

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


Essays on Economics of Innovation and International Trade

Citation

Fadeev, Evgenii. 2022. Essays on Economics of Innovation and International Trade. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link

https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37372111

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

HARVARD UNIVERSITY
Graduate School of Arts and Sciences



DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Economics

have examined a dissertation entitled

"Essays on Economics of Innovation and International Trade"

presented by Evgenii Fadeev

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature _____

Typed name: Prof. Pol Antras

Signature _____

Typed name: Prof. Andrei Shleifer

Signature _____

Typed name: Prof. Oliver Hart

Signature _____

Typed name: Prof. Joshua Lerner

Date: April 21, 2022

Essays on Economics of Innovation and International Trade

A dissertation presented

by

Evgenii Fadeev

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

Harvard University

Cambridge, Massachusetts

April 2022

© 2022 Evgenii Fadeev

All rights reserved.

Dissertation Advisors:
Professor Pol Antràs
Professor Andrei Shleifer

Author:
Evgenii Fadeev

Essays on Economics of Innovation and International Trade

Abstract

This dissertation studies how firms share knowledge with business partners, protect their intellectual property rights, and organize production around the world.

In the first chapter, I analyze patent citations which are often used to measure knowledge spillovers between firms. I show that even the most cited patents granted before 2000 received around 50% of citations from one firm only, and this concentration has increased up to 75% by 2014. I argue that firms frequently combine patents and trade secrets to protect their intellectual property, and that patent citations reflect intentional sharing of secrets between business partners. Then, I develop and test a theory of knowledge sharing between firms in a production network that accounts for the observed citation patterns.

The second chapter, co-authored with Alice Wu, documents new facts on trade secret litigation. We use the textual analysis of litigation records to study the types of information firms protect as trade secrets, such as business or technological information, and the mechanisms for the misappropriation of these secrets, such as the mobility of employees across firms. We document the variation in litigation activity and the composition of trade secrets over time, across US states, across industries, and between firms based on their size. We develop a theory of knowledge sharing between firms which shows how to evaluate the welfare consequences of changes in trade secret laws, for example, the Defend Trade Secrets Act of 2016.

The third chapter, co-authored with Professors Pol Antràs, Teresa Fort, and Felix Tintelnot, studies the organization of production by multinational companies. We link US Census data on firms' domestic establishments to BEA surveys on foreign direct investment to document a series of novel facts regarding the global assembly and global sourcing strategies of U.S. firms, and unveil systematic interdependencies in these firm decisions. We then develop a multi-country model in which firms decide on the location of their assembly plants (their assembly strategy) as well as the source of the inputs used in their plants worldwide (their sourcing strategy), and use the model to interpret our novel facts.

Contents

Title i
Copyright ii
Abstract iii
Table of Contents v
Acknowledgments xii

1 Creative Construction: Knowledge Sharing in Production Networks 1
1.1 Introduction 1
1.2 Motivating Evidence: Concentration of Citations 9
1.2.1 Results: High Concentration of Patent Citations 10
1.2.2 Sources of the Concentration of Citations 12
1.2.3 Movement of Inventors 17
1.2.4 Robustness 20
1.2.5 Discussion and Implications for the Theory 20
1.3 Theory 21
1.3.1 Simple Model 23
1.3.2 Extensions and Additional Results 32
1.3.3 Testable Predictions 37
1.4 Empirical Tests 38
1.4.1 Citations in a Production Network 40
1.4.2 Trade with China and Concentration of Knowledge Flows 44
1.5 Policy Implications 49
1.5.1 Estimation of the Value of Private Knowledge 49
1.5.2 Welfare Implications of Restrictions on Vertical Contracts 53
1.6 Conclusion 58

2 The Role of Trade Secrets in the US Economy: Evidence from Litigation 60
2.1 Introduction 60
2.2 Facts on Trade Secret Litigation 63

2.2.1 Background . . . 63
2.2.2 Data and Methodology . . . 64
2.2.3 Empirical Results . . . 65
2.3 Theory: Regulation of Trade Secrets . . . 73
2.4 Conclusion . . . 79

3 Global Sourcing and Multinational Activity: A Unified Approach 81
3.1 Introduction . . . 81
3.2 New Data and Facts on US Multinational Activity . . . 92
3.2.1 Data Description . . . 92
3.2.2 Aggregate Importance of US-based MNEs . . . 95
3.2.3 Definition of Our Sample . . . 96
3.2.4 Aggregate Import and Export Statistics . . . 99
3.2.5 Foreign Manufacturing Affiliate Statistics . . . 101
3.3 Global Sourcing and MNE Status: Reduced-Form Evidence . . . 102
3.3.1 Import and Export Behavior by Firm Type . . . 103
3.3.2 Interactions Between FDI and Importing . . . 104
3.4 Theoretical Framework . . . 110
3.4.1 Environment . . . 111
3.4.2 Manufacturing Production . . . 112
3.4.3 Firm Behavior for Fixed Assembly and Sourcing Strategies . . . 118
3.4.4 Intensive Margin Analysis: Complementarities and Cannibalization Effects . . . 122
3.4.5 Industry and General Equilibrium . . . 126
3.5 Optimal Sourcing and Assembly Strategies . . . 127
3.5.1 The Problem and General Results . . . 127
3.5.2 The Case with Pervasive Complementarities . . . 132
3.6 Zooming in on the Assembly-Sourcing Complementarity . . . 134
3.6.1 Isolating the Assembly-Sourcing Complementarity . . . 134
3.6.2 The Complementarity in Action: Two Illustrative Firm-Level Examples 138
3.7 Conclusion . . . 145

References 147

Appendix A Appendix to Chapter 1 156
A.1 Case Studies: Combining Secrecy and Patenting . . . 156
A.2 Data Appendix . . . 158

A.2.1 Data Details . . . 158
A.2.2 Details to Section 2: Concentration of Patent Citations . . . 159
A.2.3 Movement of Inventors and Citation Patterns . . . 165
A.2.4 Robustness to Section 2: Concentration of Citations . . . 166
A.2.5 Details to Section 4.1: Citations in a Production Network . . . 168
A.2.6 Details to Section 4.2: Import Competition with China . . . 170
A.2.7 Details to Section 5.1: Estimation of the Value of Private Knowledge 171
A.2.8 Joint Innovations Between Suppliers and Customers . . . 172
A.3 Figures and Tables . . . 173
A.4 Theory Appendix . . . 185
A.4.1 Proposition 1: $\mathcal{I}$'s profits under different contracts . . . 185
A.4.2 Two Suppliers and a Common Customer . . . 187
A.4.3 Welfare Implications of Restrictions on the Vertical Contracts . . . 188

Appendix B Appendix to Chapter 2 191

Appendix C Appendix to Chapter 3 195
C.1 Theory Appendix . . . 195
C.2 Data Appendix . . . 200
C.2.1 Matching the Census and BEA data . . . 200
C.2.2 Distinguishing US and Foreign-Owned Firms . . . 201
C.2.3 BEA Country Classifications . . . 204
C.2.4 Sample Description . . . 204
C.2.5 Additional Tables . . . 205
C.2.6 Gravity Regressions for Firm Exports . . . 205
C.2.7 Aggregate gravity regressions . . . 205

List of Tables

1.1 Example of a patent with a high concentration of citations 2
1.2 Quality of Final Products 25
1.3 Supply Chain Relationship and the Probability of a Citation 43
1.4 Trade with China and Increase in the Concentration of Citations 48
1.5 Value of Patents as a Function of Backward Citations 53

3.1 Aggregate statistics for US-based firms in 2007, by firm import and MNE status 96
3.2 Aggregate statistics for firms in 2007, by firm manufacturing, import, and MNE type 98
3.3 US-Based MNEs' foreign affiliate activity, by firm manufacturing and MNE type 99
3.4 Trade statistics for manufacturing firms in 2007, by import and MNE status 101
3.5 Foreign affiliate manufacturing sales for US MNEs in 2007, by destination . 102
3.6 Import and export patterns by MNE status for US manufacturing firms . . 105
3.7 Extensive margin import regressions 108
3.8 Intensive margin import regressions 110

A.1 Distribution of Citations across firms for IBM/Novellus's patent 6126798. . 183
A.2 Distribution of Citations across firms for Motorola's patent 5391517. 184
A.3 Placebo Tests: Trade with China and Increase in the Concentration of Citations 185

B.1 Keywords for the Textual Analysis in Chapter 2 192
B.2 Trade Secret Litigation Across US States 193
B.3 Trade Secret Litigation Across Industries 194

C.1 Trade statistics for non-manufacturing firms in 2007, by import and MNE status 206
C.2 Extensive margin export regressions 207

C.3 Intensive margin export regressions 208
C.4 Aggregate gravity regressions 209

List of Figures

1.1 Concentration of Citations . . . 3
1.2 Decomposition of the Concentration of Citations . . . 15
1.3 Decomposition Within Inventors Who Moved Across Firms . . . 18
1.4 Timing . . . 27
1.5 Equilibrium as a function of (α, γ) . . . 30
1.6 Equilibrium in the Model with Two Suppliers and a Common Customer . . . 36
1.7 The concentration for technological classes with different exposure to import competition. . . . 45
1.8 Changes in the concentration and in the import penetration from China. . . . 47
1.9 Value of Private Knowledge . . . 51
1.10 Equilibrium Under Restrictions on Vertical Contracts . . . 55
1.11 Optimal policy as a function of (α, γ) . . . 56

2.1 Number of Trade Secret Litigation Cases by Years . . . 66
2.2 Trade Secret Litigation over Time . . . 67
2.3 Share of Litigation Cases Across US States . . . 69
2.4 Distribution Across States of Cases Involving Employees . . . 70
2.5 Distribution Across States of Cases based on the Types of Protected Information . . . 71
2.6 Distribution Across States of Cases based on the Types of Protected Information . . . 72
2.7 Share of Litigation Cases by Firm Size . . . 73
2.8 The Composition of Trade Secrets By Firm Size . . . 74
2.9 Equilibrium as a function of (α, q) . . . 78

3.1 Trade Structure . . . 140
3.2 Equilibrium Assembly and Sourcing Decisions . . . 141
3.3 The Effect of Tariffs . . . 142
3.4 Trade Structure . . . 143
3.5 The Effect of Tariffs . . . 145

A.1 Concentration of Citations Within Technological Classes 173
A.2 Concentration of Citations: Patents, Monte-Carlo, and Publications 174
A.3 Decomposition of the Concentration of Citations Based on the Textual Similarity Between Patents 175
A.4 Joint Innovations Between Suppliers and Customers 176
A.5 Trade with China. Actual Concentration vs. Counterfactual One with Equalized Citation Rates 177
A.6 Corporate patents, top 1%, top 10%, exclude superstar firms in patenting . 178
A.7 Alternative aggregation: citation-weighted average and median within classes 179
A.8 Exclude Self-Citations and top patents in terms of backward citations . . . 180
A.9 Exclude Citations from common lawyers and group patents from one family 181
A.10 Concentration of Citations: Examiners vs. Non-examiners 181
A.11 Concentration of Citations for Lawyers Who Represented Several Firms . . 182

Acknowledgments

I am immensely grateful to my advisors Pol Antràs, Oliver Hart, Josh Lerner, and Andrei Shleifer for their continuous guidance and support. Their complementary views and advice shaped my thinking on the economic research.

I have been fortunate to benefit from insightful and detailed discussions with Gabriel Chodorow-Reich, Gordon Hanson, Elhanan Helpman, Bill Kerr, and Marc Melitz. I thank Maxim Boycko, Emmanuel Farhi, Edward Glaeser, Sergei Izmalkov, Myrto Kaloutpsidi, Robin Lee, Marc Melitz, Ariel Pakes, Ludwig Straub, and Hosny Zoabi for their helpful comments. I also would like to thank my co-authors, Pol Antràs, Teresa Fort, Felix Tintelnot, and Alice Wu.

Finally, I am indebted to my wife, Kristina Komissarova, whose encouragement, love, and care helped me to improve my research and to overcome the most difficult challenges during this project.

Chapter 1

Creative Construction: Knowledge Sharing in Production Networks

1.1 Introduction

Knowledge flows between firms are key for economic growth (Romer (1990), Grossman and Helpman (1991), Aghion and Howitt (1992)). The economic and legal rationales for patents rely centrally on these transfers: a patent owner gets monopoly rights on a technology in exchange for its disclosure to the public, so others can create follow-on innovations based on it.[1] Therefore, patents should act as a source of knowledge similar to scientific publications. Patent citations are often used to measure knowledge spillovers in order to discipline growth models (Caballero and Jaffe (1993), Eeckhout and Jovanovic (2002), Akcigit and Kerr (2018)), to evaluate the localization of spillovers in space (Jaffe *et al.* (1993), Thompson and Fox-Kean (2005), Singh and Marx (2013)), and to identify high-quality technologies (Aghion *et al.* (2021), Akcigit *et al.* (2021), Moretti (2021)).

[1]See Mazzoleni and Nelson (1998) and *Pfaff v. Wells Elecs., Inc.*, 525 U.S. 55, 63 (1998) where the U.S. Supreme Court stated that "the patent system represents a carefully crafted bargain that encourages both the creation and the public disclosure of new and useful advances in technology, in return for an exclusive monopoly for a limited period of time."

This paper shows that the distribution of citations across firms is highly concentrated, raising a question about the role of patents in the diffusion of knowledge. For example, the IBM's patent in Table 1.1 is heavily cited, but almost all of its citations come from *one firm only*, Amkor Technology Inc. In general, for each year between 1976 and 2014 I consider the set of the most cited patents filed in the U.S. Patent and Trademark Office.[2] Figure 1.1 shows that a patent granted between 1980 and 2000, on average, received around 50% of citations from one firm only, and this increased to around 70% in 2010s. As a benchmark, the concentration of citations for the most cited scientific publications declined from 8% to 4% over the same time period.

Table 1.1: Example of a patent with a high concentration of citations

Patent Number	Assignee	Total Number of Citations	% of Citations from Amkor Technology Inc
5877043	IBM	218 top 0.005%	94%

To clarify the source of the concentration, I document additional facts about patent citations. First, I show that the concentration is primarily driven by differences across firms in the probabilities of a citation rather than in the number of patents. For instance, Amkor is responsible for the majority of citations to the IBM's patent, not because Amkor has more patents than other companies, but because a large share of its patents makes citations to IBM while other companies with similar patents make no citations. Second, these citation probabilities are specific to citing firms rather than to inventors working in these companies. Specifically, I show that that an inventor who heavily cites a particular patent in one company significantly decreases her citations to this patent once she moves

[2]Specifically, in each grant year and technology class for the period 1976–2014 I track citations within a five-year window for the top 5% of the most cited patents. Taking a five-year window controls for the truncation bias that older patents have more time to accumulate citations. Comparing patents from the same technology class controls for differences across classes in citation patterns (Lerner and Seru (2021)).

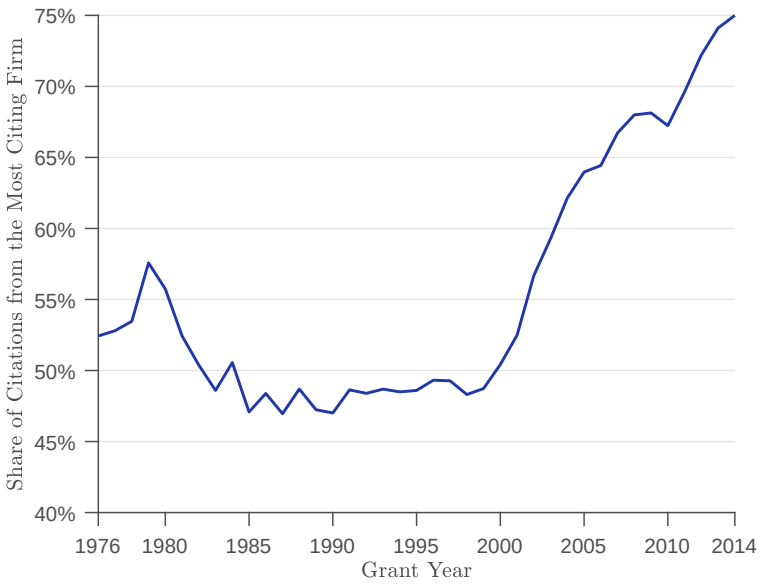


Figure 1.1: Concentration of Citations

This figure shows the average concentration of citations for the most cited patents between 1976 and 2014. In each grant year and technology class for the period 1976–2014, I track citations within a five-year window for the top 5% of the most cited patents. For each cited patent, the concentration is defined as the share of citations coming from the most citing firm. The technological classes are defined at the group level in the cooperative patent classification system. To construct the aggregate measure, I take the average concentration across patents within each class, and then the average across classes weighted by the number of patents.

to another firm. Therefore, the explanation for the concentration of citations should be based on the relationship between firms citing each other rather than on the interactions between their inventors. Finally, I control for the legal specifics of the patent system. For example, the concentration is not driven by patent examiners or lawyers.

The high concentration of citations is puzzling because valuable technologies disclosed in public patent files would be expected to generate spillovers across a broader set of firms. So, interpreting patent citations as being similar to scientific ones, in terms of measuring knowledge spillovers and the breadth of influence, might be incorrect. Instead, an alternative view of patent citations might be that they reflect cooperation or exchange of information between firms with close business ties. This is the broad idea that I pursue

in this paper.

I argue that citations between firms are correlated with the sharing of trade secrets accompanying patents. As a result, they are concentrated because only a limited set of firms gets access to private knowledge of a patent owner. Despite the disclosure requirements of the patent system, firms frequently combine patenting with secrecy "because innovations are rarely composed of a monolithic piece of knowledge" (Anton *et al.* (2006)). For instance, the debates on waiving intellectual property rights for Covid-19 vaccines emphasized that, in addition to the information in patents, a successful replication of the mRNA technology requires access to the trade secrets and technical know-how about it (Price II *et al.* (2020)).[3] In general, firms prefer patenting for knowledge that is codified and can be reverse-engineered, and secrecy for knowledge that is tacit and easier to hide (Roin (2005), Hall *et al.* (2014)).

