscript{13} Let $x_{j,b}(i)$ denote the buyer’s investment and $x_{j,s}(i)$ denote the seller’s investment. The marginal cost of investment is $c_{j,b}(i)$ for the buyer and $c_{j,s}(i)$ for the seller. Both firms take their marginal costs as exogenous.

The final good is produced by the Cobb-Douglas production function

$$q_j(i) = \left(\frac{x_{j,b}(i)}{\eta_j(i)}\right)^{\eta_j(i)} \left(\frac{x_{j,s}(i)}{1 - \eta_j(i)}\right)^{1 - \eta_j(i)},$$

where $\eta_j(i) \in (0,1)$ represents the relative importance of the buyer’s investment. $\eta_j(i)$ can be pair-specific or industry specific. When $\eta_j(i)$ is industry-specific, i.e., $\eta_j(i) = \eta_j$, $\forall i$. Previous literature treats $\eta_j(i)$ as industry-specific (Antràs, 2003; Antràs and Helpman, 2004; Antràs et al., 2008).

The timing of the game is as follows:

1. Buyer firms and seller firms form one-to-one matches.

2. Each firm pair draws a set of fixed costs $\{f_{j,BI}(i), f_{j,SI}(i), f_{j,NI}(i)\}$, where $f_{j,k}(i) = f_j + \varepsilon_{j,k}(i)$, $k \in \{BI, SI, NI\}$. $\varepsilon_{j,k}(i)$ is an error term which follows a type I extreme value distribution

$$\varepsilon_{j,k}(i) \sim Gumbel(0, \sigma_j).$$

Note that the Gumbel distribution function is independent of $k$.

3. The buyer firm and the seller firm in each pair collectively determine their ownership structure using Nash bargaining. The bargaining weights are $\gamma_j(i)$ for the buyer and $[1 - \gamma_j(i)]$ for the seller ($0 < \gamma_j(i) < 1$). They both have 0 outside options at this stage. The bargaining

\textsuperscript{12}Although I refer to the two firms as buyer and seller, this collaboration relationship does not necessarily require the transfer of a physical good. Recent work by Alfaro et al. (2015) and Atalay et al. (2014) show that integrated parties often do not transfer any physical goods. This is not contradictory with the setup in this paper. Grossman and Hart (1986) did not dictate there to be transfer of physical goods either.

\textsuperscript{13}That the seller’s industry does not produce any final good is a simplifying assumption. Alternatively, one can relax this assumption to allow sellers to produce final goods. In this case, a close resemblance to the buyer-seller relationships in this model is an input-output table, or the network structure as modeled in Acemoglu et al. (2012).
results in them choosing an ownership structure \( k \in \{BI, SI, NI\} \), where \( BI, SI \) and \( NI \) respectively stand for buyer integration, seller integration and non-integration/outsourcing, and a transfer payment \( t_j(i) \in \mathbb{R} \) from the buyer to the seller (\( t_j(i) < 0 \) is a transfer from the seller to the buyer).

4. After \( k \) is chosen and \( t_j(i) \) is paid, the buyer and the seller simultaneously choose their investment levels \( x_{j,b}(i) \) and \( x_{j,s}(i) \). The buyer firms and the seller firms face constant marginal costs of investment. The buyer’s marginal cost is \( c_{j,b} \), and the seller firm’s marginal cost is \( c_{j,s} \).

5. After investments are made, the two firms bargain over the division of their future revenue. Suppose the bargaining weights are \( \beta_j \) for the buyer and \( [1 - \beta_j] \) for the seller. \( \beta_j \in (0, 1) \).

6. The product is produced and sold on the market. The two firms split the total revenue according to the agreement reached in step 5.

In this homogeneous model, within each industry \( j \), all firm pairs face identical production functions. They also face the same marginal costs of investments \( c_{j,b} \) and \( c_{j,s} \). As mentioned before, I do not solve for the general equilibrium. To avoid unnecessary notation complexity, ignore the \( i, j \) indexes for now. The inverse demand function can be written as

\[
p = q^{\alpha - 1}, \tag{2.1}
\]

and production function as

\[
q = \left( \frac{x_b}{\eta} \right)^{\eta} \left( \frac{x_s}{1 - \eta} \right)^{1-\eta}. \tag{2.2}
\]

Combining the above two equations generates the revenue function

\[
R = q^\alpha = \left( \frac{x_b}{\eta} \right)^{\alpha \eta} \left( \frac{x_s}{1 - \eta} \right)^{\alpha (1-\eta)}. \tag{2.3}
\]

Under this simplification, the investment levels of the buyer and the seller can be written as \( x_b \) and \( x_s \). Assume that \( x_b \) and \( x_s \) are relationship-specific, so that they are of 0 values to the buyer and the seller outside this relationship. Also, assume that they are non-contractible, so that the firms cannot write an ex-ante contract on the levels of \( x_b \) and \( x_s \). The non-contractible feature of these investments imply that there will be renegotiation between the buyer and the seller after investments are made.

Although they cannot write a contract on their investment levels, the buyer and the seller
can write an ex ante contract on the ownership over these investments. As defined in Grossman and Hart (1986), ownership is a collection of residual control rights, rights that are not specified in the contract. It comes into play during the renegotiation stage, after investments $x_b$ and $x_s$ are made. If the buyer owns both parties’ investments, in the event of disagreement, the buyer seizes $x_b$ and $x_s$, leaving the seller with nothing. The buyer can only use $x_s$ inefficiently, which following Antràs (2003) is modeled as the buyer is only able to use a fraction $\delta^{1/(1-\eta)}$ of $x_s$.\footnote{One can think of $\delta$ as a measure of the court’s enforcement power. The higher $\delta$ is, the more investment the owner can retrieve, and the higher outside option the ownership delivers.}

According to the production function, this generates an output of $\delta q$ and a revenue of $\delta^\alpha R$. The seller on the other hand, gets a revenue of 0. Similarly, if the seller owns both $x_b$ and $x_s$, in the event of disagreement she seizes $x_b$ and $x_s$, suffers an efficiency loss of $(1 - \delta^{1/\eta})x_b$ when using $x_b$, produces an output of $\delta q$ and generates a revenue of $\delta^\alpha R$. The buyer gets 0 revenue in this event. A case in between is where they each own their own investments. In the event of disagreement, they both retrieve their own investments. With only one input for the production function, they both produce 0 output and hence generate 0 revenue. I refer to each of these three ownership structures as buyer integration (BI), seller integration (SI), and outsourcing (NI).\footnote{There is one additional ownership beside buyer integration, seller integration and non-integration, which is cross ownership, where the buyer owns the seller’s investment and the seller owns the buyer’s investment. This type of ownership structure does not seem to exist in reality, and it generates identical results as non-integration, and is thus left out of the discussion. The reason this generates identical results as the non-integration case is because property rights are assumed to be indivisible. When partial ownership over a firm is allowed, cross ownership, where the two firms are allowed to control only part of the other firm’s asset can allow the firms to achieve the first-best joint surplus. This is because under cross-ownership, the two firms can perfectly control each-other’s ex post share of total revenue, and hence their ex ante investment incentives. Allowing for the existence of cross partial ownership may be an interesting extension, so far there is no work modeling partial ownership in a property rights framework.}

The equilibrium of this game can be solved using backward induction.

In step 6, the two firms simply produce the product, sell it, and divide the revenue. The output function and revenue function are as specified in equations (2.2) and (2.3).

In step 5, the two firms bargain over the division of revenue, taking $k$, $x_b$ and $x_s$ as given. As discussed before, if $k = BI$, the buyer owns $x_b$ and $x_s$. His outside option is $\delta^\alpha R$. $R$ is as defined in equation (2.3). The seller’s outside option is 0. The pair’s Nash surplus is their joint revenue minus their outside options: $R - \delta^\alpha R - 0 = (1 - \delta^\alpha)R$. The buyer’s Nash surplus is simply his outside option plus a share $\beta$ of the total surplus, that is, $\delta^\alpha R + \beta(1 - \delta^\alpha)R = [\beta + (1 - \beta)\delta^\alpha]R$, so the buyer’s share of total revenue under $k = BI$ is $\beta + (1 - \beta)\delta^\alpha$. Similarly, the seller’s Nash surplus is $0 + (1 - \beta)(1 - \delta^\alpha)R(i) = (1 - \beta)(1 - \delta^\alpha)R$. Her share of total revenue is
Global Supply Chains without Principals

The two firms’ revenue shares under \( k = SI \) and \( k = NI \) can be derived using the same logic. Let \( \beta^k \) represent the buyer’s share of total revenue under ownership structure \( k \). The seller’s share of total revenue is then \( (1 - \beta^k) \). Their revenue shares under different ownership structures are summarized in Table 2.4.

**Table 2.4: Buyer and Seller’s Revenue Shares under Different Ownership Structures**

<table>
<thead>
<tr>
<th></th>
<th>( k = BI )</th>
<th>( k = SI )</th>
<th>( k = NI )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buyer’s revenue share (( \beta^k ))</td>
<td>( \beta + (1 - \beta)\delta^\alpha )</td>
<td>( \beta(1 - \delta^\alpha) )</td>
<td>( \beta )</td>
</tr>
<tr>
<td>Seller’s revenue share (( 1 - \beta^k ))</td>
<td>( (1 - \beta)(1 - \delta^\alpha) )</td>
<td>( 1 - \beta + \beta\delta^\alpha )</td>
<td>( 1 - \beta )</td>
</tr>
</tbody>
</table>

It is clear from Table 2.4 that \( \beta^{BI} > \beta^{NI} > \beta^{SI} \): buyer has the highest revenue share under \( k = BI \) and the lowest revenue share under \( k = SI \). On contrary, \( 1 - \beta^{BI} < 1 - \beta^{NI} < 1 - \beta^{SI} \): seller has the highest revenue share under \( k = SI \) and the lowest revenue share under \( k = BI \). This is driven by their ranking of outside options across the three ownership structures.

In step 4, the two firms choose their investment levels taking \( \beta^k \) and \( (1 - \beta^k) \) as given. Marginal costs of investments are \( c_b \) for the buyer and \( c_s \) for the seller. They are also taken as exogenous. The buyer chooses \( x_b \) to maximize his total surplus at step 4. His problem can be written as

\[
\max_{x_b} \beta^k \left( \frac{x_b}{\eta} \right)^{\alpha\eta} \left( \frac{x_s}{1 - \eta} \right)^{\alpha(1-\eta)} - c_b x_b, \tag{2.4}
\]

Similarly, the seller’s problem at step 4 can be written as

\[
\max_{x_s} (1 - \beta^k) \left( \frac{x_b}{\eta} \right)^{\alpha\eta} \left( \frac{x_s}{1 - \eta} \right)^{\alpha(1-\eta)} - c_s x_s. \tag{2.5}
\]

The equilibrium levels of \( x_b \) and \( x_s \) simultaneously solve the buyer’s problem and the seller’s problem. The buyer’s investment level can be expressed as

\[
x_b = \left\{ \alpha \left( \frac{\beta^k \eta}{c_b} \right)^{1-\alpha + \alpha\eta} \left[ \frac{(1 - \beta^k)(1 - \eta)}{c_s} \right]^{\alpha - \alpha\eta} \right\}^{1/(1-\alpha)}, \tag{2.6}
\]

and the seller’s investment level can be expressed as

\[
x_s^{1-\alpha} = \left\{ \alpha \left( \frac{\beta^k \eta}{c_b} \right)^{\alpha\eta} \left[ \frac{(1 - \beta^k)(1 - \eta)}{c_s} \right]^{1-\alpha\eta} \right\}^{1/(1-\alpha)}. \tag{2.7}
\]
Substituting the solutions to $x_b$ and $x_s$ back into the buyer and the seller’s problems, we can get their surpluses at step 4

$$
\pi_b(\beta^k, \eta) = \frac{(1 - \alpha \eta) \beta^k}{\left(\frac{1}{\alpha}\right)^{\eta} \left(\frac{c_b}{1 - \beta^k}\right)^{1-\eta}} - f_b^k, \quad (2.8)
$$

and

$$
\pi_s(\beta^k, \eta) = \frac{(1 - \alpha + \alpha \eta)(1 - \beta^k)}{\left(\frac{1}{\alpha}\right)^{\eta} \left(\frac{c_s}{1 - \beta^k}\right)^{1-\eta}} - f_s^k, \quad (2.9)
$$

where $f_b^k$ is the portion of fixed cost $f^k$ that is paid by the buyer firm, and $f_s^k$ the proportion of fixed cost paid by the seller. $f_b^k + f_s^k = f^k$.

Equations (2.8) (2.9) and (2.11) imply that the buyer and the seller’s surpluses from step 4 can also be expressed as fractions of their joint surplus $\pi_k(\eta)$. Let $\mu^k(\eta)$ be the buyer’s fraction and let $(1 - \mu^k(\eta))$ be the seller’s fraction, then

$$
\mu^k(\eta) = \frac{(1 - \alpha \eta) \beta^k}{1 - \alpha(\beta^k \eta + (1 - \beta^k)(1 - \eta))}; \quad 1 - \mu^k(\eta) = \frac{(1 - \beta^k)(1 - \eta)}{1 - \alpha(\beta^k \eta + (1 - \beta^k)(1 - \eta))}. \quad (2.10)
$$

The pair’s joint surplus at step 4 is simply $\pi(\beta^k, \eta) \equiv \pi_b^k(\eta) + \pi_s^k(\eta)$, which is

$$
\pi(\beta^k, \eta) = \psi(\beta^k, \eta) - f^k, \quad (2.11)
$$

where

$$
\psi(\beta^k, \eta) = \frac{1 - \alpha(\beta^k \eta + (1 - \beta^k)(1 - \eta))}{\left\{\left(\frac{1}{\alpha}\right)^{\eta} \left(\frac{c_b}{1 - \beta^k}\right)^{1-\eta}\right\}^{\frac{\alpha}{1-\alpha}}} \quad (2.12)
$$

$\psi(\beta^k, \eta)$ is pair $i$’s variable profit under ownership $k$. $f^k$ is the pair’s fixed cost associated with ownership $k$.

In step 3, the optimal ownership structure $k^*$ is chosen to maximize the pair’s joint surplus $\pi(\beta^k, \eta)$.

$$
k^* = \arg \max_{k \in \{BI, SI, NI\}} \pi(\beta^k, \eta). \quad (2.13)
$$

The two components in the profit function $\pi^k(\eta)$ are $\psi^k(\eta)$ and $f^k$. The following lemmas

\[\text{16}\text{The reason why the two firms' choose a } k \text{ to maximize their joint surplus is because any organizational form that is not profit maximizing is not Pareto optimal and is thus unstable.}\]
summarize some properties of the variable profit function $\psi(\beta^k, \eta)$.

**Lemma 3.** The variable profit function $\psi(\beta^k, \eta)$ is supermodular in $\beta^k$ and $\eta$ for $\beta^k \in (0, 1)$ and $\eta \in (0, 1)$.

The intuition behind Lemma 3 is simple. It states that given a higher $\eta$, a higher $\beta^k$ generates a higher variable profit. This is because a higher $\eta$ indicates that the buyer firms’ investment is relatively more important, in which case a higher ex post revenue share for the buyer provides him with a higher ex ante investment incentive.

In fact, if firms were allowed to freely choose a revenue share $\beta$ to maximize their variable profit, their choice of $\beta$ can be solved as a function of $\eta$

$$
\beta^*(\eta) = \eta(\alpha \eta + 1 - \alpha) - \sqrt{\eta(1 - \eta)(1 - \alpha \eta)(\alpha \eta + 1 - \alpha)} \over 2\eta - 1,
$$

where $\beta^*(\eta)$ is strictly increasing in $\eta$.

In this model, firms are not allowed to freely choose their revenue shares. Instead, they can only choose from three discrete values of $\beta^k$: $\beta^{BI}$, $\beta^{SI}$ and $\beta^{NI}$. Given the ranking of these three values ($\beta^{BI} > \beta^{NI} > \beta^{SI}$) and the supermodularity of $\psi(\beta^k, \eta)$, $\beta^{BI}$ generates a higher variable profit when $\eta$ is high, while $\beta^{SI}$ generates the highest revenue when $\eta$ is low. When $\eta$ is in between, $\beta^{NI}$ generates the highest revenue share. This is summarized in the following lemma.

**Lemma 4.** There are two threshold values of $\eta$, $\underline{\eta}$ and $\overline{\eta}$, with $0 < \underline{\eta} < \eta < \overline{\eta} < 1$, such that the following statements hold.

1. When $\eta > \overline{\eta}$, $\psi(\beta^{BI}, \eta) > \psi(\beta^{NI}, \eta) > \psi(\beta^{SI}, \eta)$.
2. When $\underline{\eta} < \eta < \overline{\eta}$, $\psi(\beta^{NI}, \eta) > \max\{\psi(\beta^{BI}, \eta), \psi(\beta^{SI}, \eta)\}$.
3. When $\underline{\eta} < \eta$, $\psi(\beta^{SI}, \eta) > \psi(\beta^{NI}, \eta) > \psi(\beta^{BI}, \eta)$.

$\underline{\eta}$ and $\overline{\eta}$ are implicitly solved by $\psi(\beta^{SI}, \underline{\eta}) = \psi(\beta^{NI}, \eta)$ and $\psi(\beta^{BI}, \overline{\eta}) = \psi(\beta^{NI}, \overline{\eta})$.

Without fixed costs, $\pi(\beta^k, \eta) = \psi(\beta^k, \eta)$. The ranking of profit functions follows the ranking of variable profit functions in Lemma 4. Given $\eta$, firm pairs choose the same $k$. There is no heterogeneity in firms’ ownership decisions. In Antràs and Helpman (2004), heterogeneous ownership decisions are driven by two assumptions, (1) heterogeneous productivity, and (2) the fixed cost of vertical integration is higher than the fixed cost of outsourcing. I show that a stochastic fixed costs assumption alone can drive heterogeneous ownership decisions.
With stochastic fixed costs, given $\beta^k$ and $\eta$, two pairs of firms have the same level of variable profit. But they may have drawn different levels of fixed costs. Suppose $k = BI$ generates the highest variable profit for both firms: One pair of firms may find it too costly to choose buyer integration because its fixed cost for buyer integration is too high.

Recall that ignoring the industry and firm indexes, the fixed cost of pair $i$ in industry $j$ can be written as

$$f^k = f + \varepsilon^k,$$

where $\varepsilon^k$ follows i.i.d. type I extreme value (Gumbel) distribution.

By McFadden (1973), the probability of $k$ being the ownership structure that generates the highest profit for pair $i$ can be written as

$$\Pr\{k^* = k\} = \frac{\exp\{\psi(\beta^k, \eta)\}}{\sum_{k' \in \{BI, SI, NI\}} \exp\{\psi(\beta^{k'}, \eta)\}}.$$ 

Although the stochastic nature of fixed costs prevents us from a deterministic description of firm-pair’s ownership decisions, the distribution of the error term in the fixed costs allows us to describe the probabilities of a firm-pair’s choosing the various ownership structures. This is summarized by Theorem 5 and Corollary 1.

**Theorem 5.** $\Pr\{k^* = BI\}$ is increasing in $\eta$; $\Pr\{k^* = SI\}$ is decreasing in $\eta$.

Theorem 5 states that the pair’s probability of choosing buyer integration decreases in $\eta$, and its probability of choosing seller integration increases in $\eta$. As stated before, when $\eta$ is high, buyer’s investment becomes more important. A higher $\beta^k$ provides the buyer with a higher ex post revenue share and thus a higher ex ante investment incentive, so when $\eta$ gets higher, buyer integration becomes more attractive to the pair of firms. A pair of firms diverts from choosing buyer integration only when their fixed cost associated with buyer integration is too high. The higher $\eta$ is, the higher fixed cost that is need for the pair of firms to deviate from choosing buyer integration. Thus the higher their probability of choosing buyer integration. In contrast, when $\eta$ is low, seller integration becomes more attractive. The lower $\eta$ is, the larger $f^{SI}$ is needed for a pair of firms to deviate. Not only does this pattern apply to the absolute probabilities, but it also applies to odds ratios, as shown in Corollary 1.

**Corollary 1.** $\Pr\{k^* = BI\}/\Pr\{k^* = NI\}$ is increasing in $\eta$; $\Pr\{k^* = SI\}/\Pr\{k^* = NI\}$ is decreasing in $\eta$. 
The last step to solving this game is \( t^k \), the transfer payment from the buyer to the seller under ownership structure \( k \). When \( t^k > 0 \), the buyer pays the seller; when \( t^k < 0 \), the seller pays the buyer. Intuitively, one can think of \( t^k \) as an entry fee to enter the production process. When the paying party is financially constrained, the two firms may not be able to initiate this production process or they may have to switch to a sub-optimal ownership structure if they both strictly prefer producing to quitting. I assume that the two firms’ outside options in step 3 are both 0, so they would both prefer to produce *something* because their surpluses from producing are always strictly positive. Theoretically, this transfer payment serves the role of changing the two firms’ shares of the joint surplus from step 4 to their shares of the joint surplus in step 3. Note that the joint surplus does not change from step 4 to step 3. With 0 outside options, the buyer and the seller’s shares of joint surplus at step 3 are respectively \( \gamma \) and \( (1 - \gamma) \). Recall that the buyer and the seller’s shares of joint surplus at step 4 are respectively \( \mu^k \) and \( 1 - \mu^k \). The transfer payment from the buyer to the seller is simply

\[
 t^k \equiv \mu^k \pi^k - \gamma \pi^k = (\mu^k - \gamma)\pi^k
\]

More specifically,

\[
 t^k = \frac{(1 - \alpha \eta)\beta^k - \gamma\{1 - \alpha[(\beta^k \eta + (1 - \beta^k)(1 - \eta)]\}}\{(1/\alpha)(c_b/\beta^k)^\eta(c_s/(1 - \beta^k))^{1-\eta}\alpha/(1-\alpha)\}.
\]

\( t^k > 0 \) if and only if \( \mu^k > \gamma \). It is shown in the appendix that \( \mu^k \) is strictly increasing in \( \beta^k \) and strictly decreasing in \( \eta \). Given ownership structure \( k \), if \( \eta \) is high, buyer’s investment is more important. The pair chooses an ownership structure that allocates a higher ex-post revenue share \( (\beta^k_b) \) to the buyer. This also generates a higher share of joint surplus \( (\mu^k) \) for the buyer. However, the buyer’s ex-ante share of total surplus \( (\gamma) \) does not depend on \( \eta \). When the buyer’s ex-post surplus \( (\mu^k \pi^k) \) exceeds his ex-ante surplus \( (\gamma \pi^k) \), the buyer has to pay the seller, meaning \( t^k > 0 \) in order for both of them to get their ex-ante surpluses. In one sentence – \( t^k \) is positive when \( \eta \) is high and negative when \( \eta \) is low. The threshold value of \( \eta \), \( \eta^k_m \) is given by the following equation:

\[
 \eta^k_m = \frac{\beta^k - [1 - \alpha(1 - \beta^k)]\gamma}{\alpha[\beta^k(1 - \gamma) + (1 - \beta^k)\gamma]}.
\]

For a given ownership structure \( k \), \( t^k > 0 \) if \( \eta < \eta^k_m \) and \( t^k < 0 \) if \( \eta > \eta^k_m \). If \( \eta^k_m < 0 \) then \( t^k > 0 \) i.e., the seller always pays the buyer. If \( \eta^k_m > 1 \) then the buyer always pays the seller.
For the purpose of this paper, a partial equilibrium does not require me to analyze the matching process, or specify the distribution function of the firms’ marginal costs, so I do not solve for steps 1 and 2.

### 2.3.2 The Heterogeneous Model

In the homogeneous model, all firm pairs have the same levels of marginal costs $c_b$ and $c_s$. In this section, I allow $c_b$ and $c_s$ to vary across firm pairs. Recall the profit function

Let $\theta_j(i)$ denote the productivity of the firm pairs producing variety $i$ in industry $j$. The production function for variety $i$ in industry $j$ is

$$q_j(i) = \theta_j(i) \left(\eta_{j,b}(i) \frac{c_{i,b}}{\eta_j(i)} \eta_{j,s}(i) \frac{c_{i,s}}{1-\eta_j(i)} \right)^{1-\eta_j(i)},$$

where $\eta_j(i)$ is either pair-specific or industry-specific. Again, since I do not solve for the general equilibrium of this model, to avoid unnecessary notation complexities, I write the production function as

$$q = \theta\left(\frac{x_b}{\eta} \left(\frac{x_s}{1-\eta}\right)^{1-\eta}\right).$$

(2.16)

From step 3 onward, the game is identical to the homogeneous model. The solution to these steps closely resemble that in the homogeneous model.

In step 4, backward induction generates the pair’s profit function

$$\pi^k(\theta, \eta) = \theta^{\alpha/(1-\alpha)} \psi^k(\eta) - f^k,$$

(2.17)

where

$$\psi^k(\eta) = \frac{1-\alpha}{\alpha} \left(1-\alpha\right) \eta \left(1-\eta\right)^{1-\eta},$$

and $f^k$ is the fixed cost associated with ownership structure $k$. $\psi^k(\eta)$ is the profit function of the homogeneous model (see equation (2.11)). In the heterogeneous model, $\psi^k(\eta)$ becomes the slope of $\pi^k$ as a function of $\theta^{\alpha/(1-\alpha)}$. By Theorem 5, there are two threshold values of $\eta$, $\eta$ and $\eta$, such that when $\eta < \eta$, $\psi^{SI}(\eta) > \psi^{NI}(\eta) > \psi^{BI}(\eta)$; when $\eta > \eta$, $\psi^{BI}(\eta) > \psi^{NI}(\eta) > \psi^{SI}(\eta)$; when $\eta < \eta < \eta$, $\psi^{NI}(\eta) > \max\{\psi^{SI}(\eta), \psi^{BI}(\eta)\}$. Figure 2.2 plots these three scenarios fixing $f^k$ at $f$.

It can be seen from Figure 2.2 that if $f = 0$, the prediction of a heterogeneous model
is very similar to that of the homogeneous model. To generate heterogeneous organizational forms across firms with different productivity levels, previous literature has assumed that $f^k$ is different for different choices of $k$ (Antràs and Helpman, 2004, 2008). More specifically, it assumes that the fixed cost of integration $f^V$ is greater than the fixed cost of outsourcing $f^O$. I adopt a different setup by allowing the firm pairs to each draw their own sets of i.i.d. fixed
costs \((f^{BI}, f^{SI}, f^{NI})\) from a Gumbel distribution.

In step 3, the buyer and the seller choose \(k\) to maximize their profit \(\pi^k\)

\[
k = \arg \max_{k \in \{BI, SI, NI\}} \pi^k, \tag{2.18}
\]

where \(\pi^k\) is as defined in equation 2.17. The solution to this problem is different from the homogeneous model in that the optimal \(k^*\) is no longer a deterministic function of \(\eta\), as firm pairs with the same value of \(\eta\) and \(\theta\) may have different levels of fixed costs, a pair with a lower \(f^{SI}\) is more likely to choose \(k = SI\), whereas a firm pair with a lower \(f^{BI}\) is more likely to choose \(k = BI\). However, the probabilities of a firm’s choice for these three organizational forms \(BI\), \(SI\) and \(NI\) are monotone functions in \(\theta\) and \(\eta\). Let \(\Pr\{k^* = BI|\theta, \eta\}\), \(\Pr\{k^* = SI|\theta, \eta\}\) and \(\Pr\{k^* = NI|\theta, \eta\}\) respectively denote a pair’s probability of choosing buyer integration, seller integration and non-integration conditional on its productivity level \((\theta)\) and the relative importance of buyer’s non-contractible, relationship-specific investment \((\eta)\).

**Theorem 6.** There are two threshold values of \(\eta\), \(\underline{\eta}\) and \(\bar{\eta}\), with \(0 < \underline{\eta} < \eta < \bar{\eta} < 1\), such that

1. When \(\eta < \underline{\eta}\), \(\Pr\{k^* = BI|\theta, \eta\}\) is increasing in \(\theta\) and \(\Pr\{k^* = SI|\theta, \eta\}\) is decreasing in \(\theta\).
2. When \(\eta > \bar{\eta}\), \(\Pr\{k^* = SI|\theta, \eta\}\) is increasing in \(\theta\) and \(\Pr\{k^* = BI|\theta, \eta\}\) is decreasing in \(\theta\).
3. When \(\underline{\eta} < \eta < \bar{\eta}\), \(\Pr\{k^* = NI|\theta, \eta\}\) is increasing in \(\theta\).

As Figure 2.2a shows, when \(\eta < \underline{\eta}\), without uncertainty, all producing firm pairs choose seller integration. However, if a pair draws a very big \(f^{SI}\) and a very small \(f^{BI}\), this pair may decide to choose buyer integration instead. To revert from seller integration to buyer integration, a less productive pair needs a smaller shock than a more productive pair. Similarly, if a pair receives a very big \(f^{BI}\) and a very small \(f^{NI}\), this pair may find non-integration more profitable. To revert from seller integration to non-integration, a less productive pair needs a smaller shock than a more productive pair. Therefore, although all pairs are more likely to choose seller integration when \(\eta < \underline{\eta}\), less productive pairs are more likely to choose otherwise than more productive pairs. In other words, the probability of choosing seller integration is increasing in \(\theta\).

On the other hand, when \(\eta > \bar{\eta}\), without uncertainty, all producing pairs choose buyer integration. To deviate from this choice, a less productive pair needs a smaller shock than a
more productive pair. Therefore, although all pairs are more likely to choose buyer integration when $\eta > \bar{\eta}$, less productive firms are more likely to choose otherwise than more productive pairs. In other words, the probability of choosing buyer integration is increasing in $\theta$.

Lastly, when $\eta < \eta < \bar{\eta}$, all pairs choose non-integration without uncertainty. To deviate from this choice, a less productive firm needs a smaller shock than a more productive pair. In other words, the probability of choosing non-integration is increasing in $\theta$.

2.3.3 Comparison to a Principal-Agent Framework

As mentioned in the beginning of this section, one importance difference between the model in this paper and other papers on firm boundary decisions is that instead of a principal firm who chooses between integration and outsourcing, the two firms collectively choose their ownership structure from buyer integration, seller integration, and outsourcing. The key difference lies in the number of options rather than there being a principal firm.

The key prediction of the Grossman-Hart-Moore framework is that property rights are allocated to minimize efficiency loss. It does not dictate whether there should be one firm or multiple firms making this allocation decision. When there are two firms involved in this relationship, there are three possible allocations (assuming that property rights are indivisible): seller owns buyer, buyer owns seller, and none of them owns the other. The organizational form with the lowest efficiency loss should be chosen, no matter who is making the decision.

In fact, it is easy to transform the model in this paper to incorporate a principal agent framework. One needs only to change the first two steps in the homogeneous model:

1. The buyer and the seller bargain over who becomes the principal. The buyer’s bargaining weight is $\gamma$ and the seller $(1 - \gamma)$. $0 < \gamma < 1$. The transfer payment from the buyer to the seller in this step is $t \in R$.

2. After the two firms have determined who is the principal, the principal firm then decides to either integrate or outsource the agent firm. Denote his choice by $k \in \{I, O\}$. Together with this choice of organizational form, the principal firm pays the agent firm an upfront payment of $\tau \in R$.

The only difference between the timing here and that in Anträss and Helpman (2004) is that there is one additional step before the game. Instead of starting with the principal firm choosing an organizational form, there is a step where the principal firm is determined. In this new setup, there are 4 types of equilibria: buyer becomes the principal and chooses integration ($BI$), buyer
becomes the principal and chooses outsourcing \((BO)\), seller becomes the principal and chooses integration \((SI)\) and seller becomes the principal and chooses outsourcing \((SO)\). \(BI\) and \(SI\) are actually equivalent to those in Theorem 5. \(BO\) and \(SO\) combined corresponds to \(NI\) in Theorem 5.

The reason why these two frameworks generate very similar conclusions is because of efficient bargaining. When bargaining is efficient, the allocation structure with minimal efficiency loss is always chosen. I do not use a principal-agent framework in this paper because the database naturally allows one to identify three types of organizational forms \((BI, SI\) and \(NI\)) instead of four \((BI, BO, SI\) and \(SO\)).

## 2.4 Empirical Results for the Homogeneous Model

In this section I first test the predictions of the homogeneous model with a multinomial logit model. To assess the magnitude of bias a unilateral integration assumption brings, I use a system of linear probability models to test the predictions of this model and those of a property rights model with unilateral integration assumption. I then extend the definition of relative R&D intensity from pair-specific to industry-specific. Last but not least, to test whether the methodology adopted by other researchers works in this database, I replicated the empirical exercise per Acemoglu et al. (2009). I do not find positive support for their theory in this database.

### 2.4.1 Multinomial Logit Model

Recall that in the homogeneous model, the profit function of pair \(i\) is

\[
\pi^k(\eta_i) = \frac{1 - \alpha[\beta^k \eta_i + (1 - \beta^k)(1 - \eta_i)]}{\{(1/\alpha) \left( \frac{c_{i,b}}{\beta^k} \right)^{\eta_i} \left( \frac{c_{i,s}}{1 - \beta^k} \right)^{1 - \eta_i} \}^{1 - \alpha}}, \quad k \in \{BI, SI, NI\},
\]

where \(i\) is a firm pair index. Suppose there is a shock in the profit function, such that it becomes

\[
\pi^k(\eta_i) = \psi^k(\eta_i) + \varepsilon^k_i,
\]

where

\[
\psi^k(\eta_i) = \frac{1 - \alpha[\beta^k \eta_i + (1 - \beta^k)(1 - \eta_i)]}{\{(1/\alpha) \left( \frac{c_{i,b}}{\beta^k} \right)^{\eta_i} \left( \frac{c_{i,s}}{1 - \beta^k} \right)^{1 - \eta_i} \}^{1 - \alpha}},
\]
and $\varepsilon_i^k$ is an i.i.d. error term drawn from a type I extreme value distribution.

Let $y_i$ denote the organizational form of pair $i$. By McFadden (1973), the probability of a pair $i$ choosing organizational form $k$ is

$$\Pr\{y_i = k\} = \frac{\exp\{\psi^k(\eta_i)\}}{\sum_{k' \in \{BI, SI, NI\}} \exp\{\psi^{k'}(\eta_i)\}},$$

Theorem 6 predicts that $\Pr\{y_i = BI\}$ increases in $\eta$ and $\Pr\{y_i = SI\}$ decreases in $\eta$. This implies that if we run a multinomial logistic regression of $y_i$ on $\eta_i$, controlling for the other observable pair, industry and country characteristics, the coefficient on $\eta_i$ should be positive for $y_i = BI$ and negative for $y_i = SI$.

To test the predictions of Theorem 6, I run the following multinomial logit regression:

$$\Pr(y_i = k) = \frac{\exp(\beta^k \cdot rd_b,i + X'_i \varphi^k)}{\sum_{k' \in \{BI,NI,SI\}} \exp(\beta^{k'} \cdot rd_b,i + X'_i \varphi^{k'})}, \quad (2.20)$$

where $i$ is an indicator of a pair of firms. $y_i$ is the organizational decision of pair $i$. $rd_b,i$ is the buyer firm’s relative R&D intensity in pair $i$, as defined in Section 2.2.3. It is an empirical proxy for $\eta$. $k \in \{BI, SI, NI\}$ refers to the different ownership structures. $X_i$ is a vector of pair characteristics (including the log difference in the buyer and seller’s age and employment, sector-pair and region-pair characteristics). $\beta^k$ and $\varphi^k$ are the ownership-specific coefficients. $\beta^k$ is the coefficient of interest. Our expectation is that $\beta^k$ should be positive for outcome $y_i = BI$ and negative for outcome $k = SI$.

Table 2.5 summarizes regression results. Columns (1)-(4) show the regression results for buyer’s relative R&D intensity; Columns (5)-(8) show the regression results for buyer’s relative capital intensity. As predicted by Theorem 5, buyer’s relative R&D and capital intensity positively affects the pair’s probability of choosing buyer integration and negatively affects the pair’s probability of choosing seller integration. This effect remains significant after adding in size controls and sector-pair and region-pair fixed effects.\(^{17}\)

Since it is difficult to interpret the coefficients of a multinomial logit regression, I calculate the average marginal effects of the variables of interest (buyer’s relative R&D and capital intensities) at different margins and report the results in Table 2.6.

\(^{17}\)There are $(10 \times 10 =) 100$ sector-pair fixed effects, and $(5 \times 5 =) 25$ region-pair fixed effects. I could not add in higher dimensions of sector-pair and region-pair fixed effects due to well-known convergence difficulties in multinomial logit models. Much higher-dimensional fixed effects are included in the linear probability models.
### Table 2.5: Multinomial Logistic Regression Results

<table>
<thead>
<tr>
<th></th>
<th>R&amp;D Intensity</th>
<th>Capital Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>k=BI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rb,i</td>
<td>2.128***</td>
<td>2.002***</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>ln(emp_b,i)-ln(emp_s,i)</td>
<td>0.101***</td>
<td>0.096***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>ln(age_b,i)-ln(age_s,i)</td>
<td>0.068***</td>
<td>0.053***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>k=SI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rb,i</td>
<td>-1.192***</td>
<td>-0.944***</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>ln(emp_b,i)-ln(emp_s,i)</td>
<td>-0.188***</td>
<td>-0.191***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>ln(age_b,i)-ln(age_s,i)</td>
<td>-0.168***</td>
<td>-0.104***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.013)</td>
</tr>
</tbody>
</table>

**Notes:** The dependent variable for all the regressions in this table is the organizational form \( k \). Columns (1)-(4) report the regression results of \( k \) on buyer’s relative R&D intensity. Columns (5)-(8) report the regression results of \( k \) on buyer’s relative capital intensity. The definitions of these two intensities are given in Section 2.2.3. The two control variables are the log difference in the buyer and seller’s employment and ages. There are \((10 \times 10 =) 100\) unique sector pairs and \((5 \times 5 =) 25\) region pairs.

The 8 columns in Table 2.6 correspond to the average marginal effects of the 8 regressions in Table 2.5. An average effect is the average of the marginal effects for all observations. The upper panel reports the average marginal effect of buyer’s relative R&D/capital intensity on the probability of buyer integration, the lower panel reports the marginal effects of the same variable on the probability of choosing seller integration. The row labels represent the different margins. Take Column (1) for example, the average marginal probability of choosing buyer integration is 24% and the average marginal probability of choosing seller integration is 16%. The numbers in parentheses report the standard errors at these margins. When the value of the buyer’s relative R&D intensity changes from the minimum to the maximum, the probability of buyer integration increases by 24%, while the probability of seller integration decreases by 16%. When the buyer’s relative R&D intensity (of every observation) changes by
Table 2.6: Average Marginal Effects of the Multinomial Logit Model

<table>
<thead>
<tr>
<th></th>
<th>R&amp;D Intensity (1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>Capital Intensity (5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>k=BI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marginal</td>
<td>0.241</td>
<td>0.214</td>
<td>0.189</td>
<td>0.158</td>
<td></td>
<td>0.376</td>
<td>0.344</td>
<td>0.345</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Min→Max</td>
<td>0.246</td>
<td>0.216</td>
<td>0.195</td>
<td>0.165</td>
<td></td>
<td>0.393</td>
<td>0.355</td>
<td>0.373</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>±1/2</td>
<td>0.243</td>
<td>0.215</td>
<td>0.19</td>
<td>0.159</td>
<td></td>
<td>0.382</td>
<td>0.352</td>
<td>0.355</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td><strong>k=SI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marginal</td>
<td>-0.163</td>
<td>-0.115</td>
<td>-0.107</td>
<td>-0.09</td>
<td></td>
<td>-0.314</td>
<td>-0.226</td>
<td>-0.214</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Min→Max</td>
<td>-0.154</td>
<td>-0.111</td>
<td>-0.107</td>
<td>-0.093</td>
<td></td>
<td>-0.3</td>
<td>-0.22</td>
<td>-0.214</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>±1/2</td>
<td>-0.163</td>
<td>-0.115</td>
<td>-0.107</td>
<td>-0.09</td>
<td></td>
<td>-0.318</td>
<td>-0.227</td>
<td>-0.214</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

Notes: This table reports the average marginal effect, which is the average of the marginal effects of all observations, of the buyer’s relative R&D/capital intensity on the pair’s probability of choosing buyer integration (the upper panel) and seller integration (the lower panel) at different margins. Columns (1)-(4) report the average marginal effects of the buyer’s relative R&D intensity. Columns (5)-(8) report the average marginal effects of the buyer’s relative capital intensity. The row labels represent the different types of marginal changes in buyer’s relative R&D/capital intensity. “Marginal” represent the derivative of the probability function with respect to buyer’s relative R&D/capital intensity. “Min→Max” represents when buyer’s relative R&D/capital intensity changes from the minimum to the maximum. “±1/2” represents a 1 unit change in buyer’s relative R&D/capital intensity.

1 unit, the probability of buyer integration increases by 24% on average, and the probability of seller integration decreases by 16%. When the magnitude of change is one standard deviation, the probability of buyer integration increases by 11% and the probability of seller integration decreases by 7%. Similar interpretations can be made for all the other columns.

2.4.2 Linear Probability Model

A multinomial logit regression provides positive support for the predictions of the homogeneous model, but it does not provide a comparison between the predictions of a bilateral integration model and a unilateral integration model, nor does it allow me to control for finer levels of fixed effects. Linear regressions help solve these problems.

A property rights model with unilateral integration predicts that the principal firm chooses
integration when the principal firm’s non-contractible, relationship-specific investment is more important, and outsourcing when the agent’s non-contractible, relationship-specific investment is more important. The theory suggests the following linear probability model

\[ \Pr(VI_i = 1) = \beta^V r_{d,i} + X_i' \varphi^V + \varepsilon_i^V, \]  

where \( i \) is a pair index. \( VI_i \) is a dummy that equals 1 if pair \( i \) is integrated and 0 otherwise. \( r_{d,i} \) is the buyer firm’s relative R&D intensity. If buyer is assumed to be the principal, \( r_{d,i} \) measures the relative importance of the principal’s non-contractible, relationship-specific investment. \( \beta^V \) is expected to be positive. If the seller is assumed to be the principal, \( r_{d,i} \) measures the relative importance of the agent firm’s non-contractible, relationship-specific investment. \( \beta^V \) is expected to be negative. In other words, a unilateral integration assumption can support any significant value of \( \beta^V \): if \( \beta^V \) is positive and significant, it supports the unilateral integration assumption with the buyer being the principal; if \( \beta^V \) is negative and significant, it supports the unilateral integration assumption with the seller being the principal. This is what Whinston (2001) points out: “the discussion above sidesteps one important issue by assuming that any observed integration is buyer integration. When this is not clear a priori, the PRT’s prediction...will depend on whether the type of integration (buyer vs. seller) is observable.”