I develop and test a theory of knowledge sharing between firms that accounts for observed citation patterns. The idea behind the theory is as follows. A firm has incentives to share valuable secrets with a producer of complementary products such as an input supplier. For example, in Table 1.1 Amkor Technology is a supplier to IBM. In turn, the supplier might want to share these secrets with its other customers because by increasing the customers' profits it can charge them a higher input price. Depending on whether the supplier leaks private knowledge to others or not, knowledge flows are diffused or concentrated. This outcome depends on the market structure and types of contracts the supplier signs with the owner of secrets.

The key assumption in the theory is that contracts on knowledge sharing are incomplete (Arrow (1962)). Courts cannot verify the quality of secrets companies share with each other, so a firm cannot sell its secrets to a competitor when it is mutually beneficial for them. Once a buyer of secrets pays for the knowledge, a seller has little incentives to

[3]Appendix A.1 provides legal background and more case studies on how firms combine patenting and secrecy.

disclose them, and, conversely, once the seller reveals its secrets, the buyer might refuse to pay. In contrast to competitors, firms in a vertical relationship have incentives to share secrets because they produce complementary products and can use an input price to split the gains from knowledge sharing.

I assume that the firm with valuable trade secrets can sign two types of contracts with a supplier to prevent a leakage of its knowledge: a confidentiality agreement and exclusive dealing. Although courts cannot verify the quality of secrets firms share, they can verify the communication between parties, and a confidentiality agreement restricts the supplier's ability to discuss topics related to the firm's technology with other companies. Exclusive dealing eliminates the supplier's incentives to leak secrets by restricting its ability to sell an input to other companies.

In equilibrium, the firm and its supplier sign a contract that maximizes their joint surplus which consists of the firm's gross profits and the supplier's input sales to other firms. If downstream competition is high, they sign exclusive dealing to give the downstream firm an advantage over its competitors. This mechanism is similar to the foreclosure literature (Rey and Tirole (2007)). If downstream competition is low, but the firm's competitors have a strong outside option to the supplier's input, the firm and the supplier sign a confidentiality agreement. They prefer to sell inputs to others due to low competition but to keep the firm's knowledge secret because the supplier cannot extract enough gains from sharing it with other firms. In both cases, the supplier does not leak the firm's secrets, so knowledge flows are concentrated. If competition is low and competitors have a weak outside option, a knowledge leakage to other companies generates greater gains in input sales to the supplier relative to losses in the firm's profits. Therefore, the supplier and the firm sign neither a confidentiality agreement nor exclusive dealing, and knowledge flows are more diffused.

The theory can be generalized to other types of business connections where firms engage in complementary activities. For example, I also consider the framework with a

common customer buying from several suppliers. I focus on a vertical relationship between firms in order to test the theory using Compustat Segments data on supplier-customer relationships between publicly traded U.S. firms (Cohen and Frazzini (2008), Barrot and Sauvagnat (2016)).

The theory leads to several testable predictions. First, I show that a vertical supply chain relationship increases the probability of a citation between firms relative to similar firms with no production relationship. Second, firms sharing a common supplier or customer are also more likely to cite each other relative to firms with exclusive suppliers or customers. Third, the probability of a citation between firms sharing a common supplier is increasing in the bargaining position of this supplier which is proxied by its size relative to other firms in the supplier's industry. The relative size of the supplier approximates the weakness of customers' outside options of not working with it.[4]

The model also predicts that a higher level of competition leads to lower knowledge sharing. I show that the decline in knowledge sharing after 2000 (the rise in the concentration of citations in Figure 1.1) is greater in the technologies more exposed to import competition from China, where the U.S. imports are instrumented by Chinese exports to other high-income countries (Autor *et al.* (2020)). Moreover, the competition affected citation patterns of firms but not of non-corporate organizations (e.g., universities and government agencies), so its effect on the decline of knowledge sharing is unlikely to be driven by the correlation between general technological changes and globalization.

Finally, I provide suggestive estimates of the value of access to private knowledge. I show that firms responsible for the majority of citations to a highly-cited patent grow around 1% faster in sales and profits. Their citing patents generate around 15% more forward citations and around 3% greater abnormal stock market response on the grant day

[4]In the framework with a common customer, I assume that the suppliers' outside options of not working with it are zero. Therefore, this framework does not give empirical predictions about the suppliers' outside options, which can be proxied by the customer's relative size in its industry. Considering a more general downstream market structure would complicate the model, and I leave it for future research.

of a patent (Kogan *et al.* (2017)). So, the access to private knowledge of a firm (proxied by the concentration of citations) is highly valuable, and this provides a rationale for a potential policy intervention.

I discuss welfare implications of restrictions on the use of confidentiality agreements and/or exclusive contracts. The upside of this policy is that it might promote knowledge diffusion in the economy. The downside of restrictions on the use of confidentiality agreements is that firms might start using exclusive dealing to prevent knowledge leakages. The downside of restrictions on both exclusive contracts and confidentiality agreements is that firms in a vertical relationship might stop sharing knowledge with each other. I provide conditions on the market structure under which this policy generates welfare gains.

Literature

Patent citations are often used to measure knowledge spillovers between firms and inventors (e.g., Jaffe and Trajtenberg (2002), Jaffe and de Rassenfosse (2019)). The evidence in this paper indicates that patent citations are more likely to reflect intentional exchange of information between business partners rather than unintentional knowledge externalities.

On the theory side, this paper studies cooperation (creative construction) between firms who *intentionally* share private knowledge with each other to improve their joint production, and highlights the role of production networks in the diffusion of knowledge across competitors. This approach differs from the Schumpeterian literature on economic growth (e.g., Aghion *et al.* (2014)) which is focused on competition between firms ("creative destruction") in the environment with knowledge spillovers. Rosenberg (1963) was one of the first to emphasize the importance of input suppliers in the technological progress and described how in the nineteenth century specialized machine tool suppliers disseminated knowledge across industries about the metal working technology developed in the armament manufacture. The importance of vertical knowledge flows is also highlighted in the literature

on R&D cooperation (e.g., Cassiman and Veughelers (2002)) and foreign direct investment (e.g., Aitken and Harrison (1999), Blalock and Gertler (2008), Alfaro-Ureña *et al.* (2021), Bai *et al.* (2021)).

Much of the literature on intellectual property (IP) protection treats patenting and secrecy as substitutes (e.g., Hall *et al.* (2014)). I argue that the combination of patents and trade secrets might be used to protect the same technology. This is in line with the surveys of firms (e.g., Cohen *et al.* (2000)), the management literature (e.g., Amara *et al.* (2008)), legal research (e.g., Jorda (2008)), and case studies on IP protection in the chemical industry and drugs (e.g., Arora (1997), Price II *et al.* (2020)).

Citations from patents to scientific publications are more likely to reflect knowledge externalities than citations to other patents (Roach and Cohen (2013)). Arora *et al.* (2021) estimate how knowledge spillouts to competitors affect firms' incentives to invest in scientific research, measured by the number of publications. I complement their paper by studying how import competition from China affected the firms' choice of protection for existing trade secrets, measured by the concentration of patent citations. Both the increase in the concentration of patent citations documented in this paper and the decline in corporate science over time (Arora *et al.* (2019)) are consistent with the rising role of trade secrets.

The empirical tests of the theory analyze citation patterns in the production network and contribute to the literature on knowledge flows between firms (e.g., Gomes-Casseres *et al.* (2006)). Autor *et al.* (2020) show that trade with China reduced corporate patenting in the U.S. I complement their evidence by showing that trade with China also reduced knowledge sharing between firms. The paper also contributes to the literature on patents' quality (e.g., Kogan *et al.* (2017), Higham *et al.* (2020)) by estimating it as a function of the knowledge inventors rely on.

On the policy side, this paper contributes to the existing recommendations for promoting innovation (e.g., Bloom *et al.* (2019)) by discussing the costs and benefits of regulations

on confidentiality agreements and exclusive dealing between suppliers and customers. The model and policy analysis are related to the antitrust literature on exclusionary vertical contracts (e.g., Whinston (2006)) but with a focus on the incentives to share knowledge. I complement this literature by discussing the role of confidentiality agreements.

1.2 Motivating Evidence: Concentration of Citations

This section provides several new facts about patent citations. Section 1.2.1 documents that even for the most cited patents the majority of citations come from one firm only, and this concentration has significantly increased since 2000. In Section 1.2.2, I show that the high concentration of citations is primarily driven by differences across firms in citation rates (probabilities) rather than by differences in the number of patents. In Section 1.2.3, I show that inventors who move between companies significantly change their citation rates, meaning that citation probabilities are firm-specific rather than inventor-specific. Section 1.2.4 provides robustness checks, and Section 1.2.5 discusses the implications of the results.

Data and Background

Patents are supposed to facilitate knowledge sharing through the disclosure of information in the award. The U.S. Supreme Court stated that patent disclosures "will stimulate ideas and the eventual development of further significant advances in the art" and that these "additions to the general store of knowledge are of such importance" that they are worth the "the high price of ... exclusive use."[5]

Patents consist of two parts: a written description of an invention including citations to prior art (patents, publications, etc), and claims defining the boundaries of intellectual

[5]See *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 481 (1974) and Ouellette (2012).

property rights. To be patentable, inventions should be patent-eligible, useful, novel, non-obvious, and the text of the application should satisfy the disclosure requirements.[6] Patent examiners use references to prior art to check whether the invention is novel and non-obvious. In the U.S., applicants have a "duty of candor" to disclose relevant prior art that they are aware of, and the failure to do so can lead to patent invalidation. In general, prior art is used to strengthen, narrow or reject certain claims. Therefore, citations serve a legal function of delimiting intellectual property rights on an invention.

I use the data on utility patents granted in the U.S. Patent and Trademark Office (USPTO) for the period 1976–2019. Most granted patents contain information about assignees (patent owners). I clean assignee names to group patents by firms, individual inventors, universities and other organizations. Autor *et al.* (2020) provide a matching of patent assignees to names of publicly traded firms in Compustat data set. I extend their matching for additional years 2015–2019. Details are given in Appendix A.2.1.

1.2.1 Results: High Concentration of Patent Citations

For each year and technological class in the period 1976–2014, I track citations within a five-year window for the top 5% of the most cited granted patents. There are two reasons to focus on the set of the most cited patents. First, the value of patents is highly skewed with many of no value to the firm, so the empirical literature on innovation is often focused on the most cited patents (e.g., Aghion *et al.* (2021)). This is justified by the positive correlation between the number of patent citations and the firm's stock market valuation (Hall *et al.* (2005), Kogan *et al.* (2017)). Second, by construction these patents are expected to generate most follow-on innovations and knowledge flows, providing a lower bound on the concentration measure. Taking a five-year window controls for the truncation bias

[6]These requirements are governed by the US Code, Title 35, sections 101, 102, 103, 112. For a review, see Scotchmer (2004), ch. 3.

that older patents have more time to accumulate citations. Comparing patents within each technological class controls for differences across classes in citation patterns (Lerner and Seru (2021)). To classify technologies, I use the group level of Cooperative Patent Classification (CPC) system which has 672 groups.[7]

For each patent among the most cited ones, I compute the distribution of citations across different organizations. Denote by $n_{k,i}$ the number of citations from organization i to patent k. Organizations are mostly firms but also include non-corporate entities such as government agencies and universities. I exclude citations from individual inventors and from patents with missing assignee information. The concentration measure for patent k is the share of citations coming from the most citing organization:

$$\mathcal{C}_k = \max_i \left\{ \frac{n_{k,i}}{\sum_{j=1}^{M_k} n_{k,j}} \right\} \tag{1.2.1}$$

where M_k is the number of citing organizations. The most citing organizations are predominantly corporate firms, so I will use terms "firms" and "organizations" interchangeably. To construct an aggregate measure, for each year I take the average of patents' concentration measures within each technology class and then the average across technological classes weighted by the number of patents in a class. Appendix A.2.2 provides more details.

Figure 1.1 on page 3 shows the resulting aggregate concentration measure. On average, a patent (among the most cited ones) granted between 1976 and 2000 received around 50% of citations from one firm only. This concentration has significantly increased since 2000: a patent granted in 2014 received around 75% of citations from one firm only. In Figure A.1 (Appendix A.3), I show that the increase in the concentration is primarily

[7]The group level of CPC system is similar to the three-digit USPC system. There are several reasons to use modern CPC system instead of the more outdated USPC system. First, CPC system provides a consistent classification over time. Second, it is available for all years while USPTO stopped updating USPC from 2015. Third, CPC system is applied internationally, not only in the U.S. Finally, its classification provides more levels of disaggregation in contrast to USPC. I will use this feature in Section 1.4.1.

driven by changes within technological classes rather than the rise of technologies with high concentration of citations. An increase in the concentration after 2000 is observed in 87% of classes.

1.2.2 Sources of the Concentration of Citations

In this subsection, I decompose the concentration of citations into three sources: differences in citation rates (probabilities) across firms, differences in patenting across firms, and the total number of firms that could potentially make citations. The number of firms provides a natural lower bound on the concentration measure: with M firms the concentration cannot be lower than $1/M$. I show that the high concentration of citations is primarily explained by the differences in citation probabilities rather than by the variance in patenting across firms. For instance, Amkor Inc is responsible for 94% of citations to the IBM's patent in Table 1.1 because each of Amkor's patents makes a citation to the IBM's patent with a significantly higher probability relative to other firms, not because Amkor has more patents than other companies.

The details of the decomposition are the following. First, I divide all granted patents into disjoint groups based on common observational characteristics. Patents that share the same characteristics are more likely to cite similar patents. It is well-known that citation patterns depend on the geographical location of inventors (e.g., Jaffe *et al.* (1993)), proximity of technologies (e.g., Acemoglu *et al.* (2016)) and time differences between inventions (e.g., Caballero and Jaffe (1993)). So, for characteristics of patents I choose a geographical location of the majority of inventors, a detailed technology class (main subgroup level in CPC), and an application year.[8] Denote the set of all these groups by $\mathcal{H}$.

Second, for each cited patent I construct the set of patents that could potentially cite

[8]The location is defined at the state level in the U.S. and at the country level for foreign inventors. For patents with inventors from multiple locations I take the location of the majority of inventors. See Appendix A.2.2.

it and rely on the knowledge disclosed in it. This set includes patents with the same characteristics as the patents that actually make citations to the cited patent. Denote by $\mathcal{H}_k \subseteq \mathcal{H}$ the set of the characteristics of the patents citing patent k. For example, patents making citations to the IBM's patent in Table 1.1 ($k = 5877043$) are divided in $|\mathcal{H}_k| = 106$ disjoint groups based on their characteristics. Most citations (10 out of 218, all from Amkor) come from the following group ($h \in \mathcal{H}_k$): technological class $H01L23$, location of inventors Arizona, and application year 2003. Overall, there were 92 patents with these characteristics coming from firms such as Intel (62), Amkor (12), Freescale Semiconductor (5), and others (13).[9] If each of these patents made a citation to the IBM's patent with the same probability, more citations would be expected to come from Intel rather than from Amkor.

Consider citations to patent k from patents with characteristics $h \in \mathcal{H}_k$. The number of citations from firm i, $n_{k,i}(h)$, can be decomposed as follows

$$n_{k,i}(h) = p_{k,i}(h) \cdot N_i(h)$$

where $N_i(h)$ is the total number of patents with characteristics $h \in \mathcal{H}_k$ (citing and non-citing) from firm i, and $p_{k,i}(h) = \frac{n_{k,i}(h)}{N_i(h)}$ is the citation rate. I construct two counterfactual distributions of the concentration measure. First, I do Monte-Carlo simulations where $n_{k,i}(h)$ citations are allocated randomly across all patents with the same characteristics $h \in \mathcal{H}_k$. This exercise equates citation rates across firms that have patents with characteristics $h \in \mathcal{H}_k$ ($p_{k,i}(h) = p_{k,j}(h)$ for $i \neq j$) and shows the counterfactual concentration that is entirely driven by the differences in the number of patents across companies and the total number of firms. Second, I allocate $n_{k,i}(h)$ citations randomly across firms assuming that all firms with patents in $h \in \mathcal{H}_k$ have the same number of patents. This exercise equates both the citation rates and the number of patents across firms ($p_{k,i}(h) = p_{k,j}(h)$

[9]For illustration, I consider citations from all years to the IBM's patent in Table 1.1, not only within a 5-year window of the grant day. The aggregate results are based on citations within a 5-year window.

and $N_i(h) = N_j(h)$ for $i \neq j$) and shows the counterfactual concentration that is entirely driven by the number of firms that could potentially cite patent k.[10]

For each patent k, I do the Monte-Carlo simulations described above for all characteristics of citing patents, $\mathcal{H}_k$. Then I compute the counterfactual concentration measures in which citation rates are equal, and in which both citation rates and the number of patents are equal across firms. I aggregate these measures across cited patents in the same way as with the actual concentration in Section 1.2.1.

Figure 1.2 shows the decomposition of the concentration of citations. The upper solid line shows the actual aggregate concentration ($\mathcal{AC}_t$). The dashed and the dotted lines in the middle show the 95th quantile and the median of the concentration measure where citation probabilities are equalized across firms (denote by $\mathcal{RC}_t$ the median). Finally, the crossed line at the bottom shows the median concentration with equalized citation rates and equalized patenting across firms ($\mathcal{PC}_t$). This measure provides a natural lower bound on the concentration of citations that is driven by the number of firms. The average decomposition over all years from 1976 to 2014

$$\underbrace{\overline{\mathcal{AC}} - \overline{\mathcal{PC}}}_{17.3\%} = \underbrace{\overline{\mathcal{AC}} - \overline{\mathcal{RC}}}_{14.1\%} + \underbrace{\overline{\mathcal{RC}} - \overline{\mathcal{PC}}}_{3.3\%}$$

shows that the concentration is primarily explained by the differences across firms in citation probabilities rather than by the variance in the number of patents. The actual and counterfactual concentrations also have different dynamics. Both $\mathcal{PC}_t$ and $\mathcal{RC}_t$ declined until 2000 and then started to increase. This means that the average number of firms that

[10]Monte-Carlo exercises can be described using the urn model in probability theory (Feller (1968)). Denote the number of firms that have at least one patent with characteristics $h \in \mathcal{H}_k$ by $M(h)$. The exercise that equates citation rates only can be described as the following urn model of random drawings without replacement: $\sum_i n_{k,i}(h)$ balls are taken from the urn that contains $\sum_{i=1}^{M(h)} N_i(h)$ balls of $M(h)$ different colors with the distribution of colors $\{N_i(h)\}_{i=1}^{M(h)}$. The exercise that equates both the citation rates and the number of patents can be described as the following urn model of random drawings with replacement: $\sum_i n_{k,i}$ drawings are taken from the urn that contains $M(h)$ balls of different colors. After each drawing, the ball is returned back to the urn.