My model with the bilateral integration assumption suggests the following linear probability models

\[ \Pr(BI_i = 1) = \beta^B r_{d,i} + X_i' \varphi^B + \varepsilon_i^B, \]  

and

\[ \Pr(SI_i = 1) = \beta^S r_{d,i} + X_i' \varphi^S + \varepsilon_i^S, \]  

where \( i, r_{d,i}, \) and \( X_i \) are the same as defined above. \( BI_i \) is a dummy variable which equals 1 if the buyer integrates the seller and 0 otherwise; \( SI_i \) is a dummy variable which equals 1 if the seller integrates the buyer and 0 otherwise. The theory predicts that \( \beta^B > 0 \) and \( \beta^S < 0 \).

Note that \( VI_i = 1 \) whenever \( BI_i = 1 \) or \( SI_i = 1 \). Since \( BI_i = 1 \) and \( SI_i = 1 \) are mutually exclusive, \( \Pr(VI_i = 1) = \Pr(BI_i = 1) + \Pr(SI_i = 1) \). This implies that equation (2.21) is a conflation of equations (2.22) and (2.23), and \( \beta^V = \beta^B + \beta^S \). If there is no seller integration in the data, \( \beta^S = 0 \), and \( \beta^V = \beta^B \); if there is no buyer integration in the data, \( \beta^B = 0 \) and \( \beta^V = \beta^S \). But when buyer integration and seller integration coexist, \( \beta^V \) is simply a combination
Table 2.7: Linear Probability Model – Pair-level R&D/Capital Intensities

<table>
<thead>
<tr>
<th>Dependent Variable: BIi</th>
<th>Dependent Variable: SIi</th>
<th>Dependent Variable: VIi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>rb,i</td>
<td>0.244***</td>
<td>0.219***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>ln(empb,i)-ln(ems,i)</td>
<td>0.011***</td>
<td>0.013***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>ln(ageb,i)-ln(agea,i)</td>
<td>0.010*</td>
<td>0.011**</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

Industry-pair fixed effects | N | N | Y | Y | N | N | Y | Y | N | N | N | N | Y | Y
Country-pair fixed effects | N | N | N | Y | N | N | N | Y | N | N | N | N | Y | Y
Observations | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783
R-squared | 0.099 | 0.121 | 0.294 | 0.351 | 0.039 | 0.116 | 0.213 | 0.270 | 0.008 | 0.018 | 0.226 | 0.297 |

(a) Linear Probability Model – R&D Intensity

<table>
<thead>
<tr>
<th>Dependent Variable: BIi</th>
<th>Dependent Variable: SIi</th>
<th>Dependent Variable: VIi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>rb,i</td>
<td>0.378***</td>
<td>0.344***</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>ln(empb,i)-ln(ems,i)</td>
<td>0.009***</td>
<td>0.009***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>ln(ageb,i)-ln(agea,i)</td>
<td>0.012*</td>
<td>0.014***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

Industry-pair fixed effects | N | N | N | Y | N | N | N | Y | N | N | N | Y
Country-pair fixed effects | N | N | N | Y | N | N | N | Y | N | N | N | Y
Observations | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 | 85783 |
R-squared | 0.119 | 0.137 | 0.328 | 0.379 | 0.075 | 0.139 | 0.229 | 0.277 | 0.003 | 0.012 | 0.228 | 0.300 |

(b) Linear Probability Model – Capital Intensity

Notes: **p < 0.05; *p < 0.1. There are 2754 industry pairs and 2077 country pairs. All regression results are two-way clustered at the industry-pair and country-pair levels. Standard errors are reported in parentheses.
of $\beta^B$ and $\beta^S$, and it is no longer a test of property rights model.

Table 2.7 reports the linear regression results for the above linear probability models. Table 2.7a reports the regression results for the buyer’s relative R&D intensity; Table 2.7b reports the regression results for the buyer’s relative capital intensity. In each table, columns (1)-(4), (5)-(8) and (9)-(12) are respectively the results with $BI_i$, $SI_i$ and $VI_i$ as dependent variables. Numbers in parentheses report standard errors. There are three rows of standard errors reported for the variables of interest. The first row are standard errors without clustering. The second row reports standard errors when clustering at the pairwise industry-levels. The third row reports standard errors when clustering at pairwise industry- and country-levels.

The coefficients on $rd_{b,i}$ are consistent with our expectations. When the dependent variable is $BI_i$, the coefficients on $rd_{b,i}$ are positive and significant. This effect remains significant after adding in size, age, industry-pair and country-pair fixed effects. When the dependent is $SI_i$, the coefficients on $rd_{b,i}$ are negative and significant. This effect remains significant after adding in size, age, industry-pair and country-pair fixed effects. Columns (1)-(8) show that the predictions of a bilateral integration model are positively supported by linear probability models. Columns (9)-(12) report the regression results for equation (2.21). The coefficients on $rd_{b,i}$ are positive and significant, which support the buyer integration assumption but rejects the seller integration assumption. As explained before, the coefficients in columns (9)-(12) are simply additions of the coefficients in (1)-(4) and (5)-(8). For example, the value in column (9), 0.09, is the summation of the value in column (1), 0.244, and the value in column (5), -0.154. A researcher assuming buyer integration finds positive support for the PRT model, but the magnitude of coefficients (0.09) are roughly one-third of the PRT model with bilateral integration assumption (0.244 and -0.154). On the other hand, a researcher assuming seller integration would find negative support for the PRT model.

The reason why we observe positive and significant coefficients for $rd_{b,i}$ in columns (9)-(12) is because the coefficients for $BI_i$ are larger in magnitudes than the coefficients for $SI_i$. If their magnitudes are similar, it is possible to observe an insignificant coefficient in columns (9)-(12). A researcher testing a PRT model with unilateral integration assumption would reject the PRT, while the data supports a PRT model with bilateral integration assumption.

A criticism that PRT models often suffer from is that one can adopt an assumption that rationalizes any pattern in the data. This problem cannot be dealt with in a firm-level database. But we show here that this problem does exist, and that it can be addressed in a relationship-
level database.

To test if the pattern is driven by the linearity assumption in the linear probability models, I also include the logit and probit regression results in Appendix 2.F. The patterns in these tables are very similar to the linear probability models discussed in this section.

2.4.3 Robustness: Industry-level R&D and Capital Intensity

Recall that the theory predicts a relationship between \( k \), the organizational choice and \( \eta \), the relative importance of buyer’s non-contractible investment is a pair-specific measure, meaning that within the same industry-pair, different firm pairs have different values of \( \eta \). My theory remains silent on the level of aggregation of \( \eta \). That is, the theory does not require \( \eta \) to be firm pair-specific or industry pair-specific. Most previous empirical research has assumed that \( \eta \) is an industry pair-specific measure. However, I believe this assumption is made due to data limitation as other databases contain investment information on at most one side of a relationship. A firm-pair specific measure is not implementable. Although I believe \( \eta \) to be a firm pair-specific measure, I do not want to restrict the width of my empirical research. This is the reason I do not extend the theory in this paper to a general equilibrium, as a general equilibrium model (with heterogeneous firms) requires one to give specific definitions of \( \eta \). My previous empirical analysis has been assuming that \( \eta \) is a firm pair-specific measure. In this section, I redefine \( \eta \) as an industry pair-specific measure.

There are two justifications for using industry measures of R&D and capital intensities. First, both R&D and capital investments are highly skewed variables (as shown in Table 2.3.). It is usually the large firms that make the majority of the R&D expenditures in an industry.\(^{18}\) Second, the theory suggests that \( \eta \) is a parameter whose value is given before the two firms choose their levels of relationship-specific investments. In this sense, we are using ex post values (R&D expenditure, total capital and total revenue, etc.) to proxy for an ex ante parameter. This proxy is valid only if the ratio between R&D/capital investment and total revenue depends on \( \eta \) and nothing else. In a more empirical language, there may be reverse causality between \( \eta \) and R&D intensity. An industry-level R&D or capital intensity takes into consideration other firms’ strategic behaviors. In this sense, they may be able to dilute the reverse causality issue (if there is one), since firms in an industry are of different ages and might be at different stages

\(^{18}\)This is not saying that an industry-level measure of R&D intensity is not representative, as shown by the large literature on learning-by-doing.
Table 2.8: Linear Probability Model - Industry-level R&D/Capital Intensities

(a) Industry R&D Intensity

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BI_i</th>
<th>SI_i</th>
<th>VI_i</th>
</tr>
</thead>
<tbody>
<tr>
<td>rd_{b,i}</td>
<td>0.207*** (0.023)</td>
<td>0.207*** (0.018)</td>
<td>-0.059*** (0.020)</td>
</tr>
<tr>
<td>ln(tr_{b,i})-ln(tr_{s,i})</td>
<td>0.013*** (0.003)</td>
<td>-0.024*** (0.005)</td>
<td>-0.011*** (0.004)</td>
</tr>
<tr>
<td>ln(emp_{b,i})-ln(emp_{s,i})</td>
<td>0.003 (0.004)</td>
<td>0.000 (0.003)</td>
<td>0.004 (0.003)</td>
</tr>
<tr>
<td>ln(age_{b,i})-ln(age_{s,i})</td>
<td>0.006 (0.004)</td>
<td>-0.007 (0.005)</td>
<td>-0.001 (0.006)</td>
</tr>
<tr>
<td>Observations</td>
<td>85780</td>
<td>85780</td>
<td>85780</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.040</td>
<td>0.089</td>
<td>0.003</td>
</tr>
</tbody>
</table>

(b) Industry Mean R&D intensity

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BI_i</th>
<th>SI_i</th>
<th>VI_i</th>
</tr>
</thead>
<tbody>
<tr>
<td>rd_{b,i}</td>
<td>0.212*** (0.025)</td>
<td>0.220*** (0.021)</td>
<td>-0.061*** (0.023)</td>
</tr>
<tr>
<td>ln(tr_{b,i})-ln(tr_{s,i})</td>
<td>0.012*** (0.003)</td>
<td>-0.024*** (0.005)</td>
<td>-0.012*** (0.004)</td>
</tr>
<tr>
<td>ln(emp_{b,i})-ln(emp_{s,i})</td>
<td>0.004 (0.004)</td>
<td>0.000 (0.003)</td>
<td>0.004 (0.003)</td>
</tr>
<tr>
<td>ln(age_{b,i})-ln(age_{s,i})</td>
<td>0.008* (0.004)</td>
<td>-0.008 (0.005)</td>
<td>0.000 (0.006)</td>
</tr>
<tr>
<td>Observations</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.037</td>
<td>0.088</td>
<td>0.003</td>
</tr>
</tbody>
</table>

of development.

I continue to use the three linear probability models defined in Section 2.4.2, replacing the pair-level variables of interest with industry-level measures. Table 2.8 reports the regression results with industry-level R&D and capital intensity measures: industry R&D intensity which is the industry’s total R&D expenditure divided by the industry’s total revenue (Table 2.8a); industry mean R&D intensity which is the average level of R&D intensity across all firms in this industry (Table 2.8b). The specification in Table 2.8 is robust to various other measures including industry median R&D or capital intensity, industry R&D or capital over value-added.
2.4.4 Robustness: Regression with Only Imputed Relationships

To test the external validity of this database, I use previously adopted empirical methodology to perform an examination of the property-rights theory. The paper I use for this purpose is Acemoglu et al. (2010).

Acemoglu et al. (2010) use the UK-plant level data to test the predictions of the property-rights theory assuming backward integration is the only form of integration. The benchmark empirical specification is

\[
v_{ijk} = \alpha s_{jk} + \beta_P R^D_P + \beta_S R^S_K + X'_{ijk} \eta + \varepsilon_{ijk}.
\]

The dependent variable is defined as

\[
v_{ijk} = \begin{cases} 
1 & \text{if firm } i \text{ owns a plant in industry } k \text{ supplying industry } j, \\
0 & \text{otherwise}.
\end{cases} \tag{2.24}
\]

\(s_{jk}\) is calculated from the UK Input-Output table. It is the share of inputs from industry \(k\) in the total cost of industry \(j\) (£ of input \(k\) necessary to produce £1 of product \(j\)). \(R^D_P\) is R&D intensity in the producing industry \(j\), \(R^S_K\) is R&D intensity in the supplying industry \(k\), \(X_{ijk}\) is a vector including the constant term and firm and industry characteristics (firm size and age, average firm size and age in the producing and supplying industries. The main coefficients of interest are \(\alpha\), \(\beta_P\) and \(\beta_S\). Their theory predicts that greater technology intensity of the producer (higher \(R^D_P\)), lower technology of the supplier (lower \(R^S_K\)), and greater share of costs of the producer accounted for by inputs from the supplier (higher \(s_{jk}\)) should make (backward) integration more likely. Thus one should expect that \(\alpha > 0\), \(\beta_P > 0\) and \(\beta_S < 0\).

They also proposed a second regression:

\[
v_{ijk} = \alpha s_{jk} + (\beta_P + \gamma_P s_{jk}) R^D_P + (\beta_S + \gamma_S s_{jk}) R^S_S + X'_{ijk} \eta + \varepsilon_{ijk}.
\]

The expectation is that \(\gamma_P > 0\) and \(\gamma_S < 0\), as the producing firm’s decisions should be more sensitive to \(R^D_P\) and \(R^S_K\) the more important the supplier’s input (or the higher is \(s_{jk}\)) is.

I use the same Input-Output table from previous analysis and tried many possible specifications. The two regression equations in this section are very sensitive to specifications. The only variable that is relatively robust and significant is \(R^S_K\), the supplying industry’s R&D
Table 2.9: Robustness Check: Regressions using Only Imputed Relationships

<table>
<thead>
<tr>
<th>Dependent variable: $v_{ijk}$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$sc_{jk}$</td>
<td>1.563***</td>
<td>1.123***</td>
<td>1.134***</td>
<td>1.111***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.174)</td>
<td>(0.242)</td>
<td>(0.245)</td>
<td>(0.262)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$RD_j^P$</td>
<td>0.909*</td>
<td>0.928***</td>
<td>1.018***</td>
<td>0.477</td>
<td>0.990**</td>
<td>0.477</td>
</tr>
<tr>
<td></td>
<td>(0.412)</td>
<td>(0.260)</td>
<td>(0.256)</td>
<td>(0.262)</td>
<td>(0.302)</td>
<td>(0.262)</td>
</tr>
<tr>
<td>$RD_j^S \times sc_{jk}$</td>
<td>3.999</td>
<td>5.446</td>
<td>0.145</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(19.950)</td>
<td>(20.850)</td>
<td>(18.934)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.346)</td>
<td>(0.293)</td>
<td>(0.306)</td>
<td>(0.297)</td>
<td>(0.318)</td>
<td></td>
</tr>
<tr>
<td>$RD_j^S \times sc_{jk}$</td>
<td>-33.299</td>
<td>-33.727</td>
<td>-28.068</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(19.131)</td>
<td>(20.064)</td>
<td>(18.275)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln difference in firm size (jk)</td>
<td>0.032***</td>
<td>0.033***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln difference in firm age (jk)</td>
<td>0.008***</td>
<td>0.010***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln average firm size in the producing industry (j)</td>
<td>-0.064*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln average firm size in the supplying industry (k)</td>
<td>-0.194***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln average firm age of the producing industry (j)</td>
<td>-0.017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln average firm age of the supplying industry (k)</td>
<td>0.016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2618392</td>
<td>2618392</td>
<td>2618392</td>
<td>2618392</td>
<td>2342841</td>
<td>2342841</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.001</td>
<td>0.063</td>
<td>0.080</td>
<td>0.081</td>
<td>0.127</td>
<td>0.133</td>
</tr>
</tbody>
</table>

Intensity. For illustration purposes, I report in Table 2.9 the regression results for one of the many regressions I tried. It is the one that puts the Acemoglu et al. (2009) theory in the most favorable light. This table reports the U.S. manufacturing firms. All results are clustered at the industry-pairs levels. The coefficient of the share of costs variable is positive and significant as predicted by the theory. This is the case for almost all the other regressions that I tried. The coefficient of the supplying industry’s R&D intensity is negative and significant, also as predicted by the theory. This is the second most robust variable. Unfortunately, I do not find support for the theory for any other variables in this regression.
2.5 Empirical Results for the Heterogeneous Model

2.5.1 Empirical Specification

The key predictions of the heterogeneous model are stated in Theorem 6, which states that when \( \eta \) is low, more productive pairs (conditional on the same level of \( \eta \)) choose seller integration and more productive firms choose outsourcing. The least productive pairs exit. When \( \eta \) is high, more productive pairs choose buyer integration and less productive pairs choose outsourcing. The least productive pairs exit. To see if my data is close to consistent with these predictions, I estimate the following linear probability models

\[
\Pr(BI_i = 1) = \alpha^B + \text{Prod}_i \sum_{q=1}^{Q} \gamma_q^B RD_q + X_i \xi^B + \varepsilon_i^B,
\]

\[
\Pr(SI_i = 1) = \alpha^S + \text{Prod}_i \sum_{q=1}^{Q} \gamma_q^S RD_q + X_i \xi^S + \varepsilon_i^S,
\]

and

\[
\Pr(NI_i = 1) = \alpha^N + \text{Prod}_i \sum_{q=1}^{Q} \gamma_q^N RD_q + X_i \xi^N + \varepsilon_i^N.
\]

\( BI_i, SI_i \) and \( NI_i \) are dummy variables which take value 1 when pair \( i \) chooses buyer integration, seller integration or outsourcing. \( \text{Prod}_i \) is a productivity measure for pair \( i \) which is defined later. \( RD \) is an empirical proxy for \( \eta \). Since \( \eta \) can either be an industry or pair characteristic, there are two empirical definitions for \( RD \). It is defined as the relative R&D intensity of either the buyer firm, or the buyer firm’s industry. \( RD \) is divided into \( Q \) quantiles so \( RD_q \) refers to the \( q \)th quantile of \( RD \). A small \( q \) represents a lower quantile and a smaller value of \( RD \). \( \gamma_q \) are the coefficients for the interaction terms between \( \text{Prod}_i \) and \( RD_q \). \( X_i \) is a vector of other characteristics of pair \( i \), such as the two firms’ ages, industry and country information. \( \alpha^B, \alpha^S \) and \( \alpha^N \) are constants. \( \varepsilon_i^B, \varepsilon_i^S \) and \( \varepsilon_i^N \) are error terms. If \( \eta \) is pair-specific, \( RD_q \) is ranked within the buyer firm and the seller firm’s industry pairs; if \( \eta \) is industry-specific, \( RD \) is ranked across all industries.

I construct \( \text{Prod}_i \) as a weighted average of the buyer firm and the seller firm’s levels of productivity:

\[
\text{Prod}_i = \frac{\text{size}_b}{\text{size}_b + \text{size}_s} \ln \text{Prod}_b + \frac{\text{size}_s}{\text{size}_b + \text{size}_s} \ln \text{Prod}_s.
\]
A firm’s productivity is measured as its total revenue per employee. size refers to a firm’s total revenue. Since the dependent variables are dummies, I normalize Prod\textsubscript{i} by replacing it with \((Prod\textsubscript{i} - Prod_{25})/(Prod_{75} - Prod_{25})\), where Prod\textsubscript{25} and Prod\textsubscript{75} are the 25th and the 75th percentiles of all pair’s productivity levels.

When \(q\) is small, the theory predicts that more productive pairs choose seller integration, less productive pairs choose outsourcing and no one chooses buyer integration, so at lower \(q\)’s, \(\gamma_{q}^{S} > 0, \gamma_{q}^{N} < 0\) and \(\gamma_{q}^{B}\) is insignificant. When \(q\) is large, the theory predicts that more productive firms choose buyer integration, less productive firms choose outsourcing and no one chooses seller integration, so at higher \(q\), \(\gamma_{q}^{B} > 0, \gamma_{q}^{N} < 0\) and \(\gamma_{q}^{S}\) is insignificant. When \(q\) is in the middle, the theory predicts that all existing firms choose outsourcing, so at intermediate \(q\), all \(\gamma\)’s should be insignificant and we should only observe outsourcing. As mentioned before, there are two definitions for \(\eta\), so there are two definitions for \(RD\).

There are two alternative measures for \(R&D\) intensities, pair-level and industry-level. I report them separately. The pair-level R&D intensity quantiles are divided within each industry pair. The industry-level R&D intensity quantiles are divided across all industry pairs. The linear regression results for the tertile R&D intensity measures are reported in the next two sections.

### 2.5.2 Regression Results for Pair-level R&D Intensities

Table 2.10 reports the linear regression results for tertiles of pair-level R&D intensities. The left panel regress each of the three dummies, BI\textsubscript{i}, SI\textsubscript{i} and NI\textsubscript{i} on a series of independent variables including the interaction of pair \(i\)’s productivity and tertile dummies RD\textsubscript{1}, RD\textsubscript{2} and RD\textsubscript{3}, and log difference in the ages of the buyer firm and the seller firm, and industry-pair and country-pair fixed effects. The right panel run similar regressions, only replacing RD\textsubscript{q}, the tertile dummies of the buyer firm’s relative R&D intensity, with \(1 - RD_{q}\), the tertile dummies of the seller firm’s relative R&D intensity. All regressions are two-way clustered at industry-pair and country-pair levels.

We can see from Table 2.10 that the predictions of the heterogeneous model are supported. Let us look at the left panel first. Columns (1)-(3) report the regression results for when the dependent variables are BI\textsubscript{i}, SI\textsubscript{i} and NI\textsubscript{i}, respectively. At the first quantile of RD (equivalent to the case where \(\eta\) is low), more productive firms are more likely going to choose seller integration and less likely going to choose buyer integration. At the third quantile of RD (equivalent to the
Table 2.10: Regression Results with Tertiles of Pair-specific $RD$ and $(1 - RD)$

(a) Buyer’s R&D intensity

| Dependent variable: | BI<sub>i</sub> | SI<sub>i</sub> | NI<sub>i</sub> | | Dependent variable: | BI<sub>i</sub> | SI<sub>i</sub> | NI<sub>i</sub> |
|---------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Prod<sub>i</sub>*RD<sub>1</sub> | -0.229*** | 0.102** | 0.126* | Prod<sub>i</sub>*RD<sub>1</sub>(1-RD<sub>1</sub>) | 0.097** | -0.163*** | 0.066 |
|                     | (0.037) | (0.037) | (0.051) | | | (0.032) | (0.038) | (0.050) |
| Prod<sub>i</sub>*RD<sub>2</sub> | 0.034 | -0.250*** | 0.216***Prod<sub>i</sub>*RD<sub>2</sub>(1-RD<sub>2</sub>) | -0.303*** | 0.045 | 0.258*** |
|                     | (0.036) | (0.046) | (0.055) | | | (0.045) | (0.038) | (0.055) |
| Prod<sub>i</sub>*RD<sub>3</sub> | 0.136*** | -0.165*** | 0.029 | Prod<sub>i</sub>*RD<sub>3</sub>(1-RD<sub>3</sub>) | -0.310*** | 0.175*** | 0.135* |
|                     | (0.039) | (0.040) | (0.053) | | | (0.043) | (0.044) | (0.054) |
| ln(age<sub>b,i</sub>) - ln(age<sub>s,i</sub>) | 0.041*** | -0.055*** | 0.015***ln(age<sub>b,i</sub>) - ln(age<sub>s,i</sub>) | 0.039*** | -0.057*** | 0.018*** |
|                     | (0.004) | (0.004) | (0.003) | | | (0.004) | (0.005) | (0.003) |

Observations: 85783 85783 85783 Observations: 85783 85783 85783
R-squared: 0.051 0.060 0.004 R-squared: 0.066 0.051 0.006

Notes: This table reports the linear regression results for tertiles of pair-level R&D intensities. The left panel regress each of the three dummies, $BI_i$, $SI_i$ and $NI_i$ on a series of independent variables including the interaction of pair $i$’s productivity and tertile dummies $RD_1$, $RD_2$ and $RD_3$, and log difference in the ages of the buyer firm and the seller firm, and industry-pair and country-pair fixed effects. The right panel run similar regressions, only replacing $RD_q$, the tertile dummies of the buyer firm’s relative R&D intensity, with $1 - RD_q$, the tertile dummies of the seller firm’s relative R&D intensity. All regressions are two-way clustered at industry-pair and country-pair levels.

case where $\eta$ is high), more productive firms are more likely going to choose buyer integration and less likely going to choose seller integration. At the second quantile of $RD$, more productive firms are more likely going to choose outsourcing and less likely going to choose the other two forms of ownership structures. The opposite is true for the right panel. At the first quantile of $(1 - RD)$, more productive pairs are more likely going to choose buyer integration and less likely going to choose seller integration. At the first quantile of $(1 - RD)$, more productive pairs are more likely going to choose seller integration and less likely going to choose buyer integration. At the second quantile of $(1 - RD)$, more productive firms are more likely going to choose outsourcing and less likely going to choose the other two ownership structures.

The regression results for quintiles and 7-quantiles of relative R&D intensities are reported in Appendix 2.G. They display similar patterns as discussed here.
2.5.3 Regression Results for Pair-level R&D Intensities

Table 2.11 reports the regression results for tertiles of industry-specific $RD$ and $(1 - RD)$. The left table suggests that at low levels of $\eta$ (the first tertile), more productive pairs are more likely going to choose outsourcing while at high levels of $\eta$ (the third tertile), more productive firms are more likely going to choose buyer integration. The right panel also suggests that at low levels of $\eta$ (the third tertile), more productive firms are more likely going to choose outsourcing. At high levels of $\eta$ (the first tertile), more productive pairs are more likely going to choose buyer integration. There is no clear pattern when $\eta$ is in the middle for either panel.

The regression results for quintiles and 7-quantiles of industry-level $RD$ and $1 - RD$ are reported in Appendix 2.G. They are not consistent with the theory’s predictions either. There does not seem to be any distinct patterns in these tables either.

2.6 Conclusion

In this paper, I compile a relationship-level database which contradicts an implicit assumption made by previous firm boundary studies, that is, within an industry-pair, only one form of integration exists, backward or forward. I show that they do coexist within 774 out of the 2,754 industry pairs in this database.

To explain why backward and forward integration coexist within certain industry pairs, I
develop a property rights model with bilateral integration decisions based on Grossman and Hart (1986) and Antrás and Helpman (2004). This model predicts that the likelihood of a pair of firms choosing buyer (or backward) integration increases in the buyer firm’s relative R&D intensity, while the likelihood of the pair choosing seller integration decreases in the buyer firm’s relative R&D intensity.

I then empirically test the predictions of this model in multinomial logit, linear, logit and probit regressions, with various measures of relative R&D intensities. The predictions of the above model are robustly supported. To see how much bias a unilateral integration assumption brings, I compare the empirical results of a bilateral integration model with those of a unilateral integration model. I show that the unilateral integration assumption results in a biased estimation of the property-rights theory.

To see whether a bilateral integration model sustains theoretical extensions, I extend the model to incorporate heterogeneous productivity. The predictions of the model are supported by pair-level R&D intensity measures, but not industry-level R&D intensity measures.

This paper contributes to the literature of supply chain management by showing that the direction of integration matters, and that it can be explained by a property rights framework. It also shows that previous empirical studies testing the property-rights theory suffer from biased estimation by assuming that the integration decisions are unilateral. There are multiple directions of future research. For example, to consider firms’ integration behavior beyond a pairwise perspective i.e., to incorporate a supply chain structure (Antràs, 2015), or a network approach (Acemoglu et al., 2012). Also, given its global feature, this database provides a good source for studying the organizational behavior of multinational firms.
Appendix

2.A Imputing Seller-Buyer Relationships

In this section I describe how I impute seller-buyer relationships from firm-parent information. This process depends heavily on the Input-Output table. Suppose we observe that firm $f$’s parent is firm $p$. Let $A$ denote the Input-Output matrix. Let $a_{ij} \in A$ denote an entry in $A$ in the $i$’s row and $j$’s column, where $i$ is the input industry and $j$ is the output industry. $a_{ij}$ is the dollar amount of industry $i$ input that is required to produce 1 dollar of industry $j$’s output. Industries $i$ and $j$ are considered to be vertically related if $a_{ij} > 0$.

Let $I_f$ and $I_p$ denote the set of industries in which firm $f$ and firm $p$ operate in. Define $a_{fp}^*$ and $a_{pf}^*$ as

$$a_{fp}^* \equiv \max_{i \in I_f, j \in I_p} \{a_{ij} | a_{ij} \in A\}$$

and

$$a_{pf}^* \equiv \max_{i \in I_p, j \in I_f} \{a_{ij} | a_{ij} \in A\}.$$

If $a_{fp}^* > 0$, then firm $f$ and firm $p$ are vertically related, with $f$ in the upstream and $p$ in the downstream. This means firm $f$ is the seller and firm $p$ is the buyer. This relationship is classified as buyer integration (or backward integration) because the buyer (firm $p$) owns the seller (firm $f$). On the other hand, if $a_{pf}^* > 0$, then firm $f$ and firm $p$ are vertically related, with firm $p$ in the upstream and firm $f$ in the downstream. This relationship is classified as seller integration because the seller (firm $p$) owns the buyer (firm $f$). If both $a_{fp}^*$ and $a_{pf}^*$ are positive, I treat $f - p$ and $p - f$ as two separate relationships. There are 8,049 unique firm-parent pairs in all the 19,071 imputed relationships. 3,647 of these 8,049 pairs are bilateral in the sense that the firm sells to the parent and the parent sells to the firm.

2.B Proof of Theorem 5

The proof of this theorem is very similar to that in Antràs and Helpman (2004). Recall that the profit function is

$$\pi^k = \frac{1 - \alpha \beta^k \eta + (1 - \beta^k)(1 - \eta)}{\left\{ (1/\alpha) \left( \frac{\alpha}{\beta^k} \right)^{\eta} \left( \frac{\alpha}{1 - \beta^k} \right)^{1 - \eta} \right\}^{\alpha}}.$$

\(^{19}\)Different thresholds may be applied. For example, in Acemoglu et al. (2010), two industries are considered to be vertically related only when $a_{ij} > 0.01$.\)
Suppose instead of choosing from three discrete values, the two firms could choose a \( \beta \in [0,1] \) to maximize the profit. This optimal \( \beta \) is

\[
\beta^*(\eta) = \eta(\alpha \eta + 1 - \alpha) - \sqrt{\eta(1 - \eta)(1 - \alpha \eta)(\alpha \eta + 1 - \alpha)} \\
2\eta - 1
\]

Figure 2.3: The Optimal Revenue Share as a Function of \( \eta \)

As Figure 2.3 shows, \( \beta^* \) is increasing in \( \eta \). \( \beta^*(0) = 0 \) and \( \beta^*(1) = 1 \). As \( \eta \) increases from 0 to 1, the seller’s investment gets less important while the buyer’s investment gets more important. \( \beta^* \), the optimal revenue share for the buyer increases as the buyer’s investment gets more important.

2.C Properties of \( \mu^k \)

\( \mu^k \) is strictly increasing in \( \beta^k \) and strictly decreasing in \( \eta \). Recall that

\[
\mu^k \equiv \frac{(1 - \alpha \eta)\beta^k}{1 - \alpha[\beta^k \eta + (1 - \beta^k)(1 - \eta)]}
\]

as defined in equation (2.10). Taking the derivative of \( \mu^k \) with respect to \( \beta^k \) and \( \eta \) gives

\[
\frac{\partial \mu^k}{\partial \beta^k} = \frac{(1 - \alpha \eta)(1 - \alpha + \alpha \eta)}{\{1 - \alpha[\beta^k \eta + (1 - \beta^k)(1 - \eta)]\}^2} > 0
\]
and
\[
\frac{\partial \mu^k}{\partial \eta} = \frac{\alpha(\alpha - 2)\beta^k(1 - \beta^k)}{(1 - \alpha[\beta^k\eta + (1 - \beta^k)(1 - \eta)])^2} < 0
\]
since \(0 < \alpha, \beta, \eta < 1\).

### 2.D Proof of Theorem 5

After introducing the i.i.d fixed costs, the profit functions of the homogeneous model can be written as
\[
\pi(\beta^k, \eta) = \psi(\beta^k, \eta) + f^k,
\]
where
\[
\psi(\beta^k, \eta) = \frac{1 - \alpha[\beta^k\eta + (1 - \beta^k)(1 - \eta)]}{(1/\alpha) \left( \frac{\alpha}{\beta^k} \right)^{\eta} \left( \frac{\alpha}{1 - \beta^k} \right)^{1 - \eta}} \frac{\alpha}{1 - \alpha}.
\]

McFadden (1973) shows that when \(f^k\) follows i.i.d type I extreme value distribution, the probability of buyer integration being the organizational form with the highest profit can be written as
\[
\Pr\{k^* = BI\} = \frac{\exp\{\psi(\beta^{BI}, \eta)\}}{\exp\{\psi(\beta^{BI}, \eta)\} + \exp\{\psi(\beta^{SI}, \eta)\} + \exp\{\psi(\beta^{NI}, \eta)\}},
\]
or
\[
\Pr\{k^* = BI\} = \frac{1}{1 + \exp\{\psi(\beta^{SI}, \eta) - \psi(\beta^{BI}, \eta)\} + \exp\{\psi(\beta^{NI}, \eta) - \psi(\beta^{BI}, \eta)\}}.
\]
If \(\psi(\beta^{SI}, \eta) - \psi(\beta^{BI}, \eta)\) and \(\psi(\beta^{NI}, \eta) - \psi(\beta^{BI}, \eta)\) are decreasing in \(\eta\), then \(\Pr\{k^* = BI\}\) is increasing in \(\eta\). Given that \(\beta^{BI} > \beta^{NI} > \beta^{SI}\), we can write \(\psi(\beta^{SI}, \eta) - \psi(\beta^{BI}, \eta)\) and \(\psi(\beta^{NI}, \eta) - \psi(\beta^{BI}, \eta)\) as
\[
\psi(\beta + \Delta \beta, \eta) - \psi(\beta, \eta),
\]
where \(\Delta \beta > 0\). If \(\psi(\beta + \Delta \beta, \eta) - \psi(\beta, \eta)\) is increasing in \(\eta\), or equivalently, if \(\psi(\beta, \eta)\) satisfies increasing differences in \((\beta, \eta)\), then \(\Pr\{k^* = BI\}\) is increasing in \(\eta\).

We showed in Theorem 5 that \(\beta^*(\eta)\) is strictly increasing in \(\eta\), where \(\beta^*(\eta)\) is solved by
\[
\frac{\partial \psi(\beta, \eta)}{\partial \beta} = 0.
\]
Taking the derivative of the above equation w.r.t. \( \eta \) gives

\[
\frac{\partial^2 \psi(\beta, \eta)}{\partial \beta \partial \eta} + \frac{\partial^2 \psi(\beta, \eta)}{\partial \beta^2} \frac{\partial \beta}{\partial \eta} = 0.
\]

Rearrange the above equation

\[
\frac{\partial^2 \psi(\beta, \eta)}{\partial \beta \partial \eta} = -\frac{\partial^2 \psi(\beta, \eta)}{\partial \beta^2} \frac{\partial \beta}{\partial \eta}.
\]

Since \( \frac{\partial^2 \psi(\beta, \eta)}{\partial \beta^2} \leq 0, \frac{\partial \beta}{\partial \eta} > 0, \) and \( \frac{\partial^2 \psi(\beta, \eta)}{\partial \beta \partial \eta} \geq 0, \psi(\beta, \eta) \) is supermodular in \((\beta, \eta)\). This is a sufficient condition for increasing differences, so \( \Pr\{k^* = BI\} \) is increasing in \( \eta \).

Similarly,

\[
\Pr\{k^* = SI\} = \frac{1}{1 + \exp\{\psi(\beta^{BI}, \eta) - \psi(\beta^{SI}, \eta)\} + \exp\{\psi(\beta^{NI}, \eta) - \psi(\beta^{SI}, \eta)\}}.
\]

Since \( \beta^{BI} > \beta^{NI} > \beta^{SI} \) and \( \psi(\beta, \eta) \) is supermodular in \((\beta, \eta)\), \( \psi(\beta^{BI}, \eta) - \psi(\beta^{SI}, \eta) \) and \( \psi(\beta^{NI}, \eta) - \psi(\beta^{SI}, \eta) \) are increasing in \( \eta \). The denominator is increasing in \( \eta \), meaning that \( \Pr\{k^* = SI\} \) is decreasing in \( \eta \).

### 2.6 Proof of Theorem 6

A pair of firms choose \( k = BI \) if and only if

\[
\pi^{BI} = \max\{\pi^{BI}, \pi^{SI}, \pi^{NI}\} > 0,
\]

where \( \pi^k \) is short for \( \pi^k(\theta, \eta) \), which is defined in equation (2.17). For notation simplicity, I will write \( \pi^k(\theta, \eta) \) as

\[
\pi^k = \theta^{\alpha/(1-\alpha)} \psi^k - f^k, \quad k \in \{BI, SI, NI\}.
\]

Note that the above statement contains two statements, \( \pi^{BI} = \max\{\pi^{BI}, \pi^{SI}, \pi^{NI}\} \) and \( \pi^{BI} > 0 \). We discuss the first statement first, which is when buyer integration is the most profitable organizational form.
As McFadden (1973) shows, if \( f^k \) follows a type I extreme value distribution,

\[
\Pr\{\pi^{BI} = \max(\pi^{BI}, \pi^{SI}, \pi^{NI})\} = \frac{\exp\{\theta^{\alpha/(1-\alpha)}\psi^{BI}\}}{\sum_{k \in \{BI, SI, NI\}} \exp\{\theta^{\alpha/(1-\alpha)}\psi^k\}} \cdot \frac{1}{1 + \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{SI} - \psi^{BI}]\} + \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{NI} - \psi^{BI}]\}}.
\]

Recall that \( \psi^k \) is the profit function in the homogeneous model. As Theorem 5 shows, when \( \eta < \eta_0, \psi^{SI} > \psi^{NI} > \psi^{BI} \). This means when \( \eta < \eta_0, \psi^{SI} - \psi^{BI} > 0 \) and \( \psi^{NI} - \psi^{BI} > 0 \), so \( \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{SI} - \psi^{BI}]\} \) and \( \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{NI} - \psi^{BI}]\} \) are increasing in \( \theta \), the denominator is increasing in \( \theta \), so the fraction is decreasing in \( \theta \). This means when \( \eta < \eta_0 \), buyer integration is less likely going to be the organizational form with the highest profit.

On the other hand, when \( \eta > \eta_0, \psi^{BI} > \psi^{NI} > \psi^{SI} \), so that \( \psi^{SI} - \psi^{BI} < 0 \) and \( \psi^{NI} - \psi^{BI} < 0 \). \( \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{SI} - \psi^{BI}]\} \) and \( \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{NI} - \psi^{BI}]\} \) are decreasing in \( \theta \). The denominator is decreasing in \( \theta \) and the fraction is increasing in \( \theta \). This means when \( \eta > \eta_0 \), buyer integration is more likely going to be the organizational form with the highest profit.

Furthermore, when \( \eta < \eta < \eta_0, \psi^{NI} > \min\{\psi^{BI}, \psi^{SI}\} \). Although \( \psi^{NI} - \psi^{BI} > 0, \psi^{SI} - \psi^{BI} \) cannot be signed. It is not clear whether the fraction is increasing or decreasing in \( \theta \).

In short, the probability of buyer integration being the organizational form with the highest profit is decreasing in \( \theta \) when \( \eta < \eta_0 \), and increasing in \( \theta \) when \( \eta > \eta_0 \). When \( \eta < \eta_0 < \eta_0 \), this probability is non-monotonic in \( \theta \). Similarly, it can be shown in the same way that the probability of seller integration being the organizational form with the highest profit is increasing in \( \theta \) when \( \eta < \eta_0 \) and decreasing in \( \theta \) when \( \eta > \eta_0 \).

Now let us turn to \( \Pr\{\pi^{NI} = \max(\pi^{BI}, \pi^{SI}, \pi^{NI})\} \), which can be written as

\[
\Pr\{\pi^{NI} = \max(\pi^{BI}, \pi^{SI}, \pi^{NI})\} = \frac{1}{1 + \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{SI} - \psi^{NI}]\} + \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{BI} - \psi^{NI}]\}}.
\]

When \( \eta < \eta < \eta_0 \), Theorem 5 implies that \( \psi^{NI} > \min\{\psi^{BI}, \psi^{SI}\} \), so \( \psi^{SI} - \psi^{NI} < 0 \) and \( \psi^{BI} - \psi^{NI} < 0 \). \( \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{SI} - \psi^{NI}]\} \) and \( \exp\{\theta^{\alpha/(1-\alpha)}[\psi^{BI} - \psi^{NI}]\} \) are decreasing in \( \theta \). The fraction is increasing in \( \theta \). In other words, when \( \eta < \eta < \eta_0 \), the probability of non-integration being the organizational form with the highest profit is increasing in \( \theta \). When \( \eta < \eta_0 \) or \( \eta > \eta_0 \), \( \Pr\{\pi^{NI} = \max(\pi^{BI}, \pi^{SI}, \pi^{NI})\} \) is non-monotonic in \( \theta \) because \( \psi^{SI} - \psi^{NI} \) and \( \psi^{BI} - \psi^{NI} \) are of opposite signs.

In conclusion, when \( \eta < \eta_0 \), the probability of seller integration being the most profitable or-
ganizational form is increasing in a firm pair’s productivity; the probability of buyer integration being the most profitable organizational form is decreasing a firm pair’s productivity. When \( \eta > \bar{\eta} \), the probability of buyer integration being the most profitable organizational form is higher, the higher is the firm pair’s productivity; the probability of seller integration being the most profitable organizational form is lower, the higher is the firm pair’s productivity. When \( \eta < \eta < \bar{\eta} \), the probability of non-integration being the most profitable organizational form is increasing in the firm pair’s productivity.

Now let us examine the second component of equation (2.25), which is \( \pi^{BI} > 0 \). This is a threshold for the latent variable \( \pi^k \), but it does not affect the stochastic ordering of the profit functions for any given level of productivity, so the above statements still hold for the producing firms.