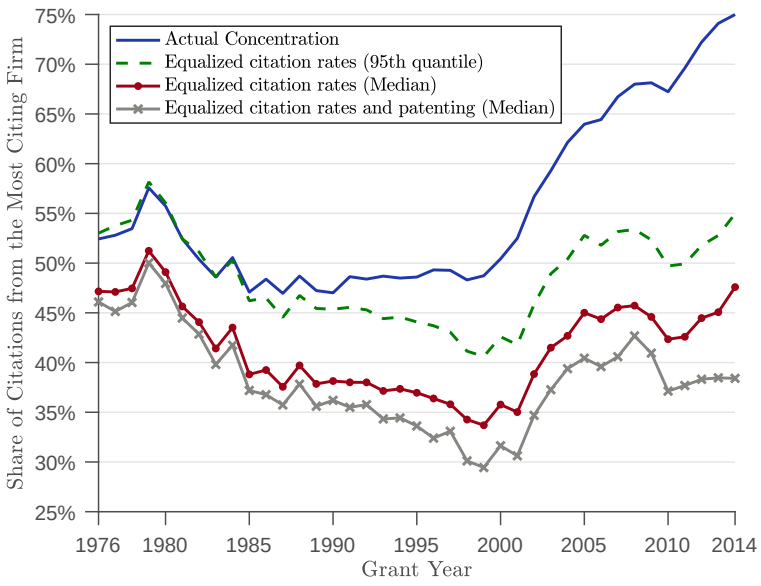


Figure 1.2: Decomposition of the Concentration of Citations

This figure compares the actual concentration of citations (the upper solid line) with the counterfactual ones in which citation rates are equalized across firms (the dashed and the dotted lines in the middle), and in which both citation rates and the number of patents are equalized across firms (the crossed line at the bottom). These counterfactual concentration measures are constructed using Monte-Carlo simulations in which citations are allocated randomly across observationally similar patents (equalized citation rates) and across firms with observationally similar patents (equalized citation rates and patenting). The details are given in Appendix A.2.2.

filed for patents with similar characteristics increased until 2000 and then decreased.[11] The actual concentration $\mathcal{AC}_t$ stayed approximately constant until 2000 and then started to increase at a faster rate than $\mathcal{PC}_t$ and $\mathcal{RC}_t$. This implies that the role of the differences in citation probabilities was increasing during all years

$$\underbrace{\Delta\mathcal{AC} - \Delta\mathcal{PC}}_{30.2\%} = \underbrace{\Delta\mathcal{AC} - \Delta\mathcal{RC}}_{22.1\%} + \underbrace{\Delta\mathcal{RC} - \Delta\mathcal{PC}}_{8.1\%}$$

where $\Delta\mathcal{AC} = \mathcal{AC}_{2014} - \mathcal{AC}_{1976}$, $\Delta\mathcal{RC} = \mathcal{RC}_{2014} - \mathcal{RC}_{1976}$, and $\Delta\mathcal{PC} = \mathcal{PC}_{2014} - \mathcal{PC}_{1976}$.

An important limitation in the construction of the counterfactual concentrations is that

[11] The difference $\mathcal{RC}_t - \mathcal{PC}_t$ has slightly increased over time, meaning that the variance in patenting has slightly increased from 1976 to 2014.

citations are randomized across patents similar to the ones that actually make citations, excluding the ones that could have potentially made citations. For example, for the IBM's patent in Table 1.1 most citations come from Amkor Technology, which is a supplier to IBM. Therefore, the citations are allocated randomly across the patents that are similar to the Amkor's patents and are likely to represent non-competing technologies to IBM. So, the randomization exercise does not take into account a lot of patents from IBM's competitors that could have made citations to it.

The counterfactual concentrations are constructed using the randomization of citations across patents that have the same detailed characteristics. Considering broader characteristics would lead to lower counterfactual concentrations. For example, the main subgroup in CPC is much more detailed level of technology classification than the one used in the literature.[12] Also, a patent file is a publicly available source of information to all inventors in the world. Figure A.2 (Appendix A.3) shows the counterfactual concentration where citations are randomized across broader technological classes, and the concentration where inventors are not required to be from the same location. These counterfactual concentrations are lower by 6 and 16 percentage points, respectively, relative to the baseline one. Figure A.3 also shows that the results of the decomposition are robust when I group patents based on their textual similarity (Arts *et al.* (2018)) rather than based on the classification into technological classes.

As another benchmark, I also compute the concentration of citations for scientific publications using Microsoft Academic Graph data (Marx and Fuegi (2020)). The concentration measure is constructed in a way similar to the measure for patents. For each year and field of science in the period 1976–2014, I track citations within a five-year window for the top 5% of the most cited publications. I use Web of Science classification

[12]The main subgroup level has 7137 detailed categories while the literature (e.g., Jaffe *et al.* (1993) and Bell *et al.* (2019)) often considers technologies to be similar if they come from the same 3-digit USPC or NBER sub-class classifications, which have 876 and 445 categories, respectively.

of scientific disciplines. Appendix A.2.2 provides details. Figure A.2 shows that in contrast to patents the concentration of citations for scientific publications declined from 8% to 4% over the same time period.

1.2.3 Movement of Inventors

The concentration of citations might be driven by inventors rather than firms. For instance, suppose a patent received all of its citations from one inventor who always worked in one firm. Then, the concentration of citations across companies for this patent would be 100%, but it could reflect the inventor's citation practice. To separate whether citations are firm- or inventor-specific, I find inventors who filed similar patents in multiple companies. Then, I compute the concentration of citations within these inventors and do the decomposition from Section 1.2.2 to check whether inventors' citation rates (probabilities) differ across firms.

Formally, consider the following statistical framework. Suppose inventor ℓ worked in two companies, i and j, and created $N_i^\ell(h)$ and $N_j^\ell(h)$ patents with characteristics $h \in \mathcal{H}$, respectively. Assume that each patent in firm i (j) makes an independent citation to patent k with probability $p_{k,i}^\ell(h)$ ($p_{k,j}^\ell(h)$). Then the expected number of citations from inventor ℓ in firm i to patent k is

$$n_{k,i}^\ell(h) = p_{k,i}^\ell(h) \cdot N_i^\ell(h),$$

and the goal is to test whether $p_{k,i}^\ell(h) = p_{k,j}^\ell(h)$. To do this, I compare the actual concentration of citations within an inventor with the counterfactual one where citation probabilities are equalized across companies ($p_{k,i}^\ell(h) = p_{k,j}^\ell(h)$).

For example, during 2008 to 2017 inventor Stefan G. Schreck from California created 16 patents in the technological class $A61F2$ while working in Endologix Inc. In 15 out of 16 patents, he made a citation to patent 5690642 assigned to Cook Incorporated. He also

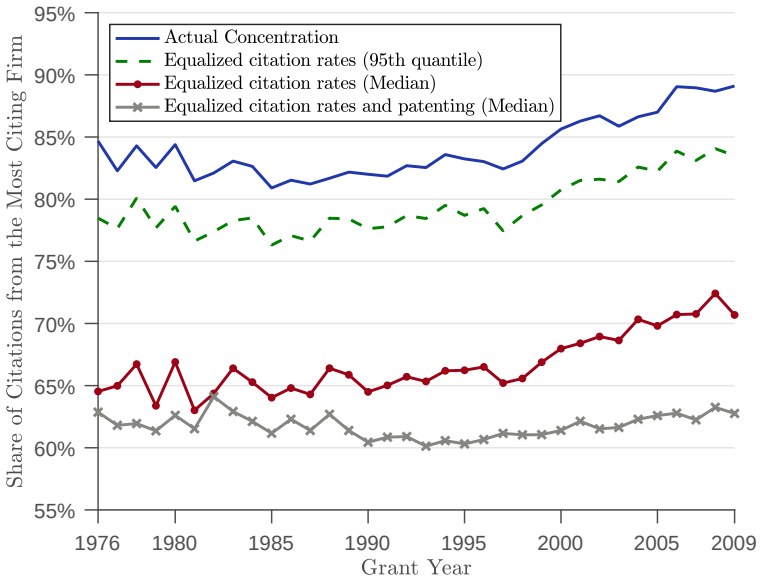


Figure 1.3: Decomposition Within Inventors Who Moved Across Firms

This figure shows the decomposition for the concentration of citations across firms within inventors who patented for multiple companies. The solid line shows the actual aggregate within-inventor concentration of citations across firms. The dashed and the dotted lines show the 95th quantile and the median of the same measure in the Monte-Carlo simulations where citations rate are equalized across firms within an inventor. The crossed line at the bottom shows the concentration (median) where both the citation rates and the number of patents are equalized across firms within an inventor. The averages are almost the same as the medians. To increase the sample size I consider citations from all years, not only 5-year window. The graph is taken until 2009 to ensure that patents have enough time to accumulate citations from inventors-movers. More details are given in Appendix A.2.3.

applied for 9 patents with similar characteristics in another company, Edwards Lifesciences Corporation, but made zero citations to patent 5690642. If 15 citations were allocated randomly across $16 + 9 = 25$ patents, the expected share of citations from Endologix Inc would be $^{9.5}/_{15} = 0.63$ and the 95th quantile would be $^{12}/_{15} = 0.8$, but the actual share ($^{15}/_{15} = 1$) is significantly higher. Notice that for an inventor who worked in two companies the concentration (the share of citations from the most citing firm) cannot be less than 50%.

I compute the average concentration of citations across firms within all inventors who moved between companies, and do the decomposition from Section 1.2.2. The details are

given in Appendix A.2.3. Figure 1.3 shows the actual average concentration within an inventor ($\mathcal{AC}_t^w$), the 95th quantile and the median of the concentration with equalized citation rates across firms ($\mathcal{RC}_t^w(q95)$ and $\mathcal{RC}_t^w$), and the median concentration in which both citation rates and patenting are equalized across firms ($\mathcal{PC}_t^w$).[13] The actual average concentration within an inventor is significantly higher relative to the what we would expect if citation probabilities were equalized across firms ($\mathcal{AC}_t^w > \mathcal{RC}_t^w(q95)$). The average decomposition over all years

$$\underbrace{\overline{\mathcal{AC}}^w - \overline{\mathcal{PC}}^w}_{22.2\%} = \underbrace{\overline{\mathcal{AC}}^w - \overline{\mathcal{RC}}^w}_{17.3\%} + \underbrace{\overline{\mathcal{RC}}^w - \overline{\mathcal{PC}}^w}_{4.9\%}$$

shows that the concentration is primarily explained by the differences across firms in citation probabilities rather than by the variance in the number of patents. The difference $\mathcal{AC}_t^w - \mathcal{RC}_t^w$ is stable over time, and there was a slight increase in the difference $\mathcal{RC}_t^w - \mathcal{PC}_t^w$, meaning that the dispersion in the number of patents across firms within an inventor has slightly increased by the end of the period.

This evidence should be interpreted with caution because inventors-movers might differ in their citation rates from inventors who always work in one company. My conjecture is that the non-movers would have higher concentration of citations across firms if they were randomly moved to another company. Below I argue that citations are correlated with access to trade secrets. Based on this interpretation, the conjecture is that inventors-movers are less restricted by contractual arrangements such as confidentiality agreements and non-compete contracts, and this leads to less concentrated citations. An interesting area for future research is to study the movement of inventors caused by exogenous shocks to firms, for example, natural disasters (Barrot and Sauvagnat (2016)) or financial constraints

[13]The variable $\mathcal{PC}_t^w$ is greater than 50% because the majority of inventor-movers worked in two firms only. For example, with two citations randomly allocated across two firms the expected concentration measure is

$$\frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 0.5 + \frac{1}{4} \cdot 0.5 + \frac{1}{4} \cdot 1 = 0.75$$

(Chodorow-Reich (2014)).

1.2.4 Robustness

Appendix A.2.4 provides multiple robustness checks. In particular, the results are robust when I consider different thresholds for the most cited patents (top 1% and top 10%). The concentration is not driven by superstar firms in patenting. The concentration is robust to the exclusion of firms' self-citations to themselves, so it is driven by citations between firms rather than self-citations. The concentration is also robust when I group citations of patents from the same within-country family (continuations, continuations-in-part, divisionals) as a single citation, so its rise is not driven by increasing patent families. Citation patterns might be affected by patent examiners (Alcácer *et al.* (2009)). I show that the concentration of citations from patent examiners is around two times lower than the concentration based on citations from non-examiners. Citations might also be affected by patent lawyers. Using citations of lawyers who represented several companies and the methodology similar to the movement of inventors, I show that citations are not lawyer-specific, and the concentration is driven by firms. Appendix A.2.4 provides more robustness checks including controls for outliers, different samples, and weighting schemes.

1.2.5 Discussion and Implications for the Theory

Knowledge spillovers might come from different sources. The U.S. courts argue that patents should provide a source of knowledge to other inventors (Roin (2005)), and the same argument is used in Romer (1990, p. 84). The literature on urban economics and technology diffusion argues that spillovers come from interactions between inventors (Duranton and Puga (2004), Perla and Tonetti (2014), Buera and Lucas (2018)). The evidence above raises a concern whether patent citations can be used to capture these types of knowledge spillovers. First, the comparison of the actual concentration with the random

one indicates that that there are significant differences in citation rates across firms with very similar patents. Second, the evidence on the movement of inventors indicates that citation rates are different across firms even within the same inventor.

I argue that patent citations might reflect cooperation between firms and exchange of tacit trade secrets that are complementary to codified knowledge disclosed in patents. As a result, citations are concentrated because only a few companies get access to private knowledge of a patent owner. In theory, due to the disclosure requirements one invention cannot be protected by both patents and trade secrets, but in practice "[b]ecause innovations are rarely composed of a monolithic piece of knowledge, a combination of patenting and secrecy is common" (Anton *et al.* (2006)). Appendix A.1 describes several examples on how firms combine secrecy and patenting. In particular, I discuss the examples of mRNA technology for Covid-19 vaccines and the Haber-Bosch process for ammonia. I also provide several legal cases on why and when the combination of patenting and secrecy does not violate the disclosure requirements of patents.

The evidence from this section motivates the structure of the theoretical framework I develop below. First, the theory is aimed to explain the differences in citation rates across firms rather than the differences in the number of patents. Second, the explanation for the differences in citation rates is based on connections between firms and their incentives to share private knowledge rather than on connections between inventors who work in these firms, and how they interact with each other.

1.3 Theory

The grant of monopoly through patents is justified by the claim that it promotes knowledge diffusion through the disclosure of inventions. However, the evidence in Section 1.2 indicates that even for the most cited patents only a limited set of firms builds on the patents' knowledge. I argue that citations are associated with the sharing of trade secrets that

accompany patents and develop a theory of knowledge sharing between firms that gives conditions under which knowledge flows are diffused or concentrated.

The idea behind the theory can be illustrated with a case study on the diffusion of knowledge about the copper interconnect technology for semiconductor chips (Lim (2009)). IBM was one of the leaders in the development of this technology and relied heavily on secrecy to protect it. IBM shared its secrets with an input supplier Novellus who developed a copper deposition tool. To protect its secrets, IBM and Novellus signed a contract that restricted Novellus's ability to discuss its partnership with IBM and to sell a deposition tool to others:

> "Under this agreement, Novellus was not permitted to reveal that it was working with IBM. . ., and IBM maintained a list of companies to which Novellus could not initially sell the tool."

In the model, IBM acts as an incumbent firm ($\mathcal{I}$) who shares its private knowledge with an input supplier ($\mathcal{S}$), Novellus. To protect its secrets from leakages to others, $\mathcal{I}$ and $\mathcal{S}$ sign a confidentiality agreement and exclusive dealing. As a result, knowledge flows are concentrated. Around 1997, IBM started to become more open about the development of the copper interconnect technology and filed for patents. Table A.1 (Appendix A.3) shows the distribution of citations for one of the most cited patents in this technology assigned to IBM and Novellus.[14] Citations for this patent are highly concentrated and mostly come from IBM's suppliers: 64% from Novellus and 30% from other suppliers.

Besides IBM, Motorola also worked on this technology. In contrast to IBM, Motorola was relatively open with its suppliers, and they diffused knowledge to others:

[14]The fact that IBM and Novellus filed for a joint patent provides another dimension of a close relationship between these firms. Only around 3% of patents granted in USPTO have more than one assignee. In Figure A.4, I document the rise in joint innovations (measured by joint patents) between input suppliers and customers. Both the share of suppliers and customers who applied for joint patents and the share of suppliers and customers among the patents with multiple assignees have increased from 1980 to 2010.

> "... Motorola worked jointly with suppliers from an early stage to develop tools for copper technology. A member of Motorola's copper team remarked that 'unlike IBM, we are quite open with vendors,...' Motorola's relative openness made it an important source of information for the rest of the industry..."

In the model, Motorola also acts as an incumbent firm who shares its private knowledge with an input supplier, for example, Applied Materials. However, the difference from IBM is that Motorola signs neither a confidentiality agreement nor exclusive dealing, so the supplier shares Motorola's knowledge with other firms, for instance, Taiwan Semiconductor Manufacturing Company (TSMC). TSMC entered into this technology much later relative to IBM and Motorola, and in the model it acts as an entrant ($\mathcal{E}$). According to interviews, "[TSMC] depended primarily on technical knowledge from equipment suppliers. TSMC worked closely with Applied Materials..." Table A.2 shows the distribution of citations for one of the most cited patents in this technology assigned to Motorola. In contrast to IBM's patent, citations to Motorola's patent are less concentrated and come from a diverse set of firms: 23% come from several suppliers (e.g, 7% from Applied Materials), and 39% come from 17 firms sharing a common supplier with Motorola (e.g., 5% from TSMC).

The goal of the theory is to derive conditions on the market structure and bargaining between firms under which business partners sign or do not sign different types of contracts that prevent knowledge flows in the economy. Section 1.3.1 presents a simple version of the model. Section 1.3.2 provides extensions and additional results to the baseline model. Section 1.3.3 lists the model's testable predictions.