**2.F Homogeneous Model: Logit and Probit Regression Results**

Table 2.12 and Table 2.13 report the logit and probit regression results. Similar to Table 2.7, there are three dependent variables in each table, \( BI_i \), \( SI_i \) and \( VI_i \). As clearly shown in these tables, buyer firms’ relative R&D and capital intensities positively affect the pair’s probability of choosing buyer integration \( (BI_i = 1) \) and negatively affect the pair’s probability of choosing seller integration \( (SI_i = 1) \). The commonly adopted specification also receives positive support: buyer firm’s relative R&D and capital intensities positively affect the pair’s probability of choosing vertical integration \( (VI_i = 1) \). The effects are much smaller than the previous two dependent variables.
Table 2.12: Logit Regression Results

(a) R&D/Capital Intensities

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{RD}_{b,i} )</td>
<td>2.287*** (0.028)</td>
<td>2.093*** (0.028)</td>
<td>2.013*** (0.031)</td>
<td>1.794*** (0.033)</td>
<td>-1.485*** (0.027)</td>
<td>-1.155*** (0.028)</td>
<td>-1.168*** (0.031)</td>
<td>-1.045*** (0.032)</td>
<td>0.446*** (0.017)</td>
<td>0.566*** (0.018)</td>
<td>0.519*** (0.021)</td>
<td>0.450*** (0.022)</td>
</tr>
<tr>
<td>( \text{ln(emp}<em>{b,i}) - \text{ln(emp}</em>{s,i}) )</td>
<td>0.118*** (0.004)</td>
<td>0.118*** (0.004)</td>
<td>0.105*** (0.004)</td>
<td>-0.197*** (-0.003)</td>
<td>-0.201*** (-0.004)</td>
<td>-0.191*** (-0.004)</td>
<td>-0.052*** (-0.003)</td>
<td>-0.058*** (-0.003)</td>
<td>-0.057*** (-0.003)</td>
<td>-0.026*** (-0.009)</td>
<td>-0.022*** (-0.010)</td>
<td></td>
</tr>
<tr>
<td>( \text{ln(age}<em>{b,i}) - \text{ln(age}</em>{s,i}) )</td>
<td>0.081*** (0.012)</td>
<td>0.069*** (0.013)</td>
<td>0.083*** (0.013)</td>
<td>-0.174*** (-0.012)</td>
<td>-0.114*** (-0.013)</td>
<td>-0.121*** (-0.013)</td>
<td>-0.040*** (-0.009)</td>
<td>-0.026*** (-0.009)</td>
<td>-0.022*** (-0.010)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sector-pair fixed effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Region-pair fixed effects</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Observations</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.128</td>
<td>0.159</td>
<td>0.213</td>
<td>0.234</td>
<td>0.053</td>
<td>0.152</td>
<td>0.189</td>
<td>0.207</td>
<td>0.007</td>
<td>0.016</td>
<td>0.088</td>
<td>0.127</td>
</tr>
</tbody>
</table>

(b) Capital intensity

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{RD}_{b,i} )</td>
<td>3.695*** (0.040)</td>
<td>3.490*** (0.041)</td>
<td>3.942*** (0.048)</td>
<td>3.699*** (0.050)</td>
<td>-2.904*** (0.038)</td>
<td>-2.288*** (0.039)</td>
<td>-2.279*** (0.042)</td>
<td>-2.049*** (0.044)</td>
<td>0.388*** (0.024)</td>
<td>0.558*** (0.025)</td>
<td>0.730*** (0.028)</td>
<td>0.777*** (0.030)</td>
</tr>
<tr>
<td>( \text{ln(emp}<em>{b,i}) - \text{ln(emp}</em>{s,i}) )</td>
<td>0.109*** (0.004)</td>
<td>0.095*** (0.004)</td>
<td>0.087*** (0.004)</td>
<td>-0.175*** (-0.004)</td>
<td>-0.182*** (-0.004)</td>
<td>-0.179*** (-0.004)</td>
<td>-0.048*** (-0.003)</td>
<td>-0.060*** (-0.003)</td>
<td>-0.062*** (-0.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{ln(age}<em>{b,i}) - \text{ln(age}</em>{s,i}) )</td>
<td>0.108*** (0.012)</td>
<td>0.123*** (0.013)</td>
<td>0.119*** (0.014)</td>
<td>-0.208*** (-0.012)</td>
<td>-0.151*** (-0.013)</td>
<td>-0.146*** (-0.013)</td>
<td>-0.043*** (-0.009)</td>
<td>-0.016</td>
<td>-0.015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sector-pair fixed effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Region-pair fixed effects</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Observations</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.159</td>
<td>0.188</td>
<td>0.272</td>
<td>0.286</td>
<td>0.100</td>
<td>0.182</td>
<td>0.215</td>
<td>0.225</td>
<td>0.003</td>
<td>0.010</td>
<td>0.088</td>
<td>0.129</td>
</tr>
</tbody>
</table>
Table 2.13: Probit Regression Results

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: BI$_i$</th>
<th></th>
<th>Dependent Variable: SI$_i$</th>
<th></th>
<th>Dependent Variable: VI$_i$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>rd$_{b,i}$</td>
<td>1.935***</td>
<td>1.810***</td>
<td>2.030***</td>
<td>1.917***</td>
<td>-1.516***</td>
<td>-1.197***</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.024)</td>
<td>(0.025)</td>
<td>(0.020)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>ln(emp$<em>{b,i}$)-ln(emp$</em>{s,i}$)</td>
<td>0.053***</td>
<td>0.050***</td>
<td>0.046***</td>
<td>-0.093***</td>
<td>-0.095***</td>
<td>-0.093***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>ln(age$<em>{b,i}$)-ln(age$</em>{s,i}$)</td>
<td>0.053***</td>
<td>0.065***</td>
<td>0.063***</td>
<td>-0.102***</td>
<td>-0.080***</td>
<td>-0.078***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Sector-pair fixed effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Region-pair fixed effects</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Observations</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.155</td>
<td>0.179</td>
<td>0.262</td>
<td>0.276</td>
<td>0.098</td>
<td>0.175</td>
</tr>
</tbody>
</table>

(a) R&D intensity

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: BI$_i$</th>
<th></th>
<th>Dependent Variable: SI$_i$</th>
<th></th>
<th>Dependent Variable: VI$_i$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>rd$_{b,i}$</td>
<td>1.202***</td>
<td>1.089***</td>
<td>1.059***</td>
<td>0.952***</td>
<td>-0.779***</td>
<td>-0.577***</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.016)</td>
<td>(0.017)</td>
<td>(0.013)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>ln(emp$<em>{b,i}$)-ln(emp$</em>{s,i}$)</td>
<td>0.060***</td>
<td>0.061***</td>
<td>0.055***</td>
<td>-0.105***</td>
<td>-0.105***</td>
<td>-0.100***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>ln(age$<em>{b,i}$)-ln(age$</em>{s,i}$)</td>
<td>0.038***</td>
<td>0.036***</td>
<td>0.043***</td>
<td>-0.084***</td>
<td>-0.062***</td>
<td>-0.065***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Sector-pair fixed effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Region-pair fixed effects</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Observations</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
<td>85783</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.128</td>
<td>0.155</td>
<td>0.209</td>
<td>0.229</td>
<td>0.053</td>
<td>0.147</td>
</tr>
</tbody>
</table>

(b) Capital intensity
2.G Heterogeneous Model: Quintile and Seventh-Quantile Regression Results

Table 2.14 and Table 2.15 report the regression results for the quintile and 7th quantile regression results with pair-specific R&D intensity measures. Table 2.16 and Table 2.17 report the regression results for the quintile and 7th quantile regression results with pair-specific R&D intensity measures.

Table 2.14 and Table 2.15 display similar patterns as Table 2.10. In the left panel, at lower quantiles of $RD$, more productive firms are more likely to choose seller integration and less likely to choose buyer integration; at higher quantiles of $RD$, more productive firms are more likely to choose buyer integration and less likely to choose seller integration. In the right panel, at lower quantiles of $(1 - RD)$, more productive firms are more likely to choose buyer integration and less likely to choose seller integration. At higher quantiles of $(1 - RD)$, more productive firms are more likely to choose seller integration and less likely to choose buyer integration.
Table 2.14: Regression Results with Quintiles of Pair-specific $RD$ and $(1 - RD)$

(a) Buyer’s R&D intensity

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BI$_i$</th>
<th>SI$_i$</th>
<th>NI$_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod,$^*$RD$_1$</td>
<td>-0.243***</td>
<td>0.152***</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.039)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Prod,$^*$RD$_2$</td>
<td>-0.123***</td>
<td>-0.188***</td>
<td>0.311***</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.040)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Prod,$^*$RD$_3$</td>
<td>-0.009</td>
<td>-0.240***</td>
<td>0.249***</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.040)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Prod,$^*$RD$_4$</td>
<td>0.177**</td>
<td>-0.267***</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.050)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Prod,$^*$RD$_5$</td>
<td>0.131***</td>
<td>-0.089*</td>
<td>-0.042</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.045)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>ln(age$<em>{b,i}$)-ln(age$</em>{s,i}$)</td>
<td>0.040***</td>
<td>-0.054***</td>
<td>0.014***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Observations 85783 85783 85783
R-squared 0.056 0.066 0.007

(b) Seller’s R&D intensity

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BI$_i$</th>
<th>SI$_i$</th>
<th>NI$_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod,$^*$RD$_1$</td>
<td>0.162***</td>
<td>-0.179***</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.038)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Prod,$^*$RD$_2$</td>
<td>-0.219***</td>
<td>-0.053</td>
<td>0.272***</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.037)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Prod,$^*$RD$_3$</td>
<td>-0.279***</td>
<td>-0.006</td>
<td>0.285***</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.061)</td>
</tr>
<tr>
<td>Prod,$^*$RD$_4$</td>
<td>-0.385***</td>
<td>0.218***</td>
<td>0.167**</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.057)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>Prod,$^*$RD$_5$</td>
<td>-0.285***</td>
<td>0.181***</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.044)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>ln(age$<em>{b,i}$)-ln(age$</em>{s,i}$)</td>
<td>0.036***</td>
<td>-0.055***</td>
<td>0.019***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Observations 85783 85783 85783
R-squared 0.078 0.057 0.009
### Table 2.15: Regression Results with 7 Quantiles of Pair-specific $RD$ and $(1 - RD)$

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BI$_i$</th>
<th>SI$_i$</th>
<th>NI$_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod*$RD_1$</td>
<td>-0.251***</td>
<td>0.175***</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.041)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Prod*$RD_2$</td>
<td>-0.146***</td>
<td>-0.148***</td>
<td>0.293***</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.038)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>Prod*$RD_3$</td>
<td>-0.106**</td>
<td>-0.212***</td>
<td>0.318***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.042)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Prod*$RD_4$</td>
<td>0.009</td>
<td>-0.243***</td>
<td>0.234***</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.042)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Prod*$RD_5$</td>
<td>0.139*</td>
<td>-0.263***</td>
<td>0.124*</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.057)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Prod*$RD_6$</td>
<td>0.169**</td>
<td>-0.181***</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.049)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Prod*$RD_7$</td>
<td>0.120*</td>
<td>-0.071</td>
<td>-0.049</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.043)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>ln(age$<em>{bi}$)-ln(age$</em>{si}$)</td>
<td>0.040***</td>
<td>-0.053***</td>
<td>0.014***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Observations: 85783
R-squared: 0.056

### Table 2.16: Regression Results with Quintiles of industry-specific $RD$ and $(1 - RD)$

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BI$_i$</th>
<th>SI$_i$</th>
<th>NI$_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod*$RD_1$</td>
<td>-0.315***</td>
<td>0.073</td>
<td>0.241**</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.043)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>Prod*$RD_2$</td>
<td>-0.192***</td>
<td>0.046</td>
<td>0.146</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.043)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>Prod*$RD_3$</td>
<td>0.105</td>
<td>0.164**</td>
<td>-0.269*</td>
</tr>
<tr>
<td></td>
<td>(0.068)</td>
<td>(0.054)</td>
<td>(0.114)</td>
</tr>
<tr>
<td>Prod*$RD_4$</td>
<td>-0.025</td>
<td>0.011</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.044)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>Prod*$RD_5$</td>
<td>0.196***</td>
<td>-0.146***</td>
<td>-0.050</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.041)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>ln(age$<em>{bi}$)-ln(age$</em>{si}$)</td>
<td>0.045***</td>
<td>-0.067***</td>
<td>0.022***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
</tbody>
</table>

Observations: 85783
R-squared: 0.065
Table 2.17: Regression Results with 7 Quantiles of industry-specific $RD$ and $(1 - RD)$

(a) Buyer’s R&D intensity

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BI$_i$</th>
<th>SI$_i$</th>
<th>NI$_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod$_i$*RD$_1$</td>
<td>-0.310***</td>
<td>0.074</td>
<td>0.236**</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.046)</td>
<td>(0.072)</td>
</tr>
<tr>
<td>Prod$_i$*RD$_2$</td>
<td>-0.210***</td>
<td>0.090*</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.041)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Prod$_i$*RD$_3$</td>
<td>-0.168***</td>
<td>0.023</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.045)</td>
<td>(0.082)</td>
</tr>
<tr>
<td>Prod$_i$*RD$_4$</td>
<td>0.180**</td>
<td>0.197***</td>
<td>-0.377***</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.053)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>Prod$_i$*RD$_5$</td>
<td>-0.181**</td>
<td>0.022</td>
<td>0.159</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.045)</td>
<td>(0.094)</td>
</tr>
<tr>
<td>Prod$_i$*RD$_6$</td>
<td>0.100*</td>
<td>-0.007</td>
<td>-0.093</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.043)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Prod$_i$*RD$_7$</td>
<td>0.247***</td>
<td>-0.172***</td>
<td>-0.075</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.040)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>ln(age$<em>{b,i}$)-ln(age$</em>{s,i}$)</td>
<td>0.047***</td>
<td>-0.066***</td>
<td>0.019***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

(b) Seller’s R&D intensity

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BI$_i$</th>
<th>SI$_i$</th>
<th>NI$_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod$_i$*(1-RD$_1$)</td>
<td>0.243***</td>
<td>-0.169***</td>
<td>-0.073</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.040)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Prod$_i$*(1-RD$_2$)</td>
<td>0.093*</td>
<td>-0.007</td>
<td>-0.086</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.043)</td>
<td>(0.074)</td>
</tr>
<tr>
<td>Prod$_i$*(1-RD$_3$)</td>
<td>0.142*</td>
<td>0.200***</td>
<td>-0.341**</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.052)</td>
<td>(0.104)</td>
</tr>
<tr>
<td>Prod$_i$*(1-RD$_4$)</td>
<td>-0.036</td>
<td>-0.004</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.056)</td>
<td>(0.094)</td>
</tr>
<tr>
<td>Prod$_i$*(1-RD$_5$)</td>
<td>-0.183***</td>
<td>0.017</td>
<td>0.166*</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.044)</td>
<td>(0.081)</td>
</tr>
<tr>
<td>Prod$_i$*(1-RD$_6$)</td>
<td>-0.205***</td>
<td>0.096*</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.041)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Prod$_i$*(1-RD$_7$)</td>
<td>-0.314***</td>
<td>0.079</td>
<td>0.235**</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.046)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>ln(age$<em>{b,i}$)-ln(age$</em>{s,i}$)</td>
<td>0.046***</td>
<td>-0.066***</td>
<td>0.020***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

Observations | 85783 | 85783 | 85783 |
R-squared    | 0.074 | 0.060 | 0.043 |
Chapter 3

Judicial Quality, Relative Efficiency, and Firm Boundary Decisions
Abstract

This paper explores the impact of judicial quality on firms’ organization decisions. I construct a model where heterogeneous buyer and seller firms form production pairs. Each firm-pair chooses one of three ownership structures: buyer integration of seller, seller integration of buyer, or outsourcing. I show that in equilibrium, two factors affect a firm-pair’s organization decision: relative judicial quality and relative efficiency. Firms allocate ownership over assets to the more efficient party because it generates a higher marginal return. This is the typical Grossman-Hart mechanism. Judicial quality acts as a substitute to integration. A country with better legal institutions provides better protection to firms’ non-contractible, relationship-specific investments. Such protection reduces firms’ need to protect their investments via integration. As a result, firms from countries with better legal institutions are less likely to be integrated. I find empirical evidence for both relative efficiency and relative judicial quality.
3.1 Introduction

This paper studies how judicial quality affects firms' organization decisions. To develop a testable hypothesis, I construct a trade model with heterogeneous buyers (downstream producers) and sellers (upstream suppliers). Buyers and sellers form one-to-one production pairs and produce a final product that requires both firms to make non-contractible, relationship-specific investments. Prior to making these investments, each firm-pair chooses from one of three organizational forms: buyer integration of the seller, seller integration of the buyer, or outsourcing. I show that in equilibrium, a firm-pair’s organizational choice depends on two factors: relative efficiency and relative judicial quality.

Relative efficiency is defined as the ratio between the buyer and the seller’s cost-adjusted productivity. It works through the classic Grossman-Hart-Moore property-rights channel. With high relative efficiency, the buyer’s investment generates a higher marginal return to the firm-pair than the seller’s investment. The firm-pair allocates residual control rights to the buyer by choosing buyer integration (of the seller) because this ownership structure increases the buyer’s investment incentive.

With low relative efficiency, the seller’s investment generates a higher marginal return than the buyer’s investment. To increase the seller’s investment incentive, the firm-pair allocates residual control rights to the seller by choosing seller integration (of the buyer). Since both types of integration incentivize one party at the expense of the other party’s incentive, in the intermediate case, the firm-pair choose outsourcing so that neither firm’s investment incentive is distorted by too much. As a result, a firm-pairs’ organization decision can be roughly expressed as a function of its relative efficiency, as shown in Figure 3.1.

Relative judicial quality is defined as the ratio between the judicial qualities in the countries where the buyer and the seller reside. It works through a more subtle, yet well-established channel. A country with better legal institutions provides better protection for a firm’s residual

---

1 I assume that all investments are non-contractible and relationship-specific so firms cannot write a contract specifying their investment levels. They can instead write a contract allocating ownership over these investments. Ownership is defined as the purchase of residual control rights.

2 Under buyer integration of the seller, the buyer has residual control rights over both firms’ investments. In the event of a negotiation failure, the buyer can retrieve both firms’ investments and generate a discounted level of output, while the seller is left with nothing. With a higher outside option, the buyer is able to negotiate a higher revenue share. The higher revenue share creates an incentive for the buyer to choose a higher investment level.

3 Many studies have shown that institutional differences are an important determinant of international trade. Levchenko (2007), Nunn (2007), Costinot (2009), and Nunn and Trefler (2014) study the relationship contracting institutions and international trade. Beck (2003) and Manova (2013) study the relationship between financial institutions and international trade. Cunat and Melitz (2012) study the relationship between labor market flexibility and international trade. Perhaps the most related paper to my work is Bernard et al. (2010a). The authors
control rights. Consider a firm-pair with the buyer coming from a country with good legal institutions and the seller from a country with bad legal institutions. The buyer firm’s residual control rights are protected by the seller country’s legal institution and the seller firm’s residual control rights are protected by the buyer country’s legal institution. Given the difference in the two countries’ judicial qualities, the buyer’s residual control rights receive poorer protection than the seller’s. As a result, buyer integration provides a smaller incentive boost for the buyer than seller integration does for the seller. As shown in Figure 3.1, an increase in relative judicial quality pushes down the relative efficiency threshold for buyer integration and pushes up the relative efficiency threshold for seller integration. Keeping everything else constant, an increase in relative judicial quality increases a firm-pair’s probability of choosing buyer integration and decreases its probability of choosing seller integration.

The same logic applies at the country-level. A country with good legal institutions provides better protection to firms’ residual control rights than a country with bad legal institutions. This reduces the likelihood of its firms being integrated by foreign firms. Effectively, judicial quality acts as a substitute for integration. This intuition is illustrated by Figure 3.2.

In Figure 3.2, all buyer-seller relationships are divided into 4 groups based on the develop... study the relationship between contractibility and within-firm trade. Also, see Chor (2010) for a comprehensive examination of the various determinants of comparative advantage.
Chapter 3. Judicial Quality, Relative Efficiency, and Firm Boundary Decisions

Figure 3.2: Distribution of Organizational Forms across Country Pairs

The data used to plot this figure is from the same database that I use for my empirical exercise. There are a total of 1,659,888 buyer-seller relationships. Starting from the upper-left figure, in clockwise order, there are respectively 180,629, 72,152, 90,431, and 1,316,676 buyer-seller relationships in each figure. Due to data availability, there are fewer observations in the regression sample than in this figure.

The development status of their countries of residence. Developed countries have higher judicial quality measures than developing countries. For each group, I plot the distribution of various ownership structures. “b”, “n” and “s” stand for buyer integration of the seller, non-integration, and seller integration of the buyer, respectively. In the upper-left and lower-right figures, buyers and sellers come from similar countries. There is no significant difference in the percentages of buyer and seller integration. When the buyer comes from a developed country and the seller from a less developed country (upper-right figure), buyer integration dominates seller integration. The opposite is true when the seller comes from a developed country and the buyer from a less developed country (lower-left figure). Figure 3.2 suggests that at the country-level, firms from countries with high judicial qualities are more likely to become the “integrators”, while firms from countries with low judicial qualities are more likely to become the “integratees”.

I empirically test the hypotheses of my model regarding relative efficiency and relative judicial quality using a similar database as Chapter 2. The empirical results are consistent with the hypotheses.

The rest of this paper is organized as follows. In Section 3.2, I set up the previously
introduced model and derive the industry equilibrium of this model. In Section 3.3, I test two predictions of the model: one on relative efficiency, the other on judicial quality. Section 3.4 concludes the paper.

3.2 Model

3.2.1 Setup

There are \( N \) identical countries and \( J + 1 \) final goods industries. The representative consumer’s utility is

\[
U = Q_0 + \frac{1}{\mu} \sum_{j=1}^{J} Q_j^\mu, \quad 0 < \mu < 1, \tag{3.1}
\]

where \( Q_0 \) is the consumption level of a numeraire good whose price is set to 1, and \( 1/(1-\mu) \) is the cross-industry elasticity of substitution. \( Q_j \) is the aggregate consumption of the differentiated goods in industry \( j \):

\[
Q_j = \left[ \int a_{ij} \, di \right]^{1/\alpha}, \quad 0 < \alpha < 1, \tag{3.2}
\]

where \( a_{ij} \) is the consumption level of variety \( i \) in industry \( j \). \( 1/(1-\alpha) \) is the within-industry elasticity of substitution. Because I do not solve for the general equilibrium of this model, there is no need to solve for the specific values of \( Q_j \). I assume that \( Q_j = 1 \) for \( j = 1, ..., N \).

Given the above preference function, the inverse demand function for variety \( i \) in industry \( j \) can be derived as

\[
p_{ij} = a_{ij}^{\alpha-1}. \tag{3.3}
\]

The final good of each industry is produced using inputs from two industries, an upstream industry and a downstream industry. For simplicity, assume that there is no input-output linkage between the upstream and downstream industries across different industries. Each industry is occupied by a continuum of upstream and downstream producers. Each upstream industry is occupied by a continuum of upstream suppliers, and each downstream industry by a continuum of downstream producers. The upstream suppliers produce an intermediate good, which they transfer to the downstream producers, who then produce and sell the final goods to consumers.

\footnote{For models with input-output linkages, see Acemoglu et al. (2012), Antràs and Chor (2013), and Costinot et al. (2013).}
From here on, I will make several notational simplifications. I use subscripts $i$, $b$, and $s$ as indexes for a pair, a buyer, and a seller, respectively. $i$ can be thought of as a tuple: $i \equiv (b, s)$. For simplicity, I refer to the upstream suppliers as the sellers, and the downstream producers as buyers. The production process in industry $j$ (ignoring the industry subscript $j$) goes as follows:

1. Buyers and sellers (in industry $j$ and each country) pay an entry cost $f_E$ to draw their productivity levels $\theta_b$ and $\theta_s$ from Pareto distributions $G^B(\theta)$ and $G^S(\theta)$.

2. Based on their productivity levels, buyers and sellers decide whether to exit the market. Exiting firms receive 0 payoffs and staying firms enter the next step.

3. To simplify the process of matching buyers with sellers I model a very simple form of one-sided (buyer) search. A buyer draws a set of sourcing costs, one for each country. Let $m_l^b$ denote buyer $b$’s sourcing cost for country $l$, $l \in \{1, 2, ..., N\}$. $m_l^b$ can be thought of as buyer $b$’s cost of being matched with a seller in country $l$. Given $\{m_1^b, ..., m_N^b\}$, the buyer chooses a single sourcing country. Denote buyer $b$’s optimal sourcing decision by $l_b^*$. Buyer $b$ is then randomly matched with a seller from country $l_b^*$ and the productivity of each partner ($\theta_b$ and $\theta_s$) is revealed to the other partner.

4. Immediately after being matched the pair costlessly draws a set of fixed costs of production, one for each organizational form. Denote these fixed costs by $f_{ki}^i$ where $i$ is a firm-pair index and $k \in \{BI, SI, NI\}$ is an index of organizational form. The pair then negotiates a pair $\{k_{i}^*, t_{i}^*\}$ where $k_{i}^*$ is the choice of organization form and $t_{i}^*$ is an ex ante transfer from the buyer to the seller.\footnote{If $t_{i}^* < 0$ then the seller pays the buyer. It does not matter who pays the fixed cost $f_{ki}^*$ or whether the buyer and seller share this cost because whatever the rule for paying the fixed cost, it is subsumed in $t_{i}^*$.}

5. The buyer and seller each choose their levels of investments $x_b$ and $x_s$. $x_b$ and $x_s$ are non-contractible and relationship-specific, which means that the pair cannot write a contract specifying the values of $x_b$ and $x_s$. The marginal costs of investments are $c_b$ for the buyer and $c_s$ for the seller.

6. After $x_b$ and $x_s$ are chosen, the buyer and seller bargain over revenue division. The outside options in this step depend on the organizational form $k$ chosen in step 5. If the two firms fail to reach an agreement, then the game ends with both firms getting their respective
outside options; if the firms reach an agreement, then the game proceeds to the next step. Final goods are produced according to the following CES production function

\[ q_i = \left[ (\theta_b x_b)^\rho + (\theta_s x_s)^\rho \right]^{1/\rho}, \tag{3.4} \]

where \( q_i \) is the output level of final good \( i \). \( \theta_b \) and \( \theta_s \) are the productivity levels of the buyer and seller. \( 1/(1 - \rho) \) is the elasticity of substitution between buyer and seller inputs. Revenues are divided according to the results in step 6.

### 3.2.2 Industry Equilibrium

In this section, I solve for the partial equilibrium for industry \( j \).

**Step 6**

From the inverse demand function in equation (3.3) and the production function in equation (3.4), pair \( i \) faces a revenue function of

\[ R_i = p_i q_i = q_i^\alpha. \tag{3.5} \]

Substituting in the production function in equation (3.4), the revenue function can be expressed as

\[ R_i = q_i^\alpha = \left[ (\theta_b x_b)^\rho + (\theta_s x_s)^\rho \right]^{\alpha/\rho}. \tag{3.6} \]

The two firms engage in Nash bargaining over the revenue as specified in equation (3.6). The bargaining weights are \( \beta_i \) for the buyer and \( (1 - \beta_i) \) for the seller. The outside options are endogenously determined by the organizational form \( k_i^* \) as chosen in step 4.

In this step, the two firms take \( k_i^* \) as given. I discuss the bargaining results in three scenarios, one for each organizational form.

If \( k_i^* = BI \), the buyer integrates the seller. I define ownership in the same sense as in Grossman and Hart (1986). Under buyer integration, property rights over the buyer and seller’s investments are allocated to the buyer. If the two firms fail to reach an agreement, then the buyer has residual control rights over the buyer and seller’s assets. This means the buyer gets to seize both \( x_b \) and \( x_s \). Suppose the buyer can use these investments to produce a certain amount of output, but since he does not possess the seller’s specific knowledge on how to
utilize \( x_b \), the buyer can only produce a fraction \( \delta_b \) of output, then the buyer’s outside option under \( k^*_i = BI \) is the revenue generated by \( \delta_b q_i \), which is \((\delta_b q_i)^\alpha = \delta_b^\alpha R_i \). On the other hand, if the two firms fail to reach an agreement, then the seller loses everything. In this case, the buyer’s surplus is his outside option \( \delta_b^\alpha R_i \), plus his bargaining weight \( \beta_i \) times total Nash surplus \((R_i - \delta_b^\alpha R_i - 0)\), so the buyer’s surplus in this step is \( \delta_b^\alpha R_i + \beta_i(R_i - \delta_b^\alpha R_i - 0) = [\beta_i + (1 - \beta_i)\delta_b^\alpha]R_i \). The seller’s surplus is her outside option 0 plus her share of Nash surplus, which adds up to 0 + \((1 - \beta_i)(R_i - \delta_b^\alpha R_i - 0)\) = \( (1 - \beta_i)(1 - \delta_b^\alpha)R_i \).

If \( k^*_i = SI \), pair \( i \) allocates all property rights to the seller. If the two firms fail to reach an agreement, the seller seizes \( x_b \) and \( x_s \), and produces a fraction \( \delta_s \) of output. The seller’s outside option is \((\delta_s q_i)^\alpha = \delta_s^\alpha R_i \). On the other hand, the buyer’s outside option is 0, because the buyer is left with nothing in the event of bargaining failure. Under seller integration of the buyer, the seller’s Nash surplus is \( \delta_s^\alpha R_i + (1 - \beta_i)(R_i - \delta_s^\alpha R_i - 0) = (1 - \beta_i + \beta_i\delta_s^\alpha)R_i \). The buyer’s Nash surplus is 0 + \(\beta_i(R_i - \delta_s^\alpha R_i - 0)\) = \( \beta_i(1 - \delta_s^\alpha)R_i \).

If \( k^*_i = NI \), the buyer and seller each retain property rights over their own investments. In the event of bargaining failure, the buyer gets \( x_b \) and the seller gets \( x_s \). Without the other firm’s input, each firm cannot produce any final product, so both firms’ outside options are 0. The buyer’s Nash surplus is 0 + \(\beta_i(R_i - 0 - 0)\) = \( \beta_i R_i \). The seller’s Nash surplus is 0 + \((1 - \beta_i)(R_i - 0 - 0)\) = \( (1 - \beta_i)R_i \).

From the above analysis, it is clear that the bargaining results under each organizational form can be categorized as a division of revenue. Denote the buyer’s revenue share under organizational form \( k \) by \( \beta^k_i \), then the seller’s revenue share is \((1 - \beta^k_i)\). Table 3.1 summarizes the buyer and seller’s revenue shares under each organizational form \( k \in \{BI, SI, NI\} \).

<table>
<thead>
<tr>
<th>( k = BI )</th>
<th>( k = SI )</th>
<th>( k = NI )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buyer’s revenue share ( (\beta^k_i) )</td>
<td>( \beta_i + (1 - \beta_i)\delta_b^\alpha )</td>
<td>( \beta_i(1 - \delta_s^\alpha) )</td>
</tr>
<tr>
<td>Seller’s revenue share ( (1 - \beta^k_i) )</td>
<td>( (1 - \beta_i)(1 - \delta_b^\alpha) )</td>
<td>( 1 - \beta_i + \beta_i\delta_s^\alpha )</td>
</tr>
</tbody>
</table>
Step 5

In this step, firms choose their investment levels to maximize their expected profits. The buyer’s problem is

$$\max_{x_b} \beta_i^k R_i - c_b x_b,$$  \hspace{1cm} (3.7)$$

and the seller’s problem is

$$\max_{x_s} (1 - \beta_i^k) R_i - c_s x_s.$$  \hspace{1cm} (3.8)$$

$R_i$ is as defined in equation (3.6). The first order conditions for the buyer and seller’s problems can be written as

$$c_b x_b = \alpha \frac{(\theta_b x_b)^\rho}{(\theta_b x_b)^\rho + (\theta_s x_s)^\rho} \beta_i^k R_i,$$  \hspace{1cm} (3.9)$$

and

$$c_s x_s = \alpha \frac{(\theta_s x_s)^\rho}{(\theta_b x_b)^\rho + (\theta_s x_s)^\rho} (1 - \beta_i^k) R_i.$$  \hspace{1cm} (3.10)$$

$R_i$ is as defined in equation (3.6). Solving equations (3.9) and (3.10) for $x_b$ and $x_s$ gives the optimal investment levels:

$$x_b^k(\theta_b, \theta_s) = \alpha^{1/(1-\alpha)} \frac{\left( \frac{\beta_i^k \theta_b}{c_b} \right)^{\rho/(1-\rho)} \beta_i^k R_i}{\left( \frac{\beta_i^k \theta_b}{c_b} \right)^{\rho/(1-\rho)} + \left[ (1 - \beta_i^k) \theta_s / c_s \right]^{\rho/(1-\rho)}}^{(\rho-\alpha)/(\rho(1-\alpha))},$$  \hspace{1cm} (3.11)$$

and

$$x_s^k(\theta_b, \theta_s) = \alpha^{1/(1-\alpha)} \frac{\left( \frac{(1 - \beta_i^k) \theta_s}{c_s} \right)^{\rho/(1-\rho)} }{\left( \frac{\beta_i^k \theta_b}{c_b} \right)^{\rho/(1-\rho)} + \left[ (1 - \beta_i^k) \theta_s / c_s \right]^{\rho/(1-\rho)}}^{(\rho-\alpha)/(\rho(1-\alpha))}. \hspace{1cm} (3.12)$$

Substituting the solutions to $x_b$ and $x_s$ back into equation (3.6) gives the following revenue function:

$$R_i^k(\theta_b, \theta_s) = \alpha \left[ \left( \frac{\beta_i^k \theta_b}{c_b} \right)^{\rho/(1-\rho)} + \left( \frac{(1 - \beta_i^k) \theta_s}{c_s} \right)^{\rho/(1-\rho)} \right]^{(1-\rho)/\rho} \frac{\alpha}{1-\alpha}.$$  \hspace{1cm} (3.13)$$
The buyer and seller’s joint surplus in this step is then

$$\pi^k_i(\theta_b, \theta_s) = R^k_i(\theta_b, \theta_s) - c_b x^k_b(\theta_b, \theta_s) - c_s x^k_s(\theta_b, \theta_s)$$

$$= \psi^k_i(\theta_b, \theta_s) - f^k_i,$$  \hspace{1cm} (3.14)

where

$$\psi^k_i(\theta_b, \theta_s) = \alpha^{1-\alpha} \left(1 - \alpha \beta^k_i \right) \left( \beta^k_i \theta^k_b/c_b \right)^{\rho/(1-\rho)} + \left[1 - \alpha (1 - \beta^k_i) \right] \left[ (1 - \beta^k_i) \theta^k_s/c_s \right]^{\rho/(1-\rho)}$$

$$\left\{ (\beta^k_i \theta^k_b/c_b)^{\rho/(1-\rho)} + [(1 - \beta^k_i) \theta^k_s/c_s]^{\rho/(1-\rho)} \right\}^{(\rho-\alpha)/\rho(1-\alpha)}$$  \hspace{1cm} (3.15)

is the variable profit, and $f^k_i$ is the fixed cost of production for pair $i$ under organizational form $k$. It is drawn from a stochastic distribution. The specific formulation of $f^k_i$ will be explained later.

Of the profit in equation (3.14), the buyer gets $\pi^k_b(\theta_b, \theta_s) = \beta^k_i R^k_i(\theta_b, \theta_s) - c_b x^k_b(\theta_b, \theta_s)$ or

$$\pi^k_b(\theta_b, \theta_s) \equiv \alpha^{\alpha/(1-\alpha)} \beta^k_i \frac{(1 - \alpha)(\beta^k_i \theta^k_b/c_b)^{\rho/(1-\rho)} + [(1 - \beta^k_i) \theta^k_s/c_s]^{\rho/(1-\rho)}}{\left\{ (\beta^k_i \theta^k_b/c_b)^{\rho/(1-\rho)} + [(1 - \beta^k_i) \theta^k_s/c_s]^{\rho/(1-\rho)} \right\}^{(\rho-\alpha)/\rho(1-\alpha)}} - f^k_b,$$  \hspace{1cm} (3.16)

and the seller gets

$$\pi^k_s(\theta_b, \theta_s) \equiv \alpha^{\alpha/(1-\alpha)} (1 - \beta^k_i) \frac{(\beta^k_i \theta^k_b/c_b)^{\rho/(1-\rho)} + (1 - \alpha)(1 - \beta^k_i) \theta^k_s/c_s]^{\rho/(1-\rho)}}{\left\{ (\beta^k_i \theta^k_b/c_b)^{\rho/(1-\rho)} + [(1 - \beta^k_i) \theta^k_s/c_s]^{\rho/(1-\rho)} \right\}^{(\rho-\alpha)/\rho(1-\alpha)}} - f^k_s.$$  \hspace{1cm} (3.17)

$f^k_b$ and $f^k_s$ are the fixed costs of production that are paid by the buyer and the seller, respectively. I require

$$f^k_b + f^k_s = f^k_i, \quad f^k_b \geq 0, f^k_s \geq 0.$$  

**Step 4**

In step 4, the buyer and seller bargain over their organizational form. They will choose the organizational form $k$ that maximizes their joint surplus as defined in equation (3.14). In other words, the optimal organizational form $k^*$ is solved by the following problem:

$$k^* = \arg \max_{k \in \{BI, SI, NI\}} \pi^k_i(\theta_b, \theta_s).$$  \hspace{1cm} (3.18)
\( \pi_i^k(\theta_b, \theta_s) \) is defined in equation (3.14). As in Antràs and Helpman (2008), this problem can be indirectly solved by extending \( \beta_i^k \) to be a continuous variable \( \beta_i \), with \( \beta_i \in (0, 1) \), and maximizing profit by choosing the optimal \( \beta_i \). Since \( \beta_i^k \) enters \( \pi_i^k(\theta_b, \theta_s) \) only through variable profits \( \psi_i^k(\theta_b, \theta_s) \), this maximization problem can be written as

\[
\beta_i^* \equiv \arg \max_{\beta \in (0, 1)} \psi_i(\beta, \theta_b, \theta_s).
\] (3.19)

where

\[
\psi_i(\beta, \theta_b, \theta_s) = \alpha \frac{(1 - \alpha \beta)(\theta_b/c_b)^{\rho/(1-\rho)} + [1 - \alpha (1 - \beta)] [(1 - \beta)\theta_s/c_s]^{\rho/(1-\rho)}}{((\theta_b/c_b)^{\rho/(1-\rho)} + [(1 - \beta)\theta_s/c_s]^{\rho/(1-\rho)})}.
\] (3.20)

Given the particular functional form, there is no explicit solution to \( \beta_i^* \). However, define the ‘relative efficiency’ of pair \( i \) as

\[
\tilde{\theta}_i \equiv \frac{(\theta_b/c_b)^{\rho/(1-\rho)}}{(\theta_s/c_s)^{(\rho/(1-\rho))}}.
\] (3.21)

The variable profit \( \psi_i(\beta, \theta_b, \theta_s) \) then satisfies the following lemma:

**Lemma 5.** \( \psi_i(\beta, \theta_b, \theta_s) \) is supermodular in \((\beta, \tilde{\theta}_i)\).

**Theorem 7.** The following relationship exists between \( \beta_i^* \) and \( \tilde{\theta}_i \):

\[
\tilde{\theta}_i = \frac{(1 - \alpha) \left[ \left( \frac{\beta_i^*}{1-\beta_i^*} \right)^2 - 1 \right] \pm \sqrt{(1 - \alpha)^2 \left[ \left( \frac{\beta_i^*}{1-\beta_i^*} \right)^2 - 1 \right]^2 - 4 (1 - \rho)^2 \left( \frac{\beta_i^*}{1-\beta_i^*} \right)^{\rho/(1-\rho)}}}{2 (1 - \rho) \left( \frac{\beta_i^*}{1-\beta_i^*} \right)^{\rho/(1-\rho)}}.
\] (3.22)

Further,

1. \( \beta_i^* \) is strictly increasing in \( \tilde{\theta}_i \);
2. \( \lim_{\tilde{\theta}_i \to 0} \beta_i^* = 0 \);
3. \( \lim_{\tilde{\theta}_i \to +\infty} \beta_i^* = 1 \).

I prove Lemma 7 in Appendix 3.B. Figure 3.3 provides a simulation of equation (3.22).

The intuition behind Theorem 7 is simple. Recall that in this step, the buyer and seller are choosing the organizational form to maximize joint profits. When \( \tilde{\theta}_i \) is high, the buyer’s productivity is relatively high so that the firm-pair would like to provide high incentives to the
buyer by choosing a high $\beta^*$. That is, a high $\tilde{\theta}_i$ leads to a high $\beta^*_i$. When $\tilde{\theta}_i$ is low, the opposite is true.

Theorem 7 implies that as $\tilde{\theta}_i$ increases, pair $i$ needs a larger $\beta^*_k$ to maximize variable profits $\psi^k_i$. However, $\beta^*_k$ is a discrete choice variable. Pair $i$’s choice of $\beta^*_k$ is thus a step function of $\tilde{\theta}_i$: there are two threshold values for $\tilde{\theta}_i$ that mark the jump in pair $i$’s choice of $k$ that maximizes variable profit $\psi^k_i$, as described in the following theorem:

**Theorem 8.** For each pair $i$ of firms, there are two threshold values of $\tilde{\theta}_i$, $\tilde{\theta}_i^H$ and $\tilde{\theta}_i^L$, with $\tilde{\theta}_i^H > \tilde{\theta}_i^L$, such that

$$\max_{k \in \{BI, SI, NI\}} \psi^k_i(\theta_b, \theta_s) = \begin{cases} 
\psi^SI_i(\theta_b, \theta_s), & \text{if } \tilde{\theta}_i \in (0, \tilde{\theta}_i^L]; \\
\psi^NI_i(\theta_b, \theta_s), & \text{if } \tilde{\theta}_i \in (\tilde{\theta}_i^L, \tilde{\theta}_i^H); \\
\psi^{BI}_i(\theta_b, \theta_s), & \text{if } \tilde{\theta}_i \in [\tilde{\theta}_i^H, \infty).
\end{cases}$$

A proof of Theorem 8 is provided in Appendix 3.C.

The value of $k$ that maximizes variable profits $\psi^k_i$ is not necessarily the $k$ that maximizes joint profit $\pi^k_i$, because $\pi^k_i$ is composed of two parts: variable profits $\psi^k_i$, and fixed cost $f^k_i$. 
Two pairs of firms that have the same level of relative efficiency \( \tilde{\theta} \) do not necessarily choose the same organizational form because they might have different draws on the fixed costs of production. However, when \( f^k_i \) follows a Gumbel distribution, I can obtain stochastic properties of a firm-pair’s organizational choice. Suppose \( f^k_i \) is composed of two parts:

\[
f^k_i = f^i_k + \varepsilon^k_i,
\]

where \( f > 0 \) is the fixed part, and \( \varepsilon^k_i \) is the stochastic part that follows a Gumbel distribution:

\[
\varepsilon^k_i \sim \text{Gumbel}(0, \sigma), \quad \sigma > 0.
\]

Given the distribution of \( f^k_i \), as McFadden (1973) shows, the probability of pair \( i \) choosing \( k \) as its optimal organizational form can be expressed as

\[
\Pr\{k^*_i = k\} = \frac{\exp\{\psi^k_i(\theta_b, \theta_s)\}}{\sum_{k' \in \{BI, SI, NI\}} \exp\{\psi^{k'}_i(\theta_b, \theta_s)\}},
\]

where \( k^*_i \) denotes pair \( i \)’s chosen organizational form, as defined in equation (3.18). \( \psi^k_i(\theta_b, \theta_s) \) and \( \psi^{k'}_i(\theta_b, \theta_s) \) are as defined in equation (3.15).

Note that \( \psi^k_i(\theta_b, \theta_s) \) in equation (3.23) is a discrete version of \( \psi_i(\beta, \theta_b, \theta_s) \) from equation (3.20). The continuous variable in \( \psi_i(\beta, \theta_b, \theta_s) \) take three discrete values \( \beta^k_i, \beta^{SI}_i, \) and \( \beta^{NI}_i \), with \( \beta^{SI}_i < \beta^{NI}_i < \beta^k_i \). The supermodularity of \( \psi_i(\beta, \theta_b, \theta_s) \) in \( (\beta, \tilde{\theta}_i) \) implies that \( \psi^k_i(\theta_b, \theta_s) \) has increasing differences in \( (\beta^k_i, \tilde{\theta}_i) \), as formally stated in Theorem 9.