1.3.1 Simple Model

Demand. There are two competing firms, an incumbent ($\mathcal{I}$) and an entrant ($\mathcal{E}$), who sell their products to a continuum of consumers of measure 1. Each consumer is one of three types. A share $1-\alpha$ of consumers buys a product from one firm only (loyal customers). Specifically, a share $\frac{1-\alpha}{2}$ buys from $\mathcal{I}$ only, and a share $\frac{1-\alpha}{2}$ buys from $\mathcal{E}$ only. Their

willingness to pay is determined by the quality of a product. So, $\mathcal{I}$-consumers are ready to pay up to $q_{\mathcal{I}}$, and $\mathcal{E}$-consumers are ready to pay up to $q_{\mathcal{E}}$, where $q_{\mathcal{I}}$ and $q_{\mathcal{E}}$ are product qualities of $\mathcal{I}$ and $\mathcal{E}$, respectively. A share α of consumers buys the product from a firm with the highest quality-price difference (switchers). Therefore, the parameter α measures the degree of competition between $\mathcal{I}$ and $\mathcal{E}$.

Note that this demand can be interpreted through the lenses of international trade where there is a continuum of measure 1 products, firms $\mathcal{I}$ and $\mathcal{E}$ are from different countries, and products with loyal customers represent non-traded goods. An increase in α can be interpreted as an increase in the set of goods that can be traded (e.g., due to the decline in trade costs).

Pricing to Consumers. Firms $\mathcal{I}$ and $\mathcal{E}$ can perfectly discriminate between different types of consumers.[15] So, they can charge loyal customers by their willingness to pay, $p_{\mathcal{I}} = q_{\mathcal{I}}$ and $p_{\mathcal{E}} = q_{\mathcal{E}}$. For the rest (switchers), $\mathcal{I}$ and $\mathcal{E}$ compete in prices. Therefore, these consumers will be served by a firm with a higher product quality at a price $p = \max\{q_{\mathcal{I}}, q_{\mathcal{E}}\} - \min\{q_{\mathcal{I}}, q_{\mathcal{E}}\}$ where I assume that in the case of indifference consumers buy from a firm with the highest quality. I also assume that consumers buy from $\mathcal{I}$ if $p_{\mathcal{I}} = p_{\mathcal{E}}$ and $q_{\mathcal{I}} = q_{\mathcal{E}}$.

Technology. Firm $\mathcal{I}$ has a valuable technology that consists of two types of knowledge: tacit and codified. Tacit knowledge is unpatentable and is easy to hide, so $\mathcal{I}$ keeps it secret. The incumbent needs an input in its production, and it has its own low-quality input or it can buy a high-quality one from an input supplier ($\mathcal{S}$). However, for $\mathcal{S}$ to produce a superior input, firm $\mathcal{I}$ has to share the *tacit* knowledge with it. Without this knowledge, $\mathcal{S}$ can only offer the same low-quality input as $\mathcal{I}$ has. Firm $\mathcal{I}$ can produce a final product of quality $q_{\mathcal{I}} = \lambda$ with a high-quality input from $\mathcal{S}$, and of quality $q_{\mathcal{I}} = (1-\gamma)\lambda$ with a low-quality input where $\gamma \leqslant 1$.

[15] With the interpretation based on international trade, this assumption means that firms can charge monopoly prices for non-traded products in their countries.

Firm $\mathcal{E}$ also needs an input, and it has its own low-quality input or it can buy an input from $\mathcal{S}$. The quality of $\mathcal{E}$'s product is $q_{\mathcal{E}} = \lambda_0$ with a superior input from $\mathcal{S}$, and $q_{\mathcal{E}} = (1-\gamma)\lambda_0$ with an inferior input where $\lambda_0 < \lambda$. However, using $\mathcal{I}$'s *tacit* knowledge it can innovate and improve the quality of its product to $q_{\mathcal{E}} = \lambda$ with a superior input from $\mathcal{S}$, and $q_{\mathcal{E}} = (1-\gamma)\lambda$ with an inferior input. The summary is given in Table 1.2.

Table 1.2: Quality of Final Products

Access to a high-quality input?	$\mathcal{I}$'s quality	$\mathcal{E}$'s quality Access to $\mathcal{I}$'s tacit knowledge?	
		Yes	No
Yes	$q_{\mathcal{I}} = \lambda$	$q_{\mathcal{E}} = \lambda$	$q_{\mathcal{E}} = \lambda_0$
No	$q_{\mathcal{I}} = (1-\gamma)\lambda$	$q_{\mathcal{E}} = (1-\gamma)\lambda$	$q_{\mathcal{E}} = (1-\gamma)\lambda_0$

This table shows the quality of $\mathcal{I}$'s and $\mathcal{E}$'s products as a function of whether they have access to a high-quality input from $\mathcal{S}$ and to $\mathcal{I}$'s tacit secrets. For $\mathcal{S}$ to produce an input of high quality it needs to know $\mathcal{I}$'s tacit secrets.

Firm $\mathcal{E}$ also values $\mathcal{I}$'s codified knowledge, which it can obtain through reverse-engineering of $\mathcal{I}$'s product. However, the codified knowledge without the tacit secrets allows $\mathcal{E}$ only to imitate $\mathcal{I}$'s quality for common customers (switchers) but not to improve the quality of its product for its loyal customers. Codified knowledge is patentable, so $\mathcal{I}$ protects it from imitation with a patent. From the firms' point of view, this imitation is a pure business stealing effect since it does not allow firms to capture more surplus from consumers. Therefore, the incumbent has no incentives to license it to $\mathcal{E}$. The idea that firms use patents to disclose knowledge that would be available to the public through other channels anyway is consistent with the arguments of legal scholars that the patent system fails its disclosure function (e.g., Roin (2005)).

To sum up, $\mathcal{I}$ patents its codified knowledge but keeps the tacit knowledge secret. The codified knowledge allows imitation only, and the tacit secrets are key for improving the

quality of $\mathcal{S}$'s input and $\mathcal{E}$'s product.

Citations. Both $\mathcal{S}$ and $\mathcal{E}$ apply for patents on their products that do not infringe $\mathcal{I}$'s patent. Without $\mathcal{I}$'s private knowledge they cite $\mathcal{I}$'s patent with probabilities $p_{\mathcal{S}}(0) \geqslant 0$ and $p_{\mathcal{E}}(0) \geqslant 0$, respectively. With $\mathcal{I}$'s private knowledge they cite $\mathcal{I}$'s patent with probabilities $p_{\mathcal{S}}(1) > p_{\mathcal{S}}(0)$ and $p_{\mathcal{E}}(1) > p_{\mathcal{E}}(0)$, respectively.

Contracting. The main contracting limitation is that courts cannot verify the quality of tacit knowledge firms share with each other and quality of their products, so $\mathcal{I}$ cannot license its tacit knowledge to $\mathcal{E}$ when it is mutually beneficial for them. Specifically, once $\mathcal{E}$ pays for knowledge, $\mathcal{I}$ has no incentives to disclose it and courts cannot check whether $\mathcal{I}$ shared relevant secrets with $\mathcal{E}$. Conversely, once $\mathcal{I}$ shares its secrets, $\mathcal{E}$ might refuse to pay claiming that the secrets are not valuable. This is a classical problem in the markets for ideas (Arrow (1962)).[16]

I assume that firms $\mathcal{I}$ and $\mathcal{E}$ can write contracts on inputs with $\mathcal{S}$. Moreover, $\mathcal{I}$ and $\mathcal{S}$ can also sign a confidentiality agreement and exclusive dealing. Although courts cannot verify the quality of secrets firms share with each other, they can verify the communication between parties. A confidentiality agreement restricts $\mathcal{S}$'s ability to discuss topics related to $\mathcal{I}$'s technology with $\mathcal{E}$. Exclusive dealing restricts $\mathcal{S}$'s ability to sell an input to $\mathcal{E}$.

Timing. The game has the following timing (see also Figure 1.4):

t1. $\mathcal{I}$ and $\mathcal{S}$ have Nash bargaining with 50/50 split of the surplus about the following contract:

- *Exclusivity* (ex). The contract states whether $\mathcal{S}$ can sell an input to $\mathcal{E}$ ($ex = 0$) or not ($ex = 1$).

[16] Arora *et al.* (2001) discuss how to license tacit knowledge to a non-competing firm by bundling it with a complementary patent. Once $\mathcal{E}$ receives $\mathcal{I}$'s secrets, it can improve the quality of its product without infringing $\mathcal{I}$'s patent, so the strategy for $\mathcal{I}$ to sell its secrets by bundling them with its patent does not work.

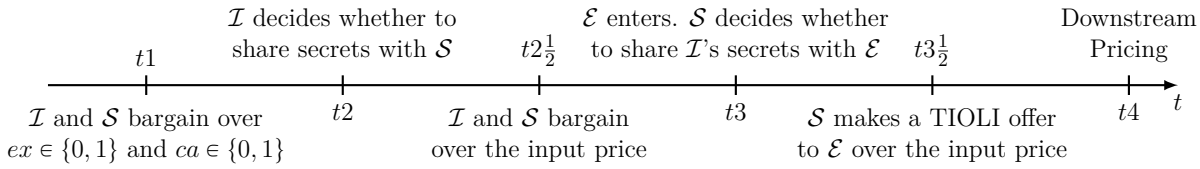


Figure 1.4: Timing

- *Confidentiality Agreement* (ca). The contract defines whether $\mathcal{S}$ can share $\mathcal{I}$'s knowledge with $\mathcal{E}$ ($ca = 0$) or not ($ca = 1$).

t2. Firm $\mathcal{I}$ decides whether to share its knowledge with $\mathcal{S}$. Then they have Nash bargaining over the input price ($t_{\mathcal{I}}$) with 50/50 split of the surplus.

t3. $\mathcal{E}$ enters. Suppose $\mathcal{I}$ shared its knowledge with $\mathcal{S}$. Without a confidentiality agreement between $\mathcal{I}$ and $\mathcal{S}$ ($ca = 0$), firm $\mathcal{S}$ decides whether to share $\mathcal{I}$'s secrets with $\mathcal{E}$. Then, under a non-exclusive dealing ($ex = 0$) it makes a take-it-or-leave-it offer about the input price ($t_{\mathcal{E}}$) to $\mathcal{E}$.

t4. Firms $\mathcal{I}$ and $\mathcal{E}$ set downstream prices, and profits are realized.

Discussion. An important assumption is that the entrant does not participate in the contracting between $\mathcal{I}$ and $\mathcal{S}$ in the first stage. For example, its identity is not known at this point. This will cause an inefficiency from the firms' point of view because $\mathcal{I}$ and $\mathcal{S}$ will maximize their joint surplus rather than the profits of all three firms. For more on bilateral contracting see Segal (1999).

Assumptions about specific production, demand and bargaining structures are not crucial for the results and will be relaxed in Section 1.3.2. For example, instead of the take-it-or-leave-it offer from $\mathcal{S}$ to $\mathcal{E}$ they could bargain over the input price. I will also consider a richer contracting space, for instance, the possibility of stipulated damages (Aghion and Bolton (1987)) in Section 1.3.2. Section 1.3.2 provides the framework with two suppliers and a common customer.

Subgame Perfect Equilibrium. If $\mathcal{S}$ and $\mathcal{I}$ do not sign an exclusive contract in the first stage, then in the third stage $\mathcal{S}$ will make a take-it-or-leave-it offer to $\mathcal{E}$ about the input price. The offer depends on whether firm $\mathcal{E}$ has access to $\mathcal{I}$'s tacit knowledge or not:

$$\begin{array}{l} \overbrace{t_{\mathcal{E}} = \frac{1-\alpha}{2} \cdot \lambda - \frac{1-\alpha}{2}(1-\gamma) \cdot \lambda = \frac{1-\alpha}{2}\gamma\lambda}^{\mathcal{E} \text{ has access to } \mathcal{I}\text{'s tacit knowledge (secrets)}} \\ \underbrace{t_{\mathcal{E}} = \frac{1-\alpha}{2} \cdot \lambda_0 - \frac{1-\alpha}{2}(1-\gamma) \cdot \lambda_0 = \frac{1-\alpha}{2}\gamma\lambda_0}_{\mathcal{E} \text{ does not have access to } \mathcal{I}\text{'s tacit knowledge (secrets)}} \end{array}$$

$\mathcal{S}$ always has incentives to share $\mathcal{I}$'s knowledge with $\mathcal{E}$ in the third stage because $\lambda > \lambda_0$. These incentives come from the complementarity between $\mathcal{S}$'s input and $\mathcal{I}$'s technology in the quality of $\mathcal{E}$'s product. A confidentiality agreement between $\mathcal{I}$ and $\mathcal{S}$ can prevent the knowledge leakage to $\mathcal{E}$ because it contractually restricts $\mathcal{S}$'s ability to communicate with $\mathcal{E}$. Exclusive dealing eliminates $\mathcal{S}$'s incentives to leak $\mathcal{I}$'s knowledge by restricting its ability to sell an input to $\mathcal{E}$.

In stage $t2$, $\mathcal{I}$ decides whether to share secrets with $\mathcal{S}$, and then they bargain over the input price. The upside of sharing secrets with $\mathcal{S}$ is that it improves the quality of $\mathcal{S}$'s input. The downsides are that $\mathcal{E}$ might also buy a high-quality input from $\mathcal{S}$ (under a non-exclusive dealing, $ex = 0$), and that $\mathcal{S}$ might leak $\mathcal{I}$'s secrets to $\mathcal{E}$ (under $ex = 0$ and $ca = 0$). For a contractual arrangement $(ex, ca) \in \{0,1\}^2$, $\mathcal{I}$ shares its tacit knowledge with $\mathcal{S}$ if

$$\begin{array}{c} \underbrace{\pi^o_{\mathcal{I}}(ex,ca) + \frac{1}{2}\Big(TS(ex,ca) - \pi^o_{\mathcal{I}}(ex,ca) - \pi^o_{\mathcal{S}}(ex,ca)\Big)}_{\mathcal{I}\text{'s profit under knowledge sharing with } \mathcal{S} \text{ and contract } (ex,ca) \in \{0,1\}^2} \quad \geqslant \\ \underbrace{\frac{1-\alpha}{2}(1-\gamma)\lambda + \alpha \cdot (1-\gamma)\left(\lambda - \lambda_0\right)}_{\mathcal{I}\text{'s profit without knowledge sharing with } \mathcal{S}} \end{array} \tag{1.3.1}$$

where the right-hand side shows $\mathcal{I}$'s profit when it does not share knowledge with $\mathcal{S}$. In this case, it works with an inferior input but so as $\mathcal{E}$, and $\mathcal{E}$ does not get $\mathcal{I}$'s secrets. The left-hand side shows $\mathcal{I}$'s profit when it shares secrets with $\mathcal{S}$. $\mathcal{I}$ and $\mathcal{S}$ bargain over the

input price and split the total surplus ($TS(ex, ca)$) 50/50 over the disagreement point $(\pi^o_{\mathcal{I}}(ex, ca), \pi^o_{\mathcal{S}}(ex, ca))$.

In equilibrium, there are three possible contractual arrangements: no contract between $\mathcal{I}$ and $\mathcal{S}$ in the first stage ($ex = 0, ca = 0$), a confidentiality agreement only ($ex = 0, ca = 1$), and exclusive dealing ($ex = 1, ca \in \{0, 1\}$). Under exclusive dealing, $\mathcal{S}$ is indifferent whether to share $\mathcal{I}$'s secrets with $\mathcal{E}$ in stage $t3$ or not, and I assume that it does not share them. Therefore, in this case a confidentiality agreement is redundant, and both arrangements ($ex = 1, ca = 0$) and ($ex = 1, ca = 1$) lead to the same outcome.[17]

In stage $t1$, $\mathcal{I}$ and $\mathcal{S}$ sign a contract (ex, ca) that maximizes their total surplus $TS(ex, ca)$ subject to constraint (1.3.1) that $\mathcal{I}$ has incentives to share its secrets with $\mathcal{S}$. Proposition 1 summarizes the Subgame Perfect Equilibrium.

Proposition 1. *The subgame perfect equilibrium contract between $\mathcal{I}$ and $\mathcal{S}$ is the following:*

$$\begin{cases} \text{Exclusive Dealing } (ex = 1, ca \in \{0, 1\}) & \text{if } \alpha > \frac{1}{3} \\ \text{CA only } (ex = 0, ca = 1) & \text{if } \min\left\{\frac{\gamma}{\gamma+2}, \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)}\right\} < \alpha \leqslant \frac{1}{3} \\ \text{Null } (ex = 0, ca = 0) & \text{if } \alpha \leqslant \min\left\{\frac{\gamma}{\gamma+2}, \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)}\right\} \end{cases} \tag{1.3.2}$$

Firm $\mathcal{I}$ always shares knowledge with $\mathcal{S}$, and $\mathcal{S}$ shares $\mathcal{I}$'s knowledge with $\mathcal{E}$ only under a non-exclusive dealing without a confidentiality agreement ($ex = 0, ca = 0$).

Proof. For each contract, the total surplus of $\mathcal{I}$ and $\mathcal{S}$ is the following

$$\text{Exclusive: } TS(ex = 1, ca \in \{0, 1\}) = \frac{1-\alpha}{2} \cdot \lambda + \alpha \cdot (\lambda - (1-\gamma)\lambda_0) + 0 \tag{1.3.3}$$

$$\text{CA only: } TS(ex = 0, ca = 1) = \frac{1-\alpha}{2} \cdot \lambda + \alpha \cdot (\lambda - \lambda_0) + \frac{1-\alpha}{2} \cdot \gamma\lambda_0 \tag{1.3.4}$$

$$\text{Null: } TS(ex = 0, ca = 0) = \frac{1-\alpha}{2} \cdot \lambda + \alpha \cdot (\lambda - \lambda) + \frac{1-\alpha}{2} \cdot \gamma\lambda \tag{1.3.5}$$

[17] If $\mathcal{I}$ expected a leakage of its knowledge when $\mathcal{S}$ is indifferent, they would sign a confidentiality agreement, but this would not change the structure of the equilibrium. Another way to break the indifference is to assume that there are small costs of knowledge sharing.

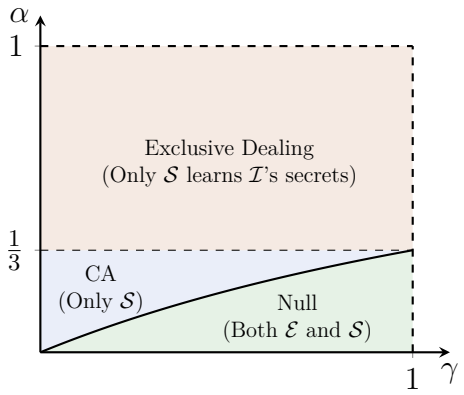


Figure 1.5: Equilibrium as a function of (α, γ)

This figure illustrates the equilibrium contract between $\mathcal{I}$ and $\mathcal{S}$ described in Proposition 1. Parameter α measures the degree of competition between $\mathcal{I}$ and $\mathcal{E}$, and parameter γ reflects the value of $\mathcal{S}$'s input relative to $\mathcal{I}$'s and $\mathcal{E}$'s outside options.

where the first two terms represent $\mathcal{I}$'s gross profits, and the last term represent $\mathcal{S}$'s profits from selling an input to $\mathcal{E}$. If $\alpha > \frac{1}{3}$, then $(ex = 1, ca \in \{0, 1\})$ maximizes total surplus $TS(ex, ca)$. If $\frac{1}{3} < \alpha \leqslant \frac{\gamma}{\gamma+2}$, then $(ex = 0, ca = 1)$ maximizes total surplus. In both cases, condition (1.3.1) is satisfied.[18] If $\alpha \leqslant \frac{\gamma}{\gamma+2}$, then $(ex = 0, ca = 0)$ maximizes total surplus, and condition (1.3.1) is satisfied only if

$$\underbrace{\frac{1-\alpha}{2}(1-\gamma)\lambda + \frac{1}{2}\cdot\frac{1-3\alpha}{2}\gamma\lambda}_{\mathcal{I}\text{'s profit under knowledge sharing}} \geqslant \underbrace{\frac{1-\alpha}{2}(1-\gamma)\lambda + \alpha\cdot(1-\gamma)(\lambda-\lambda_0)}_{\mathcal{I}\text{'s profit without knowledge sharing}}$$

which can be rewritten as

$$\alpha \leqslant \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda}) + \gamma(4\frac{\lambda_0}{\lambda} - 1)} \tag{1.3.6}$$

□

The equilibrium is illustrated in Figure 1.5, and the intuition behind Proposition 1 is

[18] Condition (1.3.1) is always satisfied for $(ex = 1, ca \in \{0, 1\})$ and is satisfied for $(ex = 0, ca = 1)$ due to simplifying functional form assumptions (see Appendix A.4.1). In Section 1.3.2, I consider a more general case.

the following. For product qualitites $(q_{\mathcal{I}}, q_{\mathcal{E}})$, the total sum of profits of all three firms is

$$\Pi_{all} = \frac{1-\alpha}{2}(q_{\mathcal{I}} + q_{\mathcal{E}}) + \alpha \cdot (q_{\mathcal{I}} - q_{\mathcal{E}}) = \frac{1+\alpha}{2} q_{\mathcal{I}} + \frac{1-3\alpha}{2} q_{\mathcal{E}}$$

For $\alpha > \frac{1}{3}$, it is decreasing in $q_{\mathcal{E}}$, so $\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing to minimize $q_{\mathcal{E}}$. This mechanism is similar to the foreclosure literature (Rey and Tirole (2007)). For $\alpha < \frac{1}{3}$, the sum of profits is increasing in both $q_{\mathcal{I}}$ and $q_{\mathcal{E}}$, and is maximized under a transfer of $\mathcal{I}$'s tacit knowledge (secrets) to $\mathcal{E}$, and when both firms use a high-quality input from $\mathcal{S}$. Since $\mathcal{S}$ makes a take-it-or-leave-it offer to $\mathcal{E}$ about the input price, it can extract the whole increase in surplus coming from its input. Therefore, $\mathcal{I}$ and $\mathcal{S}$ prefer not to sign an exclusive contract: an input price to $\mathcal{E}$ is greater than losses of $\mathcal{I}$ in profits.[19]

However, a knowledge transfer to $\mathcal{E}$ might fail for two reasons. First, since sharing of secrets in non-contractible, $\mathcal{I}$ cannot be directly compensated for it. So, if (1.3.6) is not satisfied (competition is too high), costs of a knowledge leakage will outweigh benefits of higher input quality, and $\mathcal{I}$ will not share knowledge with $\mathcal{S}$. In this case, $\mathcal{I}$ and $\mathcal{S}$ sign a confidentiality agreement. Second, if (1.3.6) is satisfied, $\mathcal{I}$ and $\mathcal{S}$ might still prefer a confidentiality agreement because $\mathcal{S}$ does not capture the whole increase in surplus associated with a knowledge transfer to $\mathcal{E}$. Specifically, $\mathcal{I}$'s secrets increase $\mathcal{E}$'s profit by $\frac{1-\alpha}{2}(\lambda - \lambda_0)$, but $\mathcal{S}$ can capture only $\gamma \cdot \frac{1-\alpha}{2}(\lambda - \lambda_0)$ through its input price. Therefore, if γ is low, then sharing of $\mathcal{I}$'s secrets with $\mathcal{E}$ generates more losses in $\mathcal{I}$'s profits relative to gains in input price $t_{\mathcal{E}}$. So, $\mathcal{I}$ and $\mathcal{S}$ prefer a confidentiality agreement to create a commitment for $\mathcal{S}$ not to share $\mathcal{I}$'s secrets with $\mathcal{E}$.

[19] If $\mathcal{E}$ made an offer to $\mathcal{S}$, then $\mathcal{I}$ and $\mathcal{S}$ would always prefer exclusive dealing. In general, suppose that $\mathcal{S}$ and $\mathcal{E}$ have a generalized Nash Bargaining where $\psi_{\mathcal{S}} \in [0,1]$ is $\mathcal{S}$'s bargaining power. Then, $\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing with a confidentiality agreement whenever $\alpha > \frac{\psi_{\mathcal{S}}}{\psi_{\mathcal{S}}+2}$. For more, see Section 1.3.2.

1.3.2 Extensions and Additional Results

This section provides extensions and additional results to the baseline model. First, I generalize the competition, production and bargaining structures. Second, I consider a possibility of writing contracts with stipulated damages (Aghion and Bolton (1987)). Finally, I consider a setup with two suppliers and a common customer.

Generalizations in the Competition, Production, and Bargaining Structures

In this section, I consider an extension of the simple model in the competition, production and bargaining structures. Denote by $h_{\mathcal{I}} = 1$ if $\mathcal{I}$ uses a high-quality input from $\mathcal{S}$ (and shares its secrets with $\mathcal{S}$) and $h_{\mathcal{I}} = 0$ otherwise, $h_{\mathcal{E}} = 1$ if $\mathcal{E}$ uses a high-quality input from $\mathcal{S}$ and $h_{\mathcal{E}} = 0$ otherwise, $\ell = 1$ if $\mathcal{S}$ leaks $\mathcal{I}$'s secrets to $\mathcal{E}$ and $\ell = 0$ otherwise. I extend the competition and production structures by considering a more general mapping from $e = (h_{\mathcal{I}}, h_{\mathcal{E}}, \ell)$ to downstream profits, $\pi_{\mathcal{I}}(e)$ and $\pi_{\mathcal{E}}(e)$.[20] I also consider a generalized Nash bargaining between parties: denote by $\psi_{\mathcal{I}}$ and $\psi_{\mathcal{E}}$ the bargaining powers of $\mathcal{S}$ in negotiations with $\mathcal{I}$ and $\mathcal{E}$, respectively.

In Section 1.3.1, the supplier always has incentives to share $\mathcal{I}$'s knowledge with $\mathcal{E}$ in the third stage. In the general framework, $\mathcal{S}$ has incentives to shares $\mathcal{I}$'s knowledge with $\mathcal{E}$ if its input is complementary to $\mathcal{I}$'s knowledge in $\mathcal{E}$'s profits

$$\Delta\pi_{\mathcal{E}}(\cdot, \ell = 1) > \Delta\pi_{\mathcal{E}}(\cdot, \ell = 0) \tag{1.3.7}$$

[20] There are minimal assumptions on profits: $\pi_{\mathcal{I}}(e)$ ($\pi_{\mathcal{E}}(e)$) is increasing (decreasing) in $h_{\mathcal{I}}$ and decreasing (increasing) in $(h_{\mathcal{E}}, \ell)$. That is, firm $\mathcal{I}$ values a high-quality input from $\mathcal{S}$, and its profit is lower if firm $\mathcal{E}$ uses $\mathcal{I}$'s knowledge and a high-quality input from $\mathcal{S}$. Firm $\mathcal{E}$ values a high-quality input from $\mathcal{S}$ and the knowledge from $\mathcal{I}$, but its profit is lower if $\mathcal{I}$ uses a high-quality input. In Section 1.3.1, profits are the following

$$\pi_{\mathcal{I}}(h_{\mathcal{I}}, h_{\mathcal{E}}, \ell) = \frac{1-\alpha}{2}\Big(1 - \gamma \cdot (1 - h_{\mathcal{I}})\Big)\lambda + \alpha\Big[\lambda - (1 - (1 - h_{\mathcal{E}})\gamma) \cdot (\ell \cdot \lambda + (1 - \ell)\lambda_0)\Big]$$
$$\pi_{\mathcal{E}}(h_{\mathcal{I}}, h_{\mathcal{E}}, \ell) = \frac{1-\alpha}{2}\Big(1 - \gamma \cdot (1 - h_{\mathcal{E}})\Big) \cdot (\ell \cdot \lambda + (1 - \ell)\lambda_0)$$

where $\Delta\pi_{\mathcal{E}}(\cdot,\ell) = \pi_{\mathcal{E}}(h_{\mathcal{I}}=1,h_{\mathcal{E}}=1,\ell) - \pi_{\mathcal{E}}(h_{\mathcal{I}}=1,h_{\mathcal{E}}=0,\ell)$. In this case, $\mathcal{S}$ can bargain a higher input price from $\mathcal{E}$ ($t_{\mathcal{E}}(\ell) = \psi_{\mathcal{E}} \cdot \Delta\pi_{\mathcal{E}}(\cdot,\ell)$) under knowledge sharing ($\ell = 1$). Assume that (1.3.7) is satisfied.

$\mathcal{I}$'s incentives to share its secrets with $\mathcal{S}$ in the second stage depend on its expectations on whether $\mathcal{E}$ will have access to a high-quality input ($h_{\mathcal{E}} \in \{0,1\}$) and whether $\mathcal{S}$ will leak $\mathcal{I}$'s secrets to $\mathcal{E}$ ($\ell \in \{0,1\}$). $\mathcal{I}$ shares its secrets with $\mathcal{S}$ if

$$\underbrace{\pi_{\mathcal{I}}(h_{\mathcal{I}}=0,h_{\mathcal{E}},\ell) + (1-\psi_{\mathcal{I}}) \cdot \Big(TS(h_{\mathcal{E}},\ell) - TS^{o}(h_{\mathcal{E}},\ell)\Big)}_{\mathcal{I}\text{'s profits under knowledge sharing with } \mathcal{S}} > \underbrace{\pi_{\mathcal{I}}(h_{\mathcal{I}}=0,h_{\mathcal{E}}=0,\ell=0)}_{\mathcal{I}\text{'s profits without knowledge sharing}} \tag{1.3.8}$$

where $TS(h_{\mathcal{E}},\ell) = \pi_{\mathcal{I}}(h_{\mathcal{I}}=1,h_{\mathcal{E}},\ell) + \psi_{\mathcal{E}} \cdot \Delta\pi_{\mathcal{E}}(h_{\mathcal{I}}=1,h_{\mathcal{E}},\ell)$ is a joint surplus of $\mathcal{I}$ and $\mathcal{S}$ when $\mathcal{I}$ uses a high-quality input from $\mathcal{S}$, and $TS^{o}(h_{\mathcal{E}},\ell) = \pi_{\mathcal{I}}(h_{\mathcal{I}}=0,h_{\mathcal{E}},\ell) + \psi_{\mathcal{E}} \cdot \Delta\pi^{o}_{\mathcal{E}}(h_{\mathcal{I}}=0,h_{\mathcal{E}},\ell)$ is a joint surplus of $\mathcal{I}$ and $\mathcal{S}$ when $\mathcal{I}$ does not use a high-quality input from $\mathcal{S}$.

A contract (ex, ca) between $\mathcal{I}$ and $\mathcal{S}$ in the first stage will define the equilibrium $(h^{\star}_{\mathcal{E}}, \ell^{\star})$. Exclusive dealing ($ex = 1$) restricts S's ability to sell its input to $\mathcal{E}$ ($h^{\star}_{\mathcal{E}} = 1 - ex$), and a confidentiality agreement ($ca = 1$) restricts S's ability to communicate with $\mathcal{E}$ ($\ell^{\star} = 1 - ca$). $\mathcal{I}$ and $\mathcal{S}$ choose a contract to maximize their joint surplus subject to condition (1.3.8) that $\mathcal{I}$ has incentives to share its secrets with $\mathcal{S}$ in Stage 2 ($h_{\mathcal{I}} = 1$):

$$(h^{\star}_{\mathcal{E}}, \ell^{\star}) = \arg\max_{(h_{\mathcal{E}},\ell)} \Big\{\pi_{\mathcal{I}}(h_{\mathcal{I}}=1,h_{\mathcal{E}},\ell) + \psi_{\mathcal{E}} \cdot \Delta\pi_{\mathcal{E}}(h_{\mathcal{I}}=1,h_{\mathcal{E}},\ell)\Big\} \text{ subject to (1.3.8)}$$

The predictions in the extended model are similar to the simple one. First, the stronger outside option (lower $\Delta\pi_{\mathcal{E}}(\cdot)$ holding $\pi_{\mathcal{I}}(\cdot)$ fixed) leads to the equilibrium without knowledge sharing. Second, higher competition between $\mathcal{I}$ and $\mathcal{E}$ leads to the equilibrium without knowledge sharing. The competition can be measured based on the responsiveness of $\pi_{\mathcal{I}}(h_{\mathcal{I}},h_{\mathcal{E}},\ell)$ to changes in $(h_{\mathcal{E}},\ell)$.

Richer Contracting: Stipulated Damages

In Section 1.3.1, $\mathcal{I}$ and $\mathcal{S}$ can sign exclusive dealing and a confidentiality agreement. These contracts prevent $\mathcal{S}$ from selling an input to $\mathcal{E}$ and sharing $\mathcal{I}$'s secrets with $\mathcal{E}$. In this section, I consider the possibility of stipulated damages (Aghion and Bolton (1987)) where $\mathcal{I}$ and $\mathcal{S}$ specify a fee F for breaking a contract. This might allow $\mathcal{I}$ and $\mathcal{E}$ to extract more surplus from $\mathcal{E}$.

Suppose that in the first stage $\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing with a confidentiality agreement and specify a fee F that $\mathcal{S}$ has to pay to $\mathcal{I}$ for breaking this contract. Assume $\mathcal{I}$ shares its knowledge with $\mathcal{S}$, and that after $\mathcal{E}$ enters, $\mathcal{S}$ can propose $\mathcal{E}$ to pay F in order to break the contract between $\mathcal{I}$ and $\mathcal{S}$. If $\mathcal{E}$ agrees to pay F, then S will share $\mathcal{I}$'s secrets with $\mathcal{E}$ and will make a take-it-or-leave-it offer about the input price $t_{\mathcal{E}} = \frac{1-\alpha}{2}\lambda - \frac{1-\alpha}{2}(1-\gamma)\lambda = \frac{1-\alpha}{2}\gamma\lambda$. Therefore, $\mathcal{E}$ agrees to pay F if

$$F < F^{\star} = \frac{1-\alpha}{2}\lambda - t_{\mathcal{E}} - \frac{1-\alpha}{2}(1-\gamma)\lambda_0 = \frac{1-\alpha}{2}(1-\gamma)\left(\lambda - \lambda_0\right)$$

Assume that in the case of indifference $\mathcal{E}$ agrees to pay F. Then, the joint surplus of $\mathcal{I}$ and $\mathcal{S}$ under the contract with stipulated damages is

$$\frac{1-\alpha}{2}\lambda + t_{\mathcal{E}} + F^{\star} = \frac{1-\alpha}{2}\left(\lambda + \gamma\lambda_0\right) + \frac{1-\alpha}{2}\left(\lambda - \lambda_0\right)$$

which maximizes the joint surplus of $\mathcal{I}$ and $\mathcal{S}$ for $\alpha < \frac{1}{3}$ (compare to (1.3.3) – (1.3.5)). To sum up, for $\alpha > \frac{1}{3}$ firms $\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing *without* stipulated damages (or set them $F > F^{\star}$). In this case, only $\mathcal{S}$ learns $\mathcal{I}$'s secrets. For $\alpha < \frac{1}{3}$, firms $\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing and a confidentiality agreement *with* stipulated damages $F = F^{\star}$. In this case, both $\mathcal{S}$ and $\mathcal{E}$ learn $\mathcal{I}$'s secrets.

The ability to use stipulated damages increases social welfare because it ensures a knowledge transfer from $\mathcal{I}$ to $\mathcal{E}$ in cases where previously $\mathcal{I}$ with $\mathcal{S}$ signed a confidentiality agreement (see Figure 1.5). This happens because $\mathcal{I}$ is compensated for the knowledge

transfer to $\mathcal{E}$ through stipulated damages while before contractual limitations prevented this compensation.

Two Suppliers and a Common Customer

Suppose an incumbent input supplier ($\mathcal{I}$) and a buyer ($\mathcal{B}$) expect the entry of a new supplier ($\mathcal{E}$). To produce a final product, $\mathcal{B}$ needs three inputs to perform a continuum of measure 1 of tasks. A share $\frac{1-\alpha}{2}$ of tasks requires one unit of an input from $\mathcal{I}$, and a share $\frac{1-\alpha}{2}$ requires an input from $\mathcal{E}$. A share $\alpha \in (0,1)$ of tasks requires an additional third input from either $\mathcal{I}$ or $\mathcal{E}$. $\mathcal{B}$'s gross profits are the following

$$\Pi_{\mathcal{B}} = U_{\mathcal{I}} \cdot \frac{1-\alpha}{2} q_{\mathcal{I}} + U_{\mathcal{E}} \cdot \frac{1-\alpha}{2} q_{\mathcal{E}} + U \cdot \alpha \tilde{q}$$

where $q_{\mathcal{I}}$ and $q_{\mathcal{E}}$ are input qualities of $\mathcal{I}$ and $\mathcal{E}$, respectively; $U_{\mathcal{I}} = 1$ ($U_{\mathcal{E}} = 1$) if $\mathcal{B}$ buys a specialized input for $\frac{1-\alpha}{2}$ tasks from $\mathcal{I}$ ($\mathcal{E}$) and $U_{\mathcal{I}} = 0$ ($U_{\mathcal{E}} = 0$) otherwise; $U = 1$ and $\tilde{q} = q_{\mathcal{I}}$ ($\tilde{q} = q_{\mathcal{E}}$) if $\mathcal{B}$ buys a third input for α tasks from $\mathcal{I}$ ($\mathcal{E}$) and $U = 0$ otherwise.