**Theorem 9.** \( \psi^k_i(\theta_b, \theta_s) \) from equation (3.15) has strictly increasing differences in \( (\beta^k_i, \tilde{\theta}_i) \). In particular,

1. \( \psi^k_i(\theta_b, \theta_s) - \psi^{NI}_i(\theta_b, \theta_s) \) is strictly increasing in \( \tilde{\theta}_i \).
2. \( \psi^k_i(\theta_b, \theta_s) - \psi^{SI}_i(\theta_b, \theta_s) \) is strictly increasing in \( \tilde{\theta}_i \).
3. \( \psi^{NI}_i(\theta_b, \theta_s) - \psi^{SI}_i(\theta_b, \theta_s) \) is strictly increasing in \( \tilde{\theta}_i \).

From equation (3.23), the probability of pair \( i \) choosing \( k^*_i = BI \) can be written as

\[
\Pr\{k^*_i = BI\} = \frac{1}{1 + \exp\{\psi^{NI}_i(\theta_b, \theta_s) - \psi^{BI}_i(\theta_b, \theta_s)\} + \exp\{\psi^{SI}_i(\theta_b, \theta_s) - \psi^{BI}_i(\theta_b, \theta_s)\}}.
\]
Chapter 3. Judicial Quality, Relative Efficiency, and Firm Boundary Decisions

Theorem 9 implies that \( \psi_{NI}^i(\theta_b, \theta_s) - \psi_{BI}^i(\theta_b, \theta_s) \) and \( \psi_{SI}^i(\theta_b, \theta_s) - \psi_{BI}^i(\theta_b, \theta_s) \) are strictly decreasing in \( \tilde{\theta}_i \). Therefore, \( \Pr\{k_i^* = BI\} \) is strictly increasing in \( \tilde{\theta} \).

Similarly, the probability of pair \( i \) choosing \( k_i^* = SI \) can be written as

\[
\Pr\{k_i^* = SI\} = \frac{1}{1 + \exp(\psi_{BI}^i(\theta_b, \theta_s) - \psi_{SI}^i(\theta_b, \theta_s)) + \exp(\psi_{NI}^i(\theta_b, \theta_s) - \psi_{SI}^i(\theta_b, \theta_s))}.
\]

Theorem 9 implies that \( \psi_{BI}^i(\theta_b, \theta_s) - \psi_{SI}^i(\theta_b, \theta_s) \) and \( \psi_{NI}^i(\theta_b, \theta_s) - \psi_{SI}^i(\theta_b, \theta_s) \) are strictly increasing in \( \tilde{\theta}_i \). Therefore, \( \Pr\{k_i^* = SI\} \) is strictly decreasing in \( \tilde{\theta}_i \).

The above results are summarized in Corollary 2.

**Corollary 2.** \( \Pr\{k_i^* = BI\} \) is increasing in \( \tilde{\theta}_i \); \( \Pr\{k_i^* = SI\} \) is decreasing in \( \tilde{\theta}_i \).

Corollary 2 predicts two linear probability models, as will be further discussed in Section 3.3.1, but it does not generate testable predictions on the multinomial logistic regressions for equation (3.23). Testable predictions for equation (3.23) are stated in Corollary 3.

**Corollary 3.** Define the log odds ratio of pair \( i \) choosing BI and SI as

\[
P_{BI}^i = \ln \left( \frac{\Pr\{k_i^* = BI\}}{\Pr\{k_i^* = NI\}} \right) = \psi_{BI}^i(\theta_b, \theta_s) - \psi_{NI}^i(\theta_b, \theta_s)
\]

and

\[
P_{SI}^i = \ln \left( \frac{\Pr\{k_i^* = SI\}}{\Pr\{k_i^* = NI\}} \right) = \psi_{SI}^i(\theta_b, \theta_s) - \psi_{NI}^i(\theta_b, \theta_s),
\]

then \( P_{BI}^i \) is increasing in \( \tilde{\theta}_i \), and \( P_{SI}^i \) is decreasing in \( \tilde{\theta}_i \).

Corollary 3 is a convenient extension of Theorem 9 because it directly builds on the fact that \( \psi_k^i(\theta_b, \theta_s) \) has increasing differences in \( (\beta_k^i, \tilde{\theta}_i) \).

**Step 3**

In this step, each buyer draws a set of sourcing costs \( \{m_1^b, m_2^b, ..., m_N^b\} \) from a common Pareto distribution,

\[
m_b^i \sim \text{Pareto}(m, \gamma), \quad (3.24)
\]

where \( m \) is the minimum possible value of fixed costs, and \( \gamma > 0 \) is a shape parameter.\(^6\)

\(^6\)The cumulative distribution function for \( m_b^i \) is \( G(m_b^i) = 1 - (m_b^i/m)^{-\gamma} \) for \( m_b^i > m \) and \( G(m_b^i) = 1 \) for \( m_b^i < m \).
Since all countries are assumed to be identical, the buyer chooses to source from the country for which its sourcing cost is the lowest. In other words, buyer b’s optimal sourcing location is determined by the following equation

\[ l_b^* = \arg \min_{l \in \{1, 2, \ldots, N\}} m_b^l. \tag{3.25} \]

Define buyer b’s minimum sourcing cost as

\[ m_b^* = \min\{m_1^b, \ldots, m_N^b\}. \]

I show in Appendix 3.D that the minimum sourcing cost \( m_b^* \) also follows a Pareto distribution:

\[ m_b^* \sim \text{Pareto}(\overline{m}, \gamma N). \tag{3.26} \]

If \( \gamma > 1/N \), the expected value of the sourcing cost is \( \gamma N \overline{m}/(\gamma N - 1) \).

**Steps 2 and 1**

Recall that in steps 1 and 2 buyers and sellers pay an entry cost \( f_E \), draw productivity \( \theta_b \) and \( \theta_s \) from distributions \( G^B(\theta) \) and \( G^S(\theta) \), and decide whether to exit or stay. Firms stay as long as their expected payoffs are greater than or equal to their outside options, which are normalized to 0. A buyer b chooses to stay as long as \( \theta_b \) satisfies the following condition:

\[ \int_{\theta_s} \pi_{b}^*(\theta_b, \theta_s)dG^S(\theta_s) \geq \frac{\gamma N \overline{m}}{\gamma N - 1}. \tag{3.27} \]

A seller s chooses to stay as long as \( \theta_s \) satisfies the following condition:

\[ \int_{\theta_b} \pi_{s}^*(\theta_b, \theta_s)dG^B(\theta_s) \geq 0. \tag{3.28} \]

\( \gamma N \overline{m}/(\gamma N - 1) \) is the buyer’s expected sourcing cost. \( \pi_{b}^*(\theta_b, \theta_s) \) and \( \pi_{s}^*(\theta_b, \theta_s) \) refer to the buyer and seller’s optimal surpluses in match \( i \):

\[ \pi_{b}^*(\theta_b, \theta_s) \equiv \pi_{k}^b(\theta_b, \theta_s)|_{k = k^*_i}, \]

\(^7\)An implicit assumption here is that there is no difference between sourcing at home or abroad.
and
\[ \pi^*_s(\theta_b, \theta_s) \equiv \pi^k_s(\theta_b, \theta_s)|_{k=k_i}. \]

Equations (3.27) and (3.28) imply that there are two threshold levels of productivity, \( \theta^B \) and \( \theta^S \), such that buyers with \( \theta_b < \theta^B \) and sellers with \( \theta_s < \theta^S \) exit the market. \( \theta^B \) and \( \theta^S \) are pinned down by the following two equations:

\[
\frac{1}{1 - G^S(\theta^S)} \int_{\theta^S}^{\infty} \pi^*_b(\theta^B, \theta_s) \, dG^S(\theta_s) - \frac{\gamma N f_m}{\gamma N - 1} = 0, \tag{3.29}
\]

and

\[
\frac{1}{1 - G^B(\theta^B)} \int_{\theta^B}^{\infty} \pi^*_s(\theta_b, \theta^S) \, dG^B(\theta_b) = 0. \tag{3.30}
\]

To summarize, I have solved for the entry thresholds \( \theta_b \) and \( \theta_s \). An industry equilibrium can be solved by balancing the mass of entry \( M_E \) with the quitters’ employments \( f_E(G^S(\theta_s) + G^B(\theta_b)) \) (if labor is the only factor and if the industry \( j \)'s employment is assumed to be constant). I do not specifically solve for the industry equilibrium because it is not pertinent to my empirical exercise.

3.3 Empirical Results

There are two testable implications of this model. The first implication is at the firm-level, the second at the country-level.

3.3.1 Organizational Form \((k^*_i)\) and Relative Efficiency \((\tilde{\theta}_i)\)

Empirical Specification

Recall from equation (3.23), the probability of pair \( i \) choosing an organizational form \( k \) can be described by a multinomial logit model. Theorem 2 further states that the probability of pair \( i \) choosing buyer integration is increasing in \( \tilde{\theta}_i \) and the probability of pair \( i \) choosing seller integration is decreasing in \( \tilde{\theta}_i \). As Theorem 2 does not provide direct predictions for multinomial logit regressions, I test the predictions of Corollary 3. More specifically, define \( y_i \) as the observed organizational form chosen by pair \( i \), and \( \tilde{\theta}_i \) as the empirical proxy for \( \tilde{\theta}_i \),

\[ ^8 \text{A multinomial logit model only identifies log odds ratios (as in Corollary 3), not probabilities (as in Theorem 2).} \]
the relative efficiency of pair \( i \). Then Corollary 3 predicts that the coefficient on \( \hat{\theta} \) should be positive for \( y_i = BI \), and negative for \( y_i = SI \).

Alternatively, the multinomial logit model can be broken down into the following linear probability models:

\[
Pr\{y_i = BI\} = \varphi_i^{BI} + \varphi_i^{BI}\hat{\theta}_i + X_i'\varphi_i^{BI} + \epsilon_i^{BI},
\]

\[
Pr\{y_i = SI\} = \varphi_i^{SI} + \varphi_i^{SI}\hat{\theta}_i + X_i'\varphi_i^{SI} + \epsilon_i^{SI}.
\]

\( y_i \) is pair \( i \)'s observed choice of organizational form. The variable of interest is \( \hat{\theta}_i \), which is an empirical proxy for \( \tilde{\theta}_i \). \( X_i \) is a vector of control variables for pair \( i \). \( \epsilon_i^{BI} \) and \( \epsilon_i^{SI} \) are error terms. Theorem 2 predicts that \( \varphi_i^{BI} > 0 \) and \( \varphi_i^{SI} < 0 \).

The empirical proxy for \( \tilde{\theta}_i, \hat{\theta}_i \), is calculated as

\[
\hat{\theta}_i = \frac{\hat{\theta}_b}{\hat{\theta}_b + \hat{\theta}_s},
\]

where \( \hat{\theta}_b \) and \( \hat{\theta}_s \) are empirical proxies for buyer and seller’s levels of productivity. They are respectively calculated as buyer and seller’s average sales over the period 2005-2012 over their average number of employees over the same period.

The data I use for my empirical exercise is similar to that used in Chapter 2. The only difference is that the buyer-seller relationships are assembled during the period 2014-2016 (as compared to 2012-2014 in Chapter 2. I also improved the algorithm for matching firm names, which significantly increases the sample size.

**Multinomial Logit Regression Results**

Table 3.3 presents the regression results from the multinomial logit regressions, as suggested in Section 3.3.1. The results are consistent with the predictions of Corollary 3. In columns (2)-(4), with an increase in \( \hat{\theta}_i \), pair \( i \) is more likely to choose buyer integration, and less likely to choose seller integration. These results hold after controlling for variables such as the log difference in the buyer and seller’s number of employees and age, pairwise sector fixed effects, and pairwise region fixed effects.\(^9\)

Since it is difficult to interpret the coefficients from a multinomial logit regression, I report the marginal effects of \( \hat{\theta}_i \) for the regression in column (4). The various types of marginal effects

\(^9\)There are a total of 10 sectors and 5 regions in the Capital IQ database. Therefore, there are a total of \( (10 \times 10 = 100) \) pairwise sector fixed effects and \( (5 \times 5 = 25) \) pairwise region fixed effects.
Table 3.2: Marginal Effects from Column (4), Table 3.3

<table>
<thead>
<tr>
<th>$\hat{\theta}_i$</th>
<th>BI</th>
<th>NI</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 1</td>
<td>0.13</td>
<td>0.105</td>
<td>-0.235</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>+1 centered</td>
<td>0.13</td>
<td>0.102</td>
<td>-0.232</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>+SD centered</td>
<td>0.034</td>
<td>0.027</td>
<td>-0.061</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Marginal</td>
<td>0.131</td>
<td>0.104</td>
<td>-0.235</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Base</td>
<td>0.296</td>
<td>0.409</td>
<td>0.295</td>
</tr>
</tbody>
</table>

Notes: Numbers in brackets are standard errors. From the top row to the bottom row, the ranges are: when $\hat{\theta}_i$ increases from 0 to 1, when $\hat{\theta}_i$ increases from mean-0.5 to mean+0.5, when $\hat{\theta}_i$ increases from mean minus $1/2$ standard error to mean plus $1/2$ standard error, and when $\hat{\theta}_i$ takes its mean value. The bottom row 'base' illustrates pair $i$’s probability of choosing $BI$, $NI$, and $SI$ when all variables in Table 3.3, column (4) take their mean values. This row provides a reference for the magnitude of the marginal effects of $\hat{\theta}_i$.

are summarized in Table 3.2. The row “Marginal” illustrates the marginal effects of $\hat{\theta}_i$ on pair $i$ choosing buyer integration, non-integration, and seller integration at the mean. At the mean of $\hat{\theta}_i$, pair $i$’s probability of choosing buyer integration, non-integration, and seller integration are $13.1\%$, $10.4\%$ and $-23.5\%$, respectively. Compared to the base probabilities at the bottom row, these marginal effects are high in magnitudes. The other rows in Table 3.2 show the change in pair $i$’s probability of choosing various organizational forms given a change in $\hat{\theta}_i$ in different ranges. In all these ranges, an increase in $\hat{\theta}_i$ results in an increase in pair $i$’s probability of choosing buyer integration, and decreases its probability of choosing seller integration.

Linear Regression Results

Table 3.4 presents the regression results for the linear probability models from equations (3.31) and (3.32). With an increase in $\hat{\theta}_i$, pair $i$ is more likely to choose buyer integration, and less likely to choose seller integration. The coefficient on $\hat{\theta}_i$ is positive and significant for $BI_i$ (except for column (1)), and negative and significant for $SI_i$. The coefficients remain constant after adding in control variables such as the log difference in buyer and seller’s size and age, and...
Table 3.3: Multinomial Logit Regression Results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}_i$</td>
<td>-0.371***</td>
<td>0.178***</td>
<td>0.527***</td>
<td>0.548***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.019)</td>
<td>(0.021)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>$\ln(\text{emp}_b)-\ln(\text{emp}_s)$</td>
<td>0.284***</td>
<td>0.311***</td>
<td>0.322***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>$\ln(\text{age}_b)-\ln(\text{age}_s)$</td>
<td>0.125***</td>
<td>0.195***</td>
<td>0.189***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>$\hat{\theta}_i$</td>
<td>-1.051***</td>
<td>-1.933***</td>
<td>-1.899***</td>
<td>-1.997***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.021)</td>
<td>(0.023)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>$\ln(\text{emp}_b)-\ln(\text{emp}_s)$</td>
<td>-0.383***</td>
<td>-0.416***</td>
<td>-0.428***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>$\ln(\text{age}_b)-\ln(\text{age}_s)$</td>
<td>-0.225***</td>
<td>-0.136***</td>
<td>-0.155***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td></td>
</tr>
</tbody>
</table>

Pairwise sector fixed effects | N | N | Y | Y |
Pairwise region fixed effects | N | N | N | Y |
Observations | 310557 | 310557 | 310557 | 310557 |
Pseudo R-squared | 0.006 | 0.258 | 0.353 | 0.377 |

Notes: $\hat{\theta}_i$ is calculated as the ratio between buyer and seller’s average sales per employee. $\ln(\text{emp}_b)-\ln(\text{emp}_s)$ is the log difference between buyer $b$ and seller $s$’s employments. $\ln(\text{age}_b)-\ln(\text{age}_s)$ is the log difference between buyer and seller’s ages.

fixed effects including pairwise industry fixed effects and pairwise country fixed effects.\(^{10}\)

3.3.2 Organizational Form ($k^*_i$) and Judicial Quality ($\delta_b, \delta_s$)

Empirical Specification

Recall that $\delta_b$ is the fraction of output that the buyer can retrieve from the seller under buyer integration, and $\delta_s$ is the fraction of output that the seller can retrieve from the buyer under seller integration. Theoretically, $\delta_b$ and $\delta_s$ determine buyer and seller’s outside options in the negotiation stage, and in turn affects their revenue shares under various ownership structures.

\(^{10}\)There are 67 industries, 181 buyer countries and 172 seller countries in the regression sample. There are a total of 3,736 pairwise industry fixed effects and a total of 4,772 pairwise country fixed effects.
### Chapter 3. Judicial Quality, Relative Efficiency, and Firm Boundary Decisions

#### Table 3.4: Linear Regression Results

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: BI&lt;sub&gt;i&lt;/sub&gt;</th>
<th></th>
<th>Dependent Variable: SI&lt;sub&gt;i&lt;/sub&gt;</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>0.015</td>
<td>0.089***</td>
<td>0.184***</td>
<td>0.168***</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.017)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>ln(emp&lt;sub&gt;b&lt;/sub&gt;) - ln(emp&lt;sub&gt;s&lt;/sub&gt;)</td>
<td>0.055***</td>
<td>0.059***</td>
<td>0.058***</td>
<td>-0.061***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>ln(age&lt;sub&gt;b&lt;/sub&gt;) - ln(age&lt;sub&gt;s&lt;/sub&gt;)</td>
<td>0.026***</td>
<td>0.034***</td>
<td>0.028***</td>
<td>-0.035***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td>0.016</td>
<td>-0.267***</td>
<td>-0.238***</td>
<td>-0.248***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>ln(age&lt;sub&gt;b&lt;/sub&gt;) - ln(age&lt;sub&gt;s&lt;/sub&gt;)</td>
<td>-0.061***</td>
<td>-0.061***</td>
<td>-0.060***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>ln(age&lt;sub&gt;b&lt;/sub&gt;) - ln(age&lt;sub&gt;s&lt;/sub&gt;)</td>
<td>-0.035***</td>
<td>-0.029***</td>
<td>-0.029***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Pairwise industry fixed effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Pairwise country fixed effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Observations</td>
<td>310557</td>
<td>310557</td>
<td>310557</td>
<td>310557</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.000</td>
<td>0.266</td>
<td>0.317</td>
<td>0.461</td>
</tr>
</tbody>
</table>

**Notes:** $\tilde{\theta}_i$ is calculated as the ratio between buyer's and seller's average sales per employee. ln(emp<sub>b</sub>) - ln(emp<sub>s</sub>) is the log difference between buyer b's and seller s's employments. ln(age<sub>b</sub>) - ln(age<sub>s</sub>) is the log difference between buyer and seller's ages.

as shown in Table 3.1.

Figure 3.4 illustrates how $\delta_b$ and $\delta_s$ affect a firm’s organizational choice. Recall from Theorem 8 that $\tilde{\theta}^L_i$ and $\tilde{\theta}^H_i$ are the threshold levels of relative efficiency, such that the firm-pair chooses seller integration if $\tilde{\theta}_i < \tilde{\theta}^L_i$, buyer integration if $\tilde{\theta}_i > \tilde{\theta}^H_i$, and non-integration if $\tilde{\theta}_i$ is in-between. Given the optimal revenue share function $\beta^*_i$, $\tilde{\theta}^L_i$ is determined by the distance between $\beta^{SI}_i$ and $\beta^{NI}_i$ while $\tilde{\theta}^H_i$ is determined by the distance between $\beta^{NI}_i$ and $\beta^{BI}_i$. $\delta_s$ and $\delta_b$ affect the firm-pair’s organizational decision only by affecting the values of $\beta^{SI}_i$ and $\beta^{BI}_i$. More specifically, for a given value of $\delta_s$, as $\delta_b$ increases, $\beta^{BI}_i$ rises, increasing the value of $\tilde{\theta}^H_i$, thus decreasing the probability of the firm pair choosing buyer integration and increasing its probability of choosing non-integration. Similarly, fixing the value of $\delta_b$ and increasing $\delta_s$ leads to lower values of $\beta^{SI}_i$ and $\tilde{\theta}^L_i$, and a lower probability of pair i choosing seller integration. To summarize, the theory predicts that an increase in $\delta_b/\delta_s$ decreases the firm-pair’s probability of choosing buyer integration and increases the firm-pair’s probability of choosing seller integration.