$\mathcal{I}$ has private knowledge about a valuable technology. If $\mathcal{B}$ learns $\mathcal{I}$'s secrets, it can better accommodate $\mathcal{I}$'s input in its production and the input's quality to $\mathcal{B}$ is $q_{\mathcal{I}} = 1$. Otherwise, $\mathcal{I}$'s input quality is $q_{\mathcal{I}} = 1 - \gamma$ where $\gamma \in [0,1]$. There are also two ways how $\mathcal{B}$ can use $\mathcal{I}$'s secrets to improve $\mathcal{E}$'s input in the final production. First, without sharing $\mathcal{I}$'s secrets with $\mathcal{E}$ it can use them to better accommodate $\mathcal{E}$'s input in the production and improve it from $q_{\mathcal{E}} = (1-\gamma)\lambda_0$ to $q_{\mathcal{E}} = \lambda_1$ where $\lambda_0 \leqslant \lambda_1 \leqslant 1$. Second, $\mathcal{B}$ can also share $\mathcal{I}$'s secrets with $\mathcal{E}$ to get the input quality $q_{\mathcal{E}} = 1$.

The timing and contracting are the same as in Section 1.3.1. First, $\mathcal{I}$ and $\mathcal{B}$ have a Nash Bargaining with 50/50 split of the surplus over exclusive dealing ($\mathcal{B}$ only buys an input from $\mathcal{I}$) and a confidentiality agreement ($\mathcal{B}$ does not communicate with $\mathcal{E}$ about $\mathcal{I}$'s secrets). Second, $\mathcal{I}$ decides whether to share its secrets with $\mathcal{B}$, and then they have a Nash Bargaining with 50/50 split of the surplus over the input price. Third, $\mathcal{B}$ decides whether to share $\mathcal{I}$'s secrets with $\mathcal{E}$ (if it is not restricted by a confidentiality agreement), and

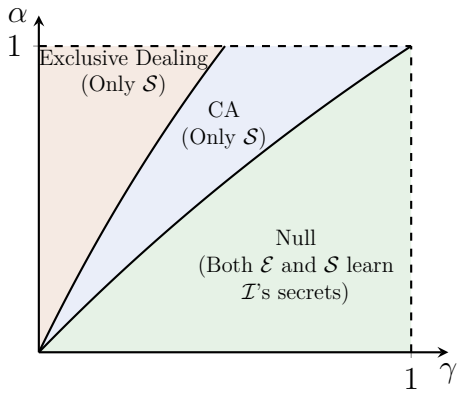


Figure 1.6: Equilibrium in the Model with Two Suppliers and a Common Customer

This figures illustrates the equilibrium contract between $\mathcal{I}$ and $\mathcal{B}$ described in Proposition 10 in Appendix A.4.2. Parameter α measures the degree of competition between $\mathcal{I}$ and $\mathcal{E}$, and parameter γ reflects the value of $\mathcal{I}$'s input in $\mathcal{B}$'s production.

makes a take-it-or-leave-it offer about the input price (if it is not restricted by exclusive dealing).

The equilibrium is summarized in Proposition 10 (Appendix A.4.2) and in Figure 1.6. Some predictions are similar to the model with a common supplier. In particular, lower competition (α) and greater importance of knowledge sharing (γ) lead to the equilibrium where all firms learn $\mathcal{I}$'s secrets. However, there are differences between common suppliers and customers in incentives to sign exclusive dealing and confidentiality agreements. The common supplier might want to maximize the difference between product qualities of its customers because they compete with each other. So, it might prefer to sign exclusive dealing to create a commitment that it will not sell inputs to $\mathcal{E}$ and will not share $\mathcal{I}$'s secrets with $\mathcal{E}$. On the other hand, profits a common customer are increasing in the quality of inputs from all its suppliers. So, the common customer ($\mathcal{B}$) does not have incentives to create such a commitment, and it signs exclusive dealing or a confidentiality agreement only to encourage a supplier to share private knowledge with it.

1.3.3 Testable Predictions

The theory predicts patterns of knowledge flows in a production network that I map to the data using patent citations. In Section 1.4, I test four predictions from the model:

1. Suppliers and customers are more likely to cite each other relative to similar firm-pairs with no vertical relationship.

2. Firms sharing a common supplier or customer are more likely to cite each other relative to firms with exclusive suppliers or customers.

3. A stronger bargaining position (γ) of the common supplier (a worse outside option of $\mathcal{E}$) is associated with a higher likelihood of a citation between its customers ($\mathcal{I}$ and $\mathcal{E}$).

4. A higher level of competition (α) leads to less knowledge sharing between $\mathcal{I}$ and $\mathcal{E}$.

Remark 1. The model with two customers and a common supplier predicts that $\mathcal{E}$ is more likely to cite $\mathcal{I}$ if it has a bad outside option to $\mathcal{S}$'s input (high γ). In the framework with two suppliers and a common customer in Section 1.3.2, I assume that suppliers' outside options of not trading with a customer are zero. Therefore, it does not give a prediction about citations between suppliers as a function of the bargaining position of a common customer. I leave the framework with a more general production network for future research.

Remark 2. In the model, each firm has one patent, and sharing of secrets affects citation probabilities between them. The framework can be extended to the case in which firms have multiple patents. Suppose that firm $\mathcal{S}$ has $N_{\mathcal{S}}$ patents, and each independently makes a citation to $\mathcal{I}$'s patent with probability $p_{\mathcal{S}}(1)$ if $\mathcal{S}$ receives $\mathcal{I}$'s secrets and $p_{\mathcal{S}}(0)$ otherwise. The expected number of citations to $\mathcal{I}$'s patent from firm $\mathcal{S}$ is $p_{\mathcal{S}}(k) \cdot N_{\mathcal{S}}$ for $k \in \{0, 1\}$. The empirical strategy in Section 1.4.1 estimates $p_{\mathcal{S}}(1) - p_{\mathcal{S}}(0)$ taking $N_{\mathcal{S}}$ as

given. The focus on citation probabilities is motivated by the evidence from Section 1.2 that the concentration of citations is primarily explained by differences across firms in citation rates rather than patenting. To explain the full distribution of citations across firms one should explain not only citation probabilities ($p_{\mathcal{S}}(k)$) but also the number of patents ($N_{\mathcal{S}}$). I leave the full analysis of $(p_{\mathcal{S}}(k), N_{\mathcal{S}}(k))$ for future research.

1.4 Empirical Tests

This section tests four theoretical predictions from Section 1.3.3. Section 1.4.1 tests predictions 1–3 using cross-sectional and time variation in supplier-customer relationships between companies. Specifically, the unit of analysis is a pair of patents, and I estimate the probability of a citation between a randomly selected pair of patents as a function of the relationship between their assignees.

Section 1.4.2 tests the fourth prediction about the effect of competition on knowledge sharing. There are at least two ways to test this prediction. The first one is a test of the equilibrium structure: compare cross-sectional variation across industries in the competition level with citation patterns between firms. There is a potential endogenity problem with this approach. For example, firms in industries with less competition might also rely less on knowledge from each other for technological reasons. The second way is a test of the equilibrium response: find within-industry shocks to the competition levels and compare *changes* in citation patterns in the model and data. I use the rise of imports from China and the instrumental variable approach from Autor *et al.* (2013) to isolate plausibly exogenous industry shocks to the competition levels. Specifically, the unit of analysis is a technological class, and I show that technologies more exposed to the rising import competition from China experienced a faster growth in the concentration of citations.

Data

I use the data on utility patents granted in the U.S. Patent and Trademark Office (USPTO), Compustat North America data on publicly traded firms, and Compustat Segments data on the supplier-customer relationship between firms. The data are taken for the period 1976–2019. Compustat data give information on sales, profits, employment and industry affiliation of publicly traded firms. Compustat Segments contains data on the supplier-customer links reported by publicly traded U.S. firms. Regulation SFAS No. 131 requires publicly traded firms to report the identity of any customer representing more than 10% of their total sales. Therefore, the data are constructed using the information provided by suppliers about their customers. All details are given in appendix A.2.1.

These data have several limitations. First, the supplier-customer data cover only publicly traded firms listed in the U.S. Private and foreign firms might differ from the U.S. publicly traded firms in terms of patenting and citation practices, and this might bias the results relative to the estimation based on the full network of firms' relationships. The differences in rates of patenting should not confound the results because the empirical strategy in Section 1.4.1 is designed to estimate the probability of a citation between two patents from different firms conditional on the existence of these patents. However, the differences in citation practices might introduce a bias if these differences are not accounted by geographic and firm controls. For example, despite filing patents in the U.S. office foreign firms might still follow citation standards of their countries. The second limitation is a selection problem because firms are required to list only the largest customers. As a result, the data might miss small customers and large suppliers selling to multiple firms who do not represent more than 10% of the supplier's revenue. I use controls for firms' sizes to partially alleviate this problem. Third, suppliers are not required to report sales to their customers.[21] Data on bilateral sales would be useful to construct bargaining

[21] Computstat Segments data include a variable for the annual sales of the reporting supplier to the

positions between firms in a vertical relationship. Instead, I use total sales of a supplier relative to other firms in the same industry to proxy for its bargaining position vis-a-vis its customers.

Despite all these limitations, Compustat Segments data set is one of the best sources for supplier-customer relationships in the U.S. over a long period of time. These data were used in studies on propagation of shocks in production networks (Barrot and Sauvagnat (2016)), network formation (Lim (2017)), and tests of investors' inattention (Cohen and Frazzini (2008)).

1.4.1 Citations in a Production Network

In this section, I test predictions 1–3 from Section 1.3.3. The first prediction is that firms are more likely to cite each other if they are connected by a vertical production relationship. Second, firms are more likely to cite others' patents if they share a common supplier or customer. Finally, the probability of a citation between firms sharing a common supplier is increasing in the bargaining position of this supplier. The strength of the supplier's bargaining position is defined by the the customer's outside options of not buying an input from the supplier. I proxy the bargaining position of common suppliers by their sales share in the industry they are affiliated with. The intuition is that the supplier's relative size approximates the difficulty of finding an alternative supplier for the customers.

Based on the econometric techniques from the spatial literature (Jaffe *et al.* (1993), Thompson and Fox-Kean (2005), Singh and Marx (2013)), I estimate the following specification

$$\begin{aligned}\mathbb{P}\left(y_{ij}=1|\mathbb{I}_{ij},X_{ij}\right)=F\big(&\beta_{sc}\cdot\mathbb{I}_{ij}^{sc}+\beta_{cs}\cdot\mathbb{I}_{ij}^{cs}+\beta_{com}(s)\cdot\mathbb{I}_{ij}^{com}(s)\cdot(1+\beta_{s}\cdot S_{ij})+\\&+\beta_{com}(c)\cdot\mathbb{I}_{ij}^{com}(c)+X_{ij}\cdot\gamma\big)\end{aligned}\tag{1.4.1}$$

reporting customer, but this variable is often missing or imputed.

where $y_{ij} = 1$ if patent j cites patent i and zero otherwise. For a pair of patents (i, j) denote the application year of patent j by t_j. The variable $\mathbb{I}_{ij}^{sc}$ ($\mathbb{I}_{ij}^{cs}$) a dummy equal to 1 if j's assignee was a supplier (customer) to i's firm in the period $[t_j - 2, t_j + 2]$. The variable $\mathbb{I}_{ij}^{com}(s)$ ($\mathbb{I}_{ij}^{com}(c)$) is a dummy equal to 1 if j's and i's firms shared at least one common supplier (customer) in the period $[t_j - 2, t_j + 2]$. Finally, the variable S_{ij} reflects the size of common suppliers. Specifically, for each common supplier I compute the average sales share in its industry over the period $[t_j - 2, t_j]$. If firms shared several suppliers, I take an average across their sales shares. The variable X_{ij} includes firm, spatial, time and technological controls.[22] I use logit, probit and linear probability models. In the linear model, I include fixed effects for the interaction between (i, j) technological classes and i's grant years ($Tech_i \times Tech_j \times Year_i$) where the technology class is defined at the group level in Cooperative Patent Classification (CPC) system. This allows to control for differences in citation patterns across technologies and in time. I also include fixed effects for cited and citing firms which allow to control for time-invariant differences in citation practices across firms.

The estimation of the specification (1.4.1) is not computationally feasible for the universe of all patent-pairs. Since the data on the supplier-customer links cover only publicly traded firms, I restrict the sample to citations between these firms only (excluding self-citations), but the set of all patent-pairs is still too large. I use the sampling procedure first proposed in Jaffe *et al.* (1993). Specifically, for each cited-citing pair of patents I randomly select a control patent from the set of patents that do not make a citation to the cited patent but have the same application year, the same detailed technological class

[22]Specifically, firm controls include log sales of i's assginee and log number of patents granted to i's assignee in the grant year of patent i, log sales of j's assignee and log number of patents applied by j's assignee in application year t_j. Spatial controls include log distance between inventors and a dummy whether they are in the same country. Time controls include i's grant year fixed effects, and dummies for a lag between j's application year and i's grant year. Technology controls include a dummy whether i and j patents are in the same main subgroup of CPC, and whether their assignees are in the same four-digit Standard Industry Classification (SIC) industry.

(main subgroup CPC) as the citing patent, and belong to a firm different from the citing one.[23] I use the sampling weighting (Manski and Lerman (1977)) to ensure that estimates reflect probabilities in the sample of cited-citing and all cited-control patent pairs, not only selected ones. Appendix A.2.5 provides details on the empirical specification and data construction.

The estimation results are given in Table 1.3. All specifications show that the probability of a citation between two patents is higher if patents' assignees have a supplier-customer relationship and/or share a common supplier or customer. Moreover, the probability of a citation between firms sharing common suppliers is increasing in the relative size of these suppliers. Suppliers, on average, are more likely to cite customers than customers suppliers, and sharing a customer is associated with a greater increase in the probability of a citation relative to sharing a supplier. Columns (1)–(4) show the results for the linear probability model. I multiply the depended variable by 100, so the estimates represent an increase in the probability of a citation in percentage points. In general, the probability of a citation between two randomly selected patents is very small, and this translates into small estimates of the linear probability model. However, they are economically large. A randomly selected pair of cited-citing patents has, on average, around 483 cited-control pairs without a citation meaning that the probability of a citation is around 0.21%. So, being a supplier increases the probability of a citation by 0.08%, which represents a $8/21 = 38\%$ proportional increase. The Logit model provides another way to evaluate the magnitudes of the estimates. Since the probability of a citation is low, the Logit coefficients

[23] Jaffe *et al.* (1993) consider control patents from the same 3-digit USPC technological class. However, Thompson and Fox-Kean (2005) argue that the 3-digit classes are too broad to capture technological similarity between patents and propose to use the 9-digit classification. The problem with the 9-digit classification is that it is too narrow and potentially introduces a lot of noise leaving the majority of patents without a control one (Henderson *et al.* (2005)). Unfortunately, the USPC system does not provide intermediate levels of aggregation between 3-digit and 9-digit classes. To address both critiques, I use instead the cooperative patent classification (CPC), which has more flexible hierarchical classification system. Specifically, I find control patents from the same main subgroup level, which has more than 7000 categories.

Table 1.3: Supply Chain Relationship and the Probability of a Citation

Empirical Specification	Linear				Logit	Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Supplier Cites a Customer, β_{sc}	0.09***	0.09***	0.08***	0.08***	0.34***	0.11***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.012)	(0.004)
Customer Cites a Supplier, β_{cs}	0.05***	0.05***	0.05***	0.07***	0.27***	0.09***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.015)	(0.005)
Share Common Customers, $\beta_{com}(c)$	0.10***	0.10***	0.09***	0.07***	0.39***	0.12***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.009)	(0.003)
Share Common Suppliers, $\beta_{com}(s)$	0.02***	0.01***	0.02***	0.03***	0.04***	0.01***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.001)
Relative Size of Common Suppliers, β_s	0.54***	0.59***	0.51***	0.47***	2.28***	0.76***
	(0.024)	(0.024)	(0.024)	(0.025)	(0.061)	(0.022)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Patent Controls		Yes	Yes	Yes	Yes	Yes
$\text{Tech}_i \times \text{Tech}_j \times \text{Year}_i$ FE			Yes	Yes		
Firm_i, Firm_j FE				Yes		

Notes: This table shows the estimation results for the specification in (1.4.1). It estimates the probability of a citation between two patents as a function of the production relationship between their assignees. The sample is restricted to patents granted between 1976 and 2019 to publicly traded firms only and excludes self-citations. I use the sampling procedure outlined in Jaffe *et al.* (1993) and Thompson and Fox-Kean (2005) to construct the sample of cited-citing and cited-control pairs of patents. For the linear probability model, the outcome variable is multiplied by 100. In the regressions, firm controls include log sales of i's assignee and log number of patents granted to i's assignee in the grant year of patent i, log sales of j's assignee and log number of patents applied by j's assignee in application year t_j, and a dummy whether i and j assignees are in the same four-digit Standard Industry Classification (SIC) industry. Patent controls include log distance between inventors and a dummy whether they are in the same country, i's grant year fixed effects, and dummies for a lag between j's application year and i's grant year, a dummy whether i and j patents are in the same main subgroup of CPC. Standard errors are clustered at the cited patent level. p-value $^{***}p \leqslant 0.01$. All specifications are sample-weighted (Manski and Lerman (1977)). More details are given in Appendix A.2.5.

approximate the proportional increase in the probability of a citation. Specifically, the marginal effect of k's variable (dummy) is

$$\Lambda(X\gamma + \beta_k) - \Lambda(X\gamma) \approx \beta_k \cdot \Lambda'(X\gamma) = \beta_k \cdot \Lambda\left(X\gamma\right) \cdot \left(1 - \Lambda\left(X\gamma\right)\right) \approx \beta_k \cdot \Lambda\left(X\gamma\right)$$

where $\Lambda(X\gamma) = \frac{1}{1+\exp(-X\gamma)}$ and $\Lambda\left(X\gamma\right) \approx 0$. So, being supplier to a cited firm is associated with a 34% proportional increase in a probability of a citation between two patents, which is comparable to 38% in the linear probability model.

1.4.2 Trade with China and Concentration of Knowledge Flows

The model in Section 1.3 predicts that a higher level of competition leads to lower knowledge sharing. I provide the evidence consistent with this prediction using the variation across technologies in their exposure to the rising import competition from China. To isolate changes in the import competition level unrelated to US demand and technological shocks, I instrument US imports from China by Chinese exports to other high-income countries (Autor *et al.* (2013)). Autor *et al.* (2020) used this methodology to show that imports from China led to the decline in corporate patenting in the manufacturing sector. I complement their evidence by studying how imports from China affected knowledge sharing between firms.

Following Autor *et al.* (2020), I define the measure of the trade exposure at the four-digit Standard Industry Classification (SIC) over the two subperiods, 1991 to 1999 and 1999 to 2007,

$$\Delta IP_{i1} = \frac{M_{i,1999} - M_{i,1991}}{Y_{i,91} + M_{i,91} - E_{i,91}} \text{ and } \Delta IP_{i2} = \frac{M_{i,2007} - M_{i,1999}}{Y_{i,91} + M_{i,91} - E_{i,91}} \tag{1.4.2}$$

where $M_{i,t}$ is the U.S. imports from China for industry i and year $t \in \{1991, 1999, 2007\}$, and $Y_{i,91} + M_{i,91} - E_{i,91}$ is the absorption at the start of the period (industry shipments plus imports minus exports). For each patent, I calculate the import penetration for its technology class using the mapping of four-digit SIC industries to technology classes implied from patents owned by publicly traded firms as in Autor *et al.* (2020).

Figure 1.7 shows the concentration measure from Section 1.2 (Figure 1.1) for different technological classes based on their exposure to import competition from China, ΔIP_{i2}. All classes experienced an increase in the concentration of citations after 2000. However, the dynamics is different based on the exposure to import competition. For classes with the least exposure (below first quartile Q_{25}), the concentration of citations was initially high (around 55%) and stable prior to 2000, and then it increased up to 75%. For classes

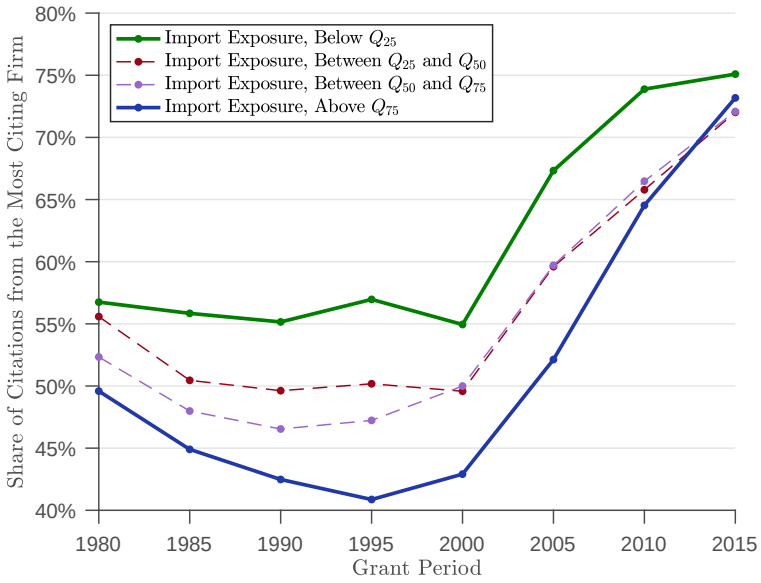


Figure 1.7: The concentration for technological classes with different exposure to import competition.

This figure shows the concentration measure from Section 1.2 (Figure 1.1) for different technological classes divided into quartiles based on their exposure to import competition from China, ΔIP_{i2} in (1.4.2). The concentration measure for year t shows the average concentration in years $[t-4, t]$.

with the most exposure (above the third quartile Q_{75}), the concentration was initially lower (around 50%), decreased to around 40% near 2000, and then it also increased up to 75%. So, technological classes with the most exposure to China shock experienced faster growth in the concentration.

I follow the methodology from Section 1.2 to separate whether the rise in the concentration in different classes is driven by increasing variance across firms in patenting (N) or in citation rates (p). In Figure A.5 (Appendix A.3), I compare the dynamics after 2000 of the actual concentration measure relative to the counterfactual one from the Monte-Carlo simulations. The counterafactual concentration shows how the concentration would evolve due to the changes in patenting across firms (N) holding the firms' citation rates (p) fixed at the level of 2000. For technologies least exposed to the competition from China, the actual and the counterfactual concentration measures closely follow each other,

and the changes in citation rates across firms (p) explain about 41% of the rise in the concentration. In contrast, for the most exposed technologies the changes in citation rates across firms (p) explain around 70% of the increase in the concentration. Therefore, trade with China affected citation patterns primarily through the changes in citation rates. An implication of this result is that firms' exit from patenting cannot fully explain the rise in the concentration of citations.

I study the change in the concentration of citations around 2000 in the regression specifications. Specifically, I define for each technology class the average concentration measure among patents granted in the 7-year period starting from 1977, 1984, 1991, 1998, 2005.[24] Appendix A.2.6 provides more details. I denote the concentration measure for technology class j and the period starting from t by $\mathcal{C}_{j,t}$, and define the following growth measures

$$\Delta y_{j1} = 100 \cdot \ln\left(\mathcal{C}_{j,1998}/\mathcal{C}_{j,1991}\right) \text{ and } \Delta y_{j2} = 100 \cdot \ln\left(\mathcal{C}_{j,2005}/\mathcal{C}_{j,1998}\right)$$

Figure 1.8 shows that technological classes more exposed to trade with China ($\Delta IP_{j\tau}$) experienced a higher growth in the concentration of citations ($y_{j\tau}$).

I estimate the following specification

$$\Delta y_{j\tau} = \beta \Delta IP_{j\tau} + \gamma X_{j0} + \varepsilon_{j\tau} \tag{1.4.3}$$

where $\tau \in \{1, 2\}$ and X_{j0} is the set of controls. To control for the aggregate trend in the concentration of citations I include time fixed effects. Since the concentration measure

[24]Periods are 1977–1983, 1984–1990, 1991–1997, 1998–2004, 2005–2011. The sample construction for patents is different from Autor *et al.* (2020) due to the nature of the outcome variable. Autor *et al.* (2020) consider patents applied in the years 1975, 1983, 1991, 1999, 2007. Their main outcome variable is the number of patents while in this paper it is the concentration of citations. The concentration measure has a meaningful interpretation only for highly-cited patents. For example, the patent with one citation always has concentration equal to one. Therefore, I focus on the set of the top 5% of the most cited patents. Since each technological class has a small number of highly-cited patents in each year, I take the average of the concentration measure over a 7-year period. I also do the analysis at the technology class level rather than at the firm level because most firms have a small number of highly-cited patents.

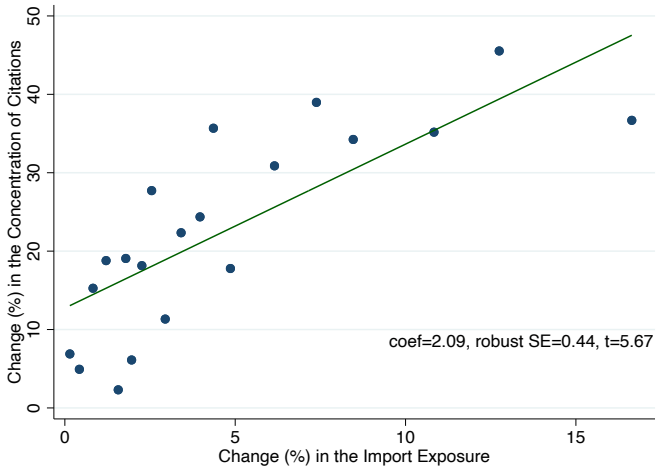


Figure 1.8: Changes in the concentration and in the import penetration from China.

This figure shows the binned scatter plot of the change in the concentration of citations and the change in the import exposure from China. The specification is weighted by the number of Compustat-matched U.S.-inventor patents in a technology class.

depends on the total number of citations, I also include the change in the average number of citations for each technological class. Moreover, I include two lags of the outcome variable to control for technology-specific trends prior to China shock. I also include fixed effects for 11 manufacturing sectors, and for 6 main NBER technological categories. Finally, I control for the rising importance of software inventions (Chattergoon and Kerr (2021)). Specifically, for each technological class I include a dummy whether it has more than 50% of software subclasses (Graham and Vishnubhakat (2013)).

Panel A in Table 1.4 shows the results of simple OLS regressions, and Panel B shows the results for the specification in which changes in US import exposure ($\Delta IP_{j\tau}$) are instrumented by changes in Chinese exports to non-U.S. high-income countries. In all specifications, there is a positive relationship between the changes in the import competition from China and the growth in the concentration of citations. Table A.3 in Appendix A.3 shows the results from two additional placebo exercises. First, I show that the relationship between China shock and the rise in the concentration is insignificant for

Table 1.4: Trade with China and Increase in the Concentration of Citations

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: OLS			
Δ Tech Class Exposure	2.09***	1.46***	1.44***	1.76***	1.43***	1.43***
to Chinese Imports	(0.44)	(0.41)	(0.42)	(0.34)	(0.39)	(0.38)
			Panel B: 2SLS			
Δ Tech Class Exposure	2.38***	1.52***	1.50***	1.92***	1.84***	1.81***
to Chinese Imports	(0.49)	(0.50)	(0.50)	(0.42)	(0.56)	(0.58)
Time FE		Yes	Yes	Yes	Yes	Yes
Δ Citations			Yes	Yes	Yes	Yes
2 Lags of outcomes				Yes	Yes	Yes
11 sectors, 6 Tech					Yes	Yes
Software Patents						Yes

Notes: This table shows the estimated coefficient β for the specification in (1.4.3). Panel A shows the results for simple OLS regressions. Panel B shows the results for the specification in which the import penetration from China is instrumented with Chinese exports to non-U.S. high-income markets (Autor *et al.* (2020)). Regressions consider the effect of higher growth in import penetration from China on the increase in the concentration of citations at the technology class level. Industry exposure to Chinese competition is mapped to technology class exposure using the mapping implied by the U.S. publicly traded firms in Compustat as in Autor *et al.* (2020). Controls include time fixed effects, a change in the average number of citations for a technology class, 2 lags of the outcome variable, fixed effects for 11 manufacturing sectors and for 6 main NBER technology categories, and a dummy for software technological classes. I define the software classes as classes where more than 50% of software subclaclasses according to Graham and Vishnubhakat (2013). All specifications are weighted by the number of Compustat-matched U.S.-inventor patents in a technology class. Standard errors are clustered at the technology class level. p-values: $^{***}p \leqslant 0.01$.

non-corporate patents (e.g., universities and government agencies). Therefore, the effect of trade competition with China is specific to the corporate sector, and the results are unlikely to be driven by the correlation between general technological changes and globalization. Second, I regress lag outcome variables (pre 1991) on future changes in imports from China. The coefficients are insignificant, so the main results are unlikely to be driven by contemporaneous changes in the technological opportunities and trade.

1.5 Policy Implications

In this section, I discuss policy implications of the evidence and theory described in the previous sections. The literature on economic growth and innovation is usually focused on incentives to invest in R&D. The most common recommendation is that R&D should be subsidized due to the presence of knowledge externalities (see, for example, Bloom *et al.* (2019)). In the analysis below, I abstract from incentives to invest in R&D and focus on firms' incentives to share knowledge with each other. I study the conditions under which the restrictions on the use of confidentiality and exclusive contracts can facilitate knowledge sharing and increase welfare.[25]

The rationale for the policy intervention is based on the assumption that private knowledge of a firm is valuable to others. In Section 1.5.1, I provide suggestive estimates on the value of access to private knowledge of a firm. Section 1.5.2 derives the socially optimal regulation of confidentiality and exclusive contracts.

1.5.1 Estimation of the Value of Private Knowledge

In this section, I use the interpretation from Section 1.2 that the concentration of citations to a patent is caused by the access of the most citing firm to private knowledge of the patent's owner. If this access is valuable, it should be reflected in the quality of (citing) patents and growth of firms that have it. I estimate patents' quality and firms' growth as a function of the knowledge they rely on (citations they make).

I identify patents where firms are likely to have access to some private knowledge relative to others based on two conditions. First, a firm should be responsible for the majority of citations to a patent from the top 5% of the most cited patents. Second,

[25]The analysis is stylized and highlights the relationship between the contractual arrangements and knowledge sharing. It is possible to incorporate investments in R&D, but the predictions about the incentives to invest in R&D will be similar to the growth literature: the ban on exclusive and/or confidentiality contracts reduces returns from R&D and leads to lower investments.

the share of citations coming from this firm should exceed the $95th$ quantile from the Monte-Carlo simulations on the randomization of citations (see Section 1.2). Based on the analysis from Section 1.2, these two conditions imply that the citing firm has a significantly higher citation rate (p) relative to other firms. I call patents satisfying these two conditions "specific" patents. For example, patents from Amkor Technology making citations to IBM's patent in Table 1.1 are "specific" ones. Appendix A.2.2 provides more details.

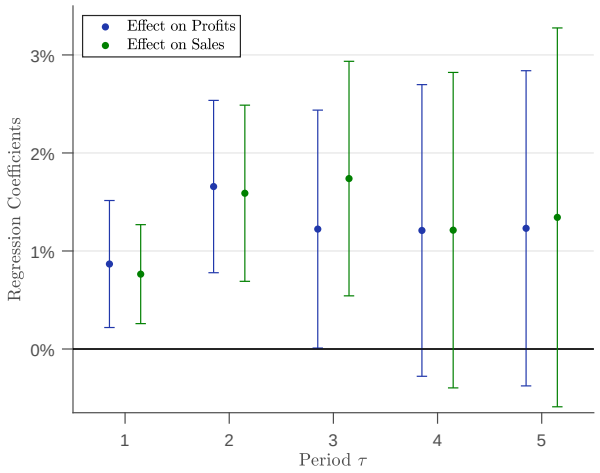
I start with the following regression of firms' growth on different types of patents.

$$y_{f,t+\tau} - y_{f,t} = \beta_s \cdot \tilde{\ln}\left(N^s_{ft}\right) + \beta_g \cdot \tilde{\ln}\left(N^g_{ft}\right) + \beta_n \cdot \tilde{\ln}\left(N_{ft}\right) + \eta_f + \zeta_{i(f),t} + \gamma X_{ft} + \varepsilon_{ft} \tag{1.5.1}$$

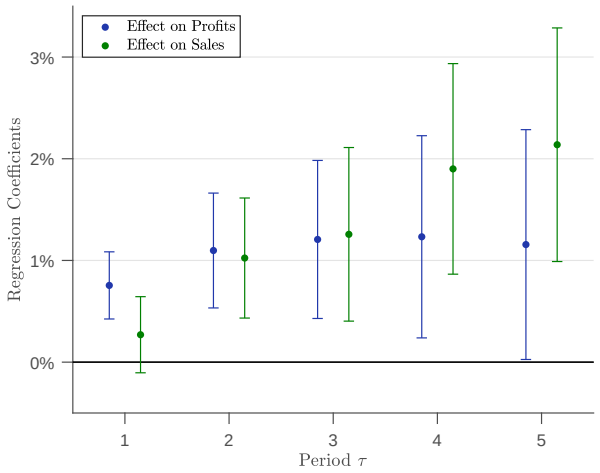
where $\tilde{\ln}(x) = \ln\left(x + (x^2+1)^{\frac{1}{2}}\right)$ is the inverse hyperbolic sine transformation to control for zeros in patenting. For firm f in year t, the variable $y_{f,t}$ denotes profits or sales, N^s_{ft} is the number of granted specific patents, N^g_{ft} is the number of granted patents that make at least one citation to top 5% of the most cited patents. The inclusion of N^g_{ft} allows to distinguish whether specific patents are associated with higher firm's growth because they are responsible for the majority of concentrated citations or because they simply rely on a highly-cited patent. N_{ft} is the total number of granted patents, η_f is a firm fixed effect, $\zeta_{i(f),t}$ is an industry-year fixed effect, and X_{ft} is the set of controls that includes firm employment, firm capital stock, and a lag of firm profits or sales. The variable ε_{ft} is an error term. Appendix A.2.7 provides details.

The estimated coefficients β_s and β_n for $\tau = 1, \ldots, 5$ are given in the figures 1.9a and 1.9b. The positive coefficient $\beta_n > 0$ means that higher number of patents is associated with higher firm's growth in sales and profits, which is a standard result in the literature. However, in addition to this result the positive coefficient $\beta_s > 0$ indicates that given the same number of patents the growth rate is higher for firms with greater share of specific patents. Therefore, specific patents are more valuable to firms relative to non-specific ones.

One of the differences between specific and non-specific patents might be related to



(a) Coefficients for specific patents, β_s.



(b) Coefficients for all patents, β_n.

Figure 1.9: Value of Private Knowledge

This figure shows the estimated coefficients β_s and β_n with corresponding 95% confidence intervals for $\tau = 1, \dots 5$ for the specification in (1.5.1). The positive coefficient $\beta_n > 0$ means that higher number of patents is associated with higher firm's growth in sales and profits. The positive coefficient $\beta_s > 0$ indicates that given the same number of patents the growth rate is higher for firms with greater share of specific patents. Blue lines show the results for the specification with growth in profits, and green lines for specification with growth in sales. The left panel shows estimates for β_s, and the right panel shows the estimates for β_n. Standard errors are clustered at the firm and year levels. Details are given in appendix A.2.7.

their quality. I explore whether specific patents are of higher quality relative to non-specific ones using two popular measures of patent quality used in the literature (Higham *et al.* (2020)). The first measure is the number of forward citations relative to the average

number of citations for patents granted in the same year and technology class (group level CPC). Citations are taken within a 5 year since the grant day. I call it a scientific value of a patent. The second measure is based on the abnormal stock market return on the grant day of a patent for the firm that owns it (Kogan *et al.* (2017)). I extend Kogan *et al.* (2017) estimates for years up to 2019 using their parameter estimates for the baseline measure (see formula (4) in Kogan *et al.* (2017)). I call it an economic value of a patent.