Empirically, the fraction of output that can be retrieved is measured by judicial quality (Kaufmann et al., 2004). More specifically, $\delta_b$ is measured by the judicial quality in the seller’s...
country, while $\delta_s$ is measured by the judicial quality in the buyer’s country.\footnote{International commercial dispute settlement follows a standard procedure: if the contract includes an arbitration clause, either private or institutional, then the dispute is settled by the arbitrator. If not, and if the two countries involved are both participants of the New York Convention, then the arbitration award is made according to the New York Convention provisions. If neither of the previous two conditions apply, then the plaintiff resorts to its local court to make the arbitration. However, if the defendant refuses to comply, then the plaintiff eventually has to resort to the defendant’s local court to settle the dispute. Even when the contract specifies an arbitration clause, the final implementation depends to a large extent on the defendant country’s judicial quality. Relating to the model in this paper, $\delta_b$ is the fraction of output that the buyer is able to retrieve from the seller, where the buyer is the plaintiff and the seller the defendant. Therefore, a correct measure for $\delta_b$ should be the judicial quality in the seller’s country of location. Similarly, the correct measure for $\delta_s$ should be the judicial quality in the buyer’s country of location.} Let $Q_b$ and $Q_s$ denote the judicial qualities in the buyer and the seller’s countries of location, respectively. The empirical measure for $\delta_b/\delta_s$ is $Q_s/Q_b$. Since an increase in $\delta_b/\delta_s$ decreases the firm-pair’s probability of choosing buyer integration and increases its probability of choosing seller integration, an increase in $Q_b/Q_s$ increases the firm-pair’s probability of choosing buyer integration and decreases its probability of choosing seller integration. These predictions are represented by the following linear probability model:
Chapter 3. Judicial Quality, Relative Efficiency, and Firm Boundary Decisions

\[
\Pr \{ y_i = BI \} = \varphi_0^B + \varphi_1^B \cdot \frac{Q_b}{Q_s} + X_i' \varphi^B + \varepsilon_i^B; \quad (3.34)
\]

\[
\Pr \{ y_i = SI \} = \varphi_0^S + \varphi_1^S \cdot \frac{Q_b}{Q_s} + X_i' \varphi^S + \varepsilon_i^S. \quad (3.35)
\]

The theory predicts that \( \varphi_1^B > 0 \) and \( \varphi^S < 0 \).

Linear Regression Results

Table 3.5 presents the regression results for the linear probability models from equations (3.34) and (3.35). \( BI_i \) and \( SI_i \) are dummy variables for buyer integration and seller integration, respectively. The judicial quality measures \( Q_b \) and \( Q_s \) are adopted from Nunn (2007), who collects the data from Kaufmann et al. (2004). The judicial quality measure is a weighted average of a number of variables measuring individuals’ perceptions of the effectiveness and predictability of the judiciary and the enforcement of contracts in each country between 1997 and 1998. The results in Table 3.5 are consistent with the theory’s implications: when \( Q_b/Q_s \) increases, pair \( i \) is more likely to choose buyer integration, and less likely to choose seller integration. This pattern remains after adding in firm- and country-level control variables.

Table 3.5: Linear Regression Results

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: ( BI_i )</th>
<th></th>
<th>Dependent Variable: ( SI_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>( Q_b/Q_s )</td>
<td>0.187*** 0.101*** 0.084***</td>
<td></td>
<td>-0.149*** -0.046*** -0.063***</td>
</tr>
<tr>
<td></td>
<td>(0.018) (0.013) (0.014)</td>
<td></td>
<td>(0.019) (0.011) (0.014)</td>
</tr>
<tr>
<td>( \ln(\text{emp}_b) - \ln(\text{emp}_s) )</td>
<td>0.054*** 0.057***</td>
<td></td>
<td>-0.059*** -0.059***</td>
</tr>
<tr>
<td></td>
<td>(0.002) (0.002)</td>
<td></td>
<td>(0.002) (0.002)</td>
</tr>
<tr>
<td>( \ln(\text{age}_b) - \ln(\text{age}_s) )</td>
<td>0.024*** 0.035***</td>
<td></td>
<td>-0.040*** -0.032***</td>
</tr>
<tr>
<td></td>
<td>(0.003) (0.003)</td>
<td></td>
<td>(0.004) (0.004)</td>
</tr>
<tr>
<td>Pairwise industry fixed effects</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Observations</td>
<td>303750</td>
<td>303750</td>
<td>303750</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.009</td>
<td>0.265</td>
<td>0.308</td>
</tr>
</tbody>
</table>

Notes: The firm-level controls include the log differences in the buyer and seller’s revenue, age, and employments. The country-level controls include the buyer and seller countries’ log GDP per capita, human capital per worker, and capital per worker.
Multinomial Logit Regression Results

Table 3.6: Multinomial Logit Regression Results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_i = BI )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Q_b / Q_s )</td>
<td>0.652***</td>
<td>0.524***</td>
<td>0.404***</td>
<td>0.470***</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.020)</td>
<td>(0.022)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>( \ln(e_{mp_b}) - \ln(e_{mp_s}) )</td>
<td>0.282***</td>
<td>0.302***</td>
<td>0.312***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>( \ln(age_{b}) - \ln(age_{s}) )</td>
<td>0.110***</td>
<td>0.191***</td>
<td>0.186***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>( y_i = SI )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Q_b / Q_s )</td>
<td>-0.540***</td>
<td>-0.324***</td>
<td>-0.425***</td>
<td>-0.576***</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.024)</td>
<td>(0.026)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>( \ln(e_{mp_b}) - \ln(e_{mp_s}) )</td>
<td>-0.352***</td>
<td>-0.385***</td>
<td>-0.395***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>( \ln(age_{b}) - \ln(age_{s}) )</td>
<td>-0.256***</td>
<td>-0.163***</td>
<td>-0.184***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td></td>
</tr>
</tbody>
</table>

Pairwise sector fixed effects  | N | N | Y | Y |
Pairwise region fixed effects  | N | N | N | Y |
Observations                   | 303750 | 303750 | 303750 | 303750 |
Pseudo R-squared               | 0.005 | 0.245 | 0.341 | 0.365 |

Notes: \( Q_b \) and \( Q_s \) are respectively the judicial qualities in the buyer and the seller’s countries. I borrow them from Nunn (2007). \( \ln(e_{mp_b}) - \ln(e_{mp_s}) \) is the log difference between buyer and seller’s employments. \( \ln(age_{b}) - \ln(age_{s}) \) is the log difference between buyer and seller’s ages. There are a total of 10 × 10 = 100 pairwise sector fixed effects and 5 × 5 = 25 pairwise region fixed effects. I could not control for finer levels of fixed effects due to the well-known convergence issue with multinomial logistic regressions.

Table 3.6 presents the regression results from the multinomial logistic equivalence of equations (3.34) and (3.35). The dependent variable is \( y_i \), which is pair \( i \)’s observed organizational choice. With an increase in \( Q_b / Q_s \), pair \( i \) is more likely to choose buyer integration, and less likely to choose seller integration.

Since it is difficult to interpret the coefficients of multinomial logit regressions, I report the marginal effects of \( Q_b / Q_s \) from the regression in column (4) of Table 3.6. These marginal effects appear in Table 3.7. In the row ”Marginal”, see that at the mean value of \( Q_b / Q_s \),
Table 3.7: Marginal Effects from Column (4), Table 3.6

<table>
<thead>
<tr>
<th>$Q_b / Q_s$</th>
<th>BI</th>
<th>NI</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 1</td>
<td>0.075</td>
<td>0.01</td>
<td>-0.085</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>+1 centered</td>
<td>0.078</td>
<td>0.003</td>
<td>-0.081</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>centered</td>
<td>0.019</td>
<td>0.001</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Marginal</td>
<td>0.079</td>
<td>0.003</td>
<td>-0.081</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Base</td>
<td>0.296</td>
<td>0.408</td>
<td>0.295</td>
</tr>
</tbody>
</table>

Notes: Numbers in brackets are standard errors. From the top row to the bottom row, the ranges are: when $Q_b / Q_s$ increases from 0 to 1, when $Q_b / Q_s$ increases from mean-0.5 to mean+0.5, when $Q_b / Q_s$ increases from mean minus 1/2 standard error to mean plus 1/2 standard error, when $Q_b / Q_s$ increases from its minimum value to its maximum value, when $Q_b / Q_s$ takes its mean value. The bottom row ‘base’ illustrates pair i’s probability of choosing BI, NI, and SI when all variables in Table 3.3, column (4) take their mean values. This row provides a reference for the magnitude of the marginal effects of $Q_b / Q_s$.

pair i’s probability of choosing buyer integration, non-integration, and seller integration are respectively 7.9%, 0.3%, and −8.1%. These magnitudes are much smaller than the ones from Table 3.2, which is understandable because $Q_b / Q_s$ are country-level variables, and so should have weaker explanatory power than firm-level variables such as $\hat{\theta}_i$.

### 3.4 Conclusion

I examine the impact of two factors on firms’ organization decisions: relative efficiency and relative judicial quality. Both factors affect a firm’s probability of becoming the “integrator”, but through different channels.

Relative efficiency measures a firm’s relative importance in a production relationship. A firm with higher relative efficiency generates a higher marginal return relative to its partner. When one of the two firms is significantly more important than the other, firms find it optimal to incentivize this firm by allocating more property rights to it, even when it comes at the
expense of the other firm’s incentive.

Relative judicial quality effectively acts as a substitute to integration. Finding a partner from a country with good legal institutions provides better protection for a firm’s property rights. Even without implementing the legal procedure, a better legal institution improves this firm’s negotiation terms and increases its investment incentive, thus reduces the firm’s need to integrate its partner. Therefore, firms from a country with good legal institutions are less likely to be integrated.

In addition to the institutions and comparative advantage literature mentioned in footnote 3 in Section 3.1, this paper also contributes to the vigorous literature on contractual incompleteness and firm boundary decisions following the seminal works by Grossman and Hart (1986) and Hart and Moore (1990). Antràs (2003), Antràs and Helpman (2004), and Antràs and Helpman (2008) are the influential papers that connect property rights theory with international trade. These theoretical papers spurred an extensive empirical literature including Antràs and Chor (2013), Nunn and Trefler (2013), Antràs et al. (2014), Bernard et al. (2014), and Alfaro et al. (2015). See Antràs (2015) and Antràs (2013) for a more comprehensive review of this literature.
Appendix

3.A Proof of Lemma 5

Recall that $\psi_i(\beta, \theta_b, \theta_s)$ in equation (3.20) can be rearranged as

$$\psi_i(\beta, \beta_b, \beta_s) = \alpha^{\alpha/(1-\alpha)} \left( \frac{\theta_s}{c_s} \right)^{\alpha/(1-\alpha)} g_i(\beta, \tilde{\theta}_i),$$

(3.36)

where

$$g_i(\beta, \tilde{\theta}_i) \equiv \frac{(1 - \alpha \beta) (\beta)^{\rho/(1-\rho)} \tilde{\theta}_i + [1 - \alpha (1 - \beta)](1 - \beta)^{\rho/(1-\rho)}}{\{\beta^{\rho/(1-\rho)} \tilde{\theta}_i + (1 - \beta)^{\rho/(1-\rho)}\}^{(\rho - \alpha)/(\rho(1-\alpha))}},$$

(3.37)

and

$$\tilde{\theta}_i \equiv \left( \frac{\theta_b/c_b}{\theta_s/c_s} \right)^{\rho/(1-\rho)}.$$

Since $\beta$ and $\tilde{\theta}$ enter $\psi_i(\beta, \theta_b, \theta_s)$ only through $g_i(\beta, \tilde{\theta}_i)$, $\psi_i(\beta, \theta_b, \theta_s)$ is supermodular in $(\beta, \tilde{\theta}_i)$ if and only if $g_i(\beta, \tilde{\theta})$ is supermodular in $(\beta, \tilde{\theta}_i)$. By definition, this requires

$$\frac{\partial^2 g_i(\beta, \tilde{\theta}_i)}{\partial \beta \partial \tilde{\theta}_i} \geq 0.$$

Since

$$\frac{\partial^2 \ln g_i(\beta, \tilde{\theta}_i)}{\partial \beta \partial \tilde{\theta}_i} = \frac{\partial}{\partial \beta} \frac{\partial \ln g_i(\beta, \tilde{\theta}_i)}{\partial \tilde{\theta}_i} = \frac{\partial}{\partial \beta} \frac{1}{g_i(\beta, \tilde{\theta}_i)} \frac{\partial g_i(\beta, \tilde{\theta}_i)}{\partial \tilde{\theta}_i} = \frac{1}{g_i(\beta, \tilde{\theta}_i)} \frac{\partial^2 g_i(\beta, \tilde{\theta}_i)}{\partial \beta \partial \tilde{\theta}_i},$$

and $g_i(\beta, \tilde{\theta}_i)$ is always positive,

$$\frac{\partial^2 g_i(\beta, \tilde{\theta}_i)}{\partial \beta \partial \tilde{\theta}_i} \geq 0 \iff \frac{\partial^2 \ln g_i(\beta, \tilde{\theta}_i)}{\partial \beta \partial \tilde{\theta}_i} \geq 0.$$

$$\ln g_i(\beta, \tilde{\theta}_i) = \ln \left[ (1 - \alpha \beta)\beta^{\rho/(1-\rho)} \tilde{\theta}_i + [1 - \alpha (1 - \beta)](1 - \beta)^{\rho/(1-\rho)} \right] - \frac{\rho - \alpha}{\rho(1-\alpha)} \ln \left[ \beta^{\rho/(1-\rho)} \tilde{\theta}_i + (1 - \beta)^{\rho/(1-\rho)} \right].$$
With simplification,
\[
\frac{\partial \ln g_i(\beta, \tilde{\theta}_i)}{\partial \tilde{\theta}_i} = \frac{\alpha}{1 - \alpha} \left( \frac{1 - \rho}{\rho} \theta + \left( \frac{1 - \beta}{\beta} \right)^{\rho/(1 - \rho)} \right).
\]
Since \(0 < \alpha, \rho < 1\), it is obvious that the above expression is strictly increasing in \(\beta\). Therefore,
\[
\frac{\partial^2 \ln g_i(\beta, \tilde{\theta}_i)}{\partial \beta \partial \tilde{\theta}_i} > 0.
\]
In turn,
\[
\frac{\partial^2 g_i(\beta, \tilde{\theta}_i)}{\partial \beta \partial \tilde{\theta}_i} > 0.
\]
Since \(g_i(\beta, \tilde{\theta}_i)\) is supermodular in \((\beta, \tilde{\theta}_i)\), \(\psi_i(\beta, \theta_b, \theta_s)\) is also supermodular in \((\beta, \tilde{\theta}_i)\). This further implies that the discrete version of \(\psi_i(\beta, \theta_b, \theta_s)\), \(\psi^k_i(\theta_b, \theta_s)\) has increasing differences in \((\beta^k_i, \tilde{\theta}_i)\). Hence Theorem 9 holds as well.

### 3.B Proof of Theorem 7

Same as Appendix 3.A. \(\psi_i(\beta, \theta_b, \theta_s)\) from equation (3.20) can be rearranged as
\[
\psi_i(\beta, \theta_b, \theta_s) = \alpha^{\alpha/(1 - \alpha)} \left( \frac{\theta_s}{c_s} \right)^{\alpha/(1 - \alpha)} g_i(\beta),
\]
where
\[
g_i(\beta, \tilde{\theta}_i) = \left( 1 - \alpha \beta \right) (\beta)^{\rho/(1 - \rho)} \tilde{\theta}_i + \left( 1 - \alpha (1 - \beta) \right)(1 - \beta)^{\rho/(1 - \rho)} \left\{ \beta^{\rho/(1 - \rho)} \tilde{\theta}_i + (1 - \beta)^{\rho/(1 - \rho)} \right\}^{(\rho - \alpha)/(\rho(1 - \alpha))}.
\]
Chapter 3. Judicial Quality, Relative Efficiency, and Firm Boundary Decisions

\[ \tilde{\theta}_i \equiv \left( \frac{\theta_b}{\theta_s} \right)^{\rho/(1-\rho)}. \]

Since \( \beta \) enters \( \psi_i(\beta, \theta_b, \theta_s) \) only through \( g_i(\beta, \tilde{\theta}_i) \), the optimal \( \beta \) that maximizes \( g_i(\beta, \tilde{\theta}) \) also maximizes \( \psi_i(\beta, \theta_b, \theta_s) \). Denote the optimal \( \beta \) by \( \beta_i^* \), then \( \beta_i^* \) is obtained by solving \( \frac{\partial g_i(\beta, \tilde{\theta}_i)}{\partial \beta} = 0 \). Since \( g_i(\beta, \tilde{\theta}_i) > 0 \) for \( \beta \in (0, 1) \),

\[ \frac{\partial \ln g_i(\beta, \tilde{\theta}_i)}{\partial \beta} = \frac{1}{1-\rho} \left[ \frac{\beta^\rho/(1-\rho) - 1}{1 - \rho} \tilde{\theta}_i - (1 - \beta)^\rho/(1-\rho) - 1 \right] - \alpha \left[ \beta^\rho/(1-\rho) - 1 \right] \tilde{\theta}_i - (1 - \beta)^\rho/(1-\rho) - 1 \] \[ \frac{\partial \ln g_i(\beta, \tilde{\theta}_i)}{\partial \beta} = \frac{\rho - \alpha}{\rho(1 - \alpha)} \ln \left[ \beta^\rho/(1-\rho) \tilde{\theta}_i + (1 - \beta)^\rho/(1-\rho) \right]. \]

It is obvious from the above equation that when \( \tilde{\theta}_i \to \infty \), \( \beta \to 1 \); when \( \tilde{\theta}_i \to 0 \), \( \beta \to 0 \). This equation is a quadratic function of \( \tilde{\theta}_i \), so \( \tilde{\theta}_i \) can be solved as
\[
\tilde{\theta}_i = \frac{(1 - \alpha) \left[ \left( \frac{\beta^*_i}{1 - \beta^*_i} \right)^2 - 1 \right] + \sqrt{(1 - \alpha)^2 \left[ \left( \frac{\beta^*_i}{1 - \beta^*_i} \right)^2 - 1 \right]^2 - 4(1 - \rho)^2 \left( \frac{\beta^*_i}{1 - \beta^*_i} \right)^2}}{2(1 - \rho) \left( \frac{\beta^*_i}{1 - \beta^*_i} \right)^{\rho/(1 - \rho)}}.
\]

### 3.C Proof of Theorem 8

Lemma 7 shows the existence and uniqueness of \( \tilde{\theta}_H \) and \( \tilde{\theta}_L \). They are implicitly solved by \( \psi_{SI} = \psi_{NI} \) and \( \psi_{BI} = \psi_{NI} \). More specifically, \( \tilde{\theta}_H \) is solved by the following equation:

\[
\frac{(1 - \alpha \beta^*_i) \left( \beta^*_{BI} \right)^{\rho/(1 - \rho)} \tilde{\theta}_H + (1 - \alpha + \alpha \beta^*_i) \left( 1 - \beta^*_{BI} \right)^{\rho/(1 - \rho)}}{\left\{ (\beta^*_i)^{\rho/(1 - \rho)} \tilde{\theta}_H + (1 - \beta^*_{SI})^{\rho/(1 - \rho)} \right\}^{(\rho - \alpha)/\rho(1 - \alpha)}} = \frac{(1 - \alpha \beta^*_{NI}) \left( \beta^*_{NI} \right)^{\rho/(1 - \rho)} \tilde{\theta}_H + (1 - \alpha + \alpha \beta^*_{NI}) \left( 1 - \beta^*_{NI} \right)^{\rho/(1 - \rho)}}{\left\{ (\beta^*_i)^{\rho/(1 - \rho)} \tilde{\theta}_H + (1 - \beta^*_{NI})^{\rho/(1 - \rho)} \right\}^{(\rho - \alpha)/\rho(1 - \alpha)}},
\]

and \( \tilde{\theta}_L \) is solved by the following equation:

\[
\frac{(1 - \alpha \beta^*_i) \left( \beta^*_{SI} \right)^{\rho/(1 - \rho)} \tilde{\theta}_L + (1 - \alpha + \alpha \beta^*_i) \left( 1 - \beta^*_{SI} \right)^{\rho/(1 - \rho)}}{\left\{ (\beta^*_i)^{\rho/(1 - \rho)} \tilde{\theta}_L + (1 - \beta^*_{SI})^{\rho/(1 - \rho)} \right\}^{(\rho - \alpha)/\rho(1 - \alpha)}} = \frac{(1 - \alpha \beta^*_{NI}) \left( \beta^*_{NI} \right)^{\rho/(1 - \rho)} \tilde{\theta}_L + (1 - \alpha + \alpha \beta^*_{NI}) \left( 1 - \beta^*_{NI} \right)^{\rho/(1 - \rho)}}{\left\{ (\beta^*_i)^{\rho/(1 - \rho)} \tilde{\theta}_L + (1 - \beta^*_{NI})^{\rho/(1 - \rho)} \right\}^{(\rho - \alpha)/\rho(1 - \alpha)}}.
\]
3.D The Distribution of $m_b^*$

$m_b^l$ follows a Pareto distribution with minimum value $m$ and shape parameter $\gamma$. The density of such a Pareto distribution is $\gamma m^\gamma/(m_b^l)^{\gamma+1}$. The survival function for $m_b^l$ is

$$
\Pr\{M > m_b^l\} = \begin{cases} 
(m/m_b^l)^\gamma, & m_b^l \geq m; \\
0, & m_b^l < m.
\end{cases} \tag{3.38}
$$

The probability of the random vector $(M_b^1, ... M_b^N)$ having a minimum $f$ is

$$
\Pr\{\min(M_b^1, ... M_b^N) = f\} = \sum_{k=1}^{N} \frac{\gamma m^\gamma}{f^{\gamma+1}} \prod_{l \neq k} \left[ \frac{m}{f} \right] 
= \sum_{k} \frac{\gamma m^\gamma}{f^{\gamma+1}} \frac{m^\gamma}{f^{\gamma(N-1)}} 
= \gamma N \frac{m^\gamma N}{f^{\gamma N+1}}.
$$

The above function is the probability distribution function of a Pareto distribution with minimum value $m$ and shape parameter $\gamma N$, so the minimum $f$ follows a Pareto distribution with minimum $m$ and shape parameter $\gamma N$. 
Bibliography


BIBLIOGRAPHY


Helpman, Elhanan, _Understanding Global Trade_ 2011.