I run the following regression at the patent level

$$\ln(Q_p) = \beta_s \cdot \mathbb{I}_p^s + \beta_g \cdot \mathbb{I}_p^g + \gamma X_p + \varepsilon_p \tag{1.5.2}$$

where Q_p is the value of patent p (scientific or economic), $\mathbb{I}_p^s = 1$ if patent p is specific and zero otherwise, $\mathbb{I}_p^g = 1$ whether patent p makes a citation to a patent from the top 5% of the most cited patents. The control $\mathbb{I}_p^g$ allows to distinguish whether specific patents are of higher quality because of the concentration of citations or because they simply rely on a highly-cited patent. The variable X_p includes the set of controls for the patent and the firm that owns it. The variable ε_p is an error term.

The results are given in the Table 1.5. All specifications show that specific patents are of higher quality relative to non-specific ones. For the scientific value of patents, the most restrictive specification is given in the column (3), which includes firm-year-technological class fixed effects meaning that specific patents are of higher scientific quality relative to non-specific ones even within the same firm, grant year and a technological class. The specific patents have on average 15% more forward citations relative to non-specific ones that cite a top patent, and 49% more citations relative to patents that do not cite top 5% of the most cited patents.

For the economic value of patents, such a restrictive specification is not feasible due to data limitations.[26] The most restrictive specification is given in the column (6), which

[26]The main problem is that patents are granted only on Tuesdays of each week (except federal holidays),

Table 1.5: Value of Patents as a Function of Backward Citations

	Scientific Value			Economic Value		
	(1)	(2)	(3)	(4)	(5)	(6)
Specific Patent, β_s	0.25***	0.19***	0.15***	0.24***	0.20***	0.03*
	(0.03)	(0.03)	(0.02)	(0.06)	(0.04)	(0.016)
Citation to a Top Patent, β_g	0.35***	0.33***	0.34***	0.13***	0.09***	0.01**
	(0.01)	(0.01)	(0.01)	(0.03)	(0.02)	(0.005)
Firm Controls	Yes	Yes		Yes	Yes	Yes
Industry-Year FE	Yes	Yes			Yes	Yes
Firm FE		Yes				Yes
Firm × Year × Tech Class FE			Yes			

Notes: This table shows the estimation results from the regression 1.5.2 of the value of a patent on a dummy whether this patent is specific or not. Columns (1)–(3) consider the specification with the scientific value of a patent, which is defined as the number of forward citations relative to the average number of citations for patents granted in the same year and technological class. Columns (4)–(6) consider the specification with the economic value of a patent, which is defined based on the abnormal stock market return on the grant day of a patent for the firm that owns it (Kogan *et al.* (2017)). Firm controls include lag log of profits, log of employment and log of capital stock. Standard errors are clustered by firm and year. p-value $^{***}p \leqslant 0.01$, $^{**}p \leqslant 0.05$, $^{***}p \leqslant 0.01$.

includes firm controls, firm and industry-year fixed effects. The specific patents have on average 3% higher economic value relative to non-specific ones that cite a top patent, and 4% higher economic value relative to patents that do not cite the top 5% of the most cited patents.

1.5.2 Welfare Implications of Restrictions on Vertical Contracts

The model in Section 1.3.1 predicts the equilibrium vertical contract as a function of the competition level (α) and the supplier's bargaining position (γ). In this section, I study the welfare properties of the equilibrium and policies that can increase social welfare. Social welfare is defined as the sum of consumer and producer surpluses and is determined

and it is impossible to distinguish the value of patents granted on the same day. In my specifications, I define patents to be specific if all patents granted on a given day were specific. Distinguishing them by technology classes would significantly reduce the sample size.

by the consumers' willingness to pay

$$SW(q_{\mathcal{I}}, q_{\mathcal{E}}) = \frac{1-\alpha}{2} \cdot (q_{\mathcal{I}} + q_{\mathcal{E}}) + \alpha \cdot \max\{q_{\mathcal{I}}, q_{\mathcal{E}}\}$$

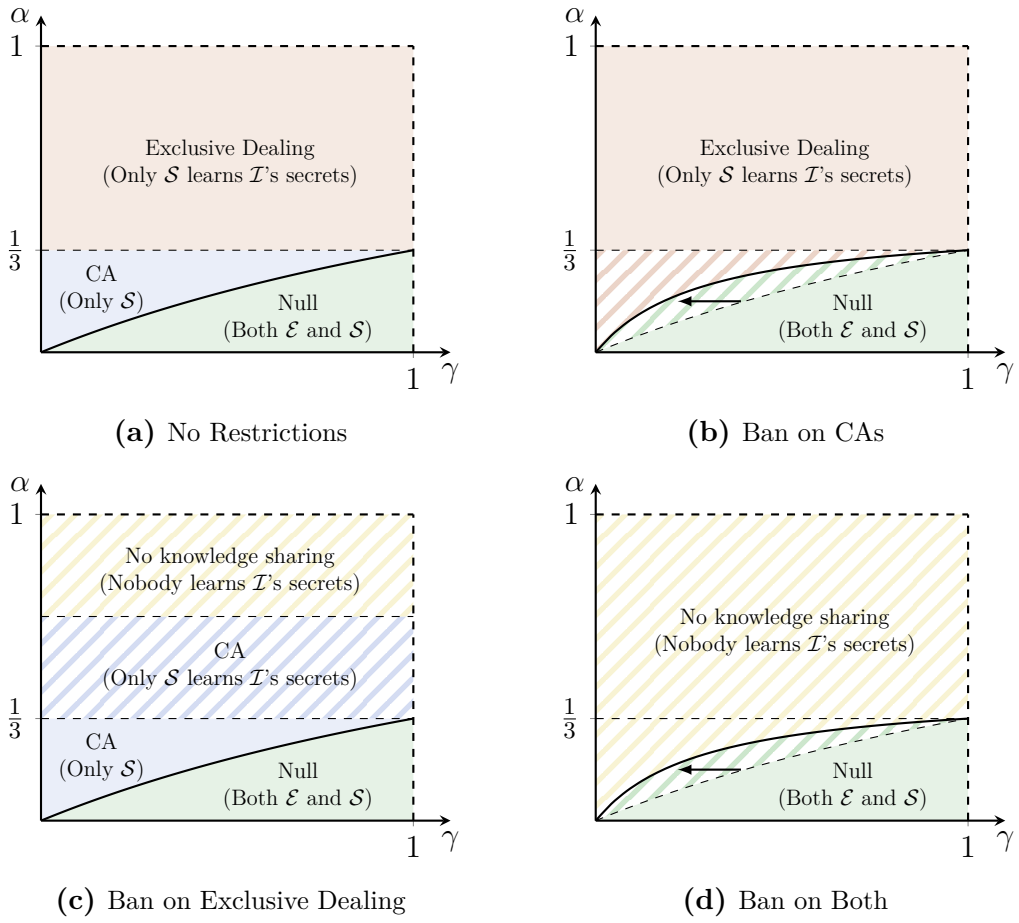
where $q_{\mathcal{I}}$ and $q_{\mathcal{E}}$ are qualities of products from firms $\mathcal{I}$ and $\mathcal{E}$, respectively.

The game in Section 1.3.1 has four possible outcomes with different welfare implications. First, both $\mathcal{E}$ and $\mathcal{S}$ learn $\mathcal{I}$'s private knowledge, and $\mathcal{E}$ uses a high-quality input from $\mathcal{S}$. This is a socially optimal outcome that gives the largest total welfare. It is achieved when $\mathcal{I}$ and $\mathcal{S}$ do not sign any contracts in the first stage (neither exclusive dealing nor a confidentiality agreement), and $\mathcal{I}$ shares knowledge with $\mathcal{S}$. The second outcome is when only $\mathcal{S}$ learns $\mathcal{I}$'s private knowledge, but $\mathcal{E}$ can still use $\mathcal{S}$'s high-quality input. This equilibrium has lower welfare because $\mathcal{E}$-product is of lower quality. This happens when $\mathcal{I}$ and $\mathcal{S}$ sign a confidentiality agreement. The third outcome is when only $\mathcal{S}$ learns $\mathcal{I}$'s private knowledge and $\mathcal{E}$ uses a low-quality input. This equilibrium gives lower welfare than the previous two cases because firm $\mathcal{E}$ works with an inferior input and without $\mathcal{I}$'s knowledge. This happens when $\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing. The final outcome is when nobody learns $\mathcal{I}$'s secrets, and both $\mathcal{I}$ and $\mathcal{E}$ work with an inferior input. However, this outcome is never an equilibrium if $\mathcal{I}$ and $\mathcal{S}$ can sign exclusive dealing. To sum up, the equilibrium outcomes (contracts) can be ranked in terms welfare as

$$\text{Null} \succ_w \text{CA} \succ_w \text{Exclusive} \succ_w \text{No Knowledge Sharing}$$

Policy. Proposition 1 shows that the socially optimal equilibrium arises only under a low level of competition (low α) and high quality of the supplier's input (γ). I discuss benefits and costs of three potential policy interventions: the restrictions on the use of confidentiality agreements, the restrictions on the use of exclusive contracts, and the restrictions on both. The equilibrium in each of these cases is described in Proposition 11, Appendix A.4.3.

Figure 1.10b shows the equilibrium under the restriction on the use of confidentiality



(a) No Restrictions

(b) Ban on CAs

(c) Ban on Exclusive Dealing

(d) Ban on Both

Figure 1.10: Equilibrium Under Restrictions on Vertical Contracts

Equilibrium in (α, γ) space. Figure 3.2 shows the equilibrium without any policy interventions (Proposition 1). Figures 1.10b, 1.10c, and 1.10d show the equilibrium under the restrictions on the use of confidentiality agreements, exclusive dealing, and both, respectively. Hatched areas show the regions where the equilibrium has changed due to a policy regulation. Colors reflect an equilibrium that has replaced the initial one.

agreements (CA). In this case, only an exclusive contract is available to $\mathcal{I}$ and $\mathcal{S}$ in the first stage. Hatched areas in Figure 1.10b show the regions (α, γ) where the equilibrium has changed due to a policy regulation. Colors reflect an equilibrium that has replaced the initial one. If the initial equilibrium is inefficient due to a confidentiality agreement, the ban on CA can achieve a socially optimal outcome if $\mathcal{I}$ would continue sharing its secrets with $\mathcal{S}$ without any contract (condition (1.3.6)). Otherwise, $\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing that gives lower social welfare relative to the initial equilibrium because $\mathcal{E}$ does

not get access to $\mathcal{I}$'s secrets and works with an inferior input.

Figure 1.10c shows the equilibrium under the restriction on the use of exclusive contracts. If the initial equilibrium is inefficient due to exclusive dealing ($\alpha > \frac{1}{3}$), the ban on exclusive contracts can never achieve the socially optimal outcome, but it can increase welfare by providing firm $\mathcal{E}$ the access to the high-quality input from $\mathcal{S}$. However, for sufficiently high level of competition (α) this ban might lead to lower welfare because $\mathcal{I}$ refuses to share its secrets with $\mathcal{S}$.

Figure 1.10d shows the equilibrium under the restriction on both confidentiality agreements and exclusive contracts. In this case, $\mathcal{I}$ shares knowledge with $\mathcal{S}$ only if the competition (α) is low enough and the gain in the input quality (γ) is high enough (condition (1.3.6)). Otherwise, $\mathcal{I}$ does share its secrets with anybody, and this gives the lowest possible welfare.

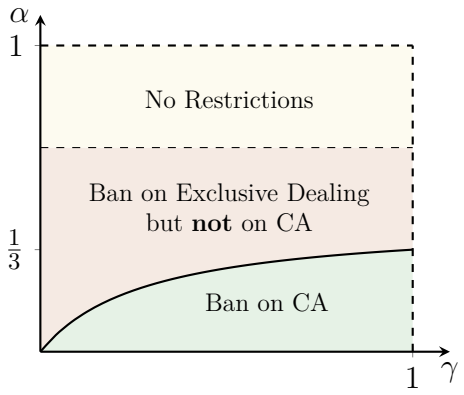


Figure 1.11: Optimal policy as a function of (α, γ)

This figure illustrates the optimal restrictions on confidentiality and exclusive contracts as a function of (α, γ). Parameter α measures the degree of competition between $\mathcal{I}$ and $\mathcal{E}$, and parameter γ reflects the value of $\mathcal{S}$'s input relative to $\mathcal{I}$'s and $\mathcal{E}$'s outside options.

Summary. Optimal Policy

Figure 1.11 illustrates the optimal set of restrictions on confidentiality and exclusive contracts in the model with a common supplier from Section 1.3.1. The precise

characterization is given in equation (A.4.14) of Appendix A.4.3. If the competition between $\mathcal{I}$ and $\mathcal{E}$ is high, the restriction on exclusive dealing can only decrease welfare because $\mathcal{I}$ will not share its knowledge with $\mathcal{S}$. For the intermediate level of competition, the ban on exclusive dealing can increase welfare because it allows $\mathcal{E}$ to get access to a high-quality input from $\mathcal{S}$. However, the planner should leave confidentiality agreements to ensure that $\mathcal{I}$ continues to share its knowledge with $\mathcal{S}$. If the competition is low, the restriction on confidentiality agreements will increase welfare only when the input from $\mathcal{S}$ is sufficiently valuable relative to $\mathcal{I}$'s and $\mathcal{E}$'s outside options (high γ).

Notice that for the social planner who does not know the details of the economy $(\alpha, \gamma, \lambda_0/\lambda)$ it is never optimal to simultaneously ban both the exclusive and confidentiality contracts. If the economy is such that the restrictions on vertical contracts can lead to the first best outcome, the ban on confidentiality agreement will be sufficient. Otherwise, the simultaneous restriction on both types of contracts will lead to the equilibrium with no knowledge sharing.

From a policy perspective, the framework with a common customer who can buy inputs from several suppliers is different from the environment with a common supplier who can sell inputs to multiple downstream firms. In the model from Section 1.3.2, a customer who can buy inputs from several suppliers signs a confidentiality agreement or exclusive dealing with the incumbent supplier in order to create incentives for this supplier to share its private knowledge. They do not use these contracts to increase their total surplus by excluding another supplier. Therefore, the restrictions on vertical contracts can only decrease welfare because firms will stop sharing knowledge with each other.

The results in this section should be interpreted with caution because they are based on the highly stylized framework. The central idea is that contractual arrangements between firms affect the diffusion of knowledge in the economy, and there is a room for policy interventions. The analysis of a more general environment is a fruitful area for future

research.[27]

1.6 Conclusion

This paper shows that patents play a limited role in the diffusion of knowledge, and patent citations provide a poor measure of knowledge spillovers. I argue that patent citations reflect intentional knowledge sharing between firms producing complementary products. I develop and test a theory of knowledge sharing in a production network, and how common suppliers and customers diffuse knowledge across competitors.

Much of the literature on economic growth (Jones (2005), Aghion *et al.* (2014)) is based on the idea of knowledge spillovers, and their systematic measurement relies on patent citations (Trajtenberg and Jaffe (2002)). The evidence in this paper raises concerns whether spillovers in the growth models can be disciplined by patent citations. The inability to measure knowledge externalities brings back the critique by Krugman (1991) that "... there is nothing [in the data] to prevent the theorist from assuming anything about them that she likes". Patent citations still provide valuable information about collaboration between firms, but the growth theories are often focused on creative destruction (competition) rather than creative construction (cooperation). More studies of strategic alliances between firms coupled with patent citations can bring new insights about innovations and technological progress.

This paper has also implications for innovation policy. The presence of knowledge spillovers provides a rationale for R&D subsidies. While some activities such as scientific research in corporate labs, academia and government agencies generate a lot of spillovers to others (Moretti (2012), Arora *et al.* (2019), Gruber and Johnson (2019)), it is less clear with technologies that companies choose to patent. For these technologies, firms might

[27]See Whinston (2006, ch. 4) for the comprehensive review of the literature on exclusionary vertical contracts.

control the diffusion of their knowledge through contractual arrangements with business partners, and I discuss benefits and costs of regulation of these contracts. Empirical and theoretical studies on how contracting between companies or within firms with their employees affects the diffusion of knowledge is an interesting area for future research.

Finally, an important area for future research is related to the boundaries of the firm and management of knowledge flows. Contracting within a firm might differ from contracting across firms. For example, it is easier to enforce confidentiality agreements with employees rather than with other companies (Almeling *et al.* (2010, 2011)). A firm might integrate a supplier in order to keep its knowledge secret, or it might outsource certain activities to a supplier serving multiple firms in order to learn from competitors. Patent self-citations of a firm to itself relative to citations with other companies can provide a measure of within-firm transactions, which is crucial for empirical tests of the theories on the boundaries of the firm.

Chapter 2

The Role of Trade Secrets in the US Economy: Evidence from Litigation[1]

2.1 Introduction

Firms protect their intellectual property (IP) though a range of mechanisms, including patents and trade secrets. The economics literature has mostly focused on patents (Hall *et al.* (2014)) despite that multiple surveys (e.g., Cohen *et al.* (2000)), case studies (e.g., Arora (1997)), and legal papers (e.g., Almeling (2012)) emphasize that firms often rely heavier on secrecy rather than on patents to appropriate returns on innovation. One of the main differences between secrecy and patents is that patents require the disclosure of knowledge to the public, while trade secrets can be hidden for an infinite period of time.[2] Therefore, the choice between secrecy and patents has implications for the diffusion of knowledge and economic growth.

One reason for a scarcity of research on trade secrets is that it is hard to track them:

[1] Co-authored with Alice Wu.

[2] The disclosure function of patents is not perfect (Roin (2005), Fadeev (2022)), but they still provide valuable information about the types of technologies firms develop and basic ideas behind them.

secrets are not observable and are not registered in government agencies. In this paper, we are tying to enhance our understanding of trade secrets using litigation data. We collect data on lawsuits filed in the US federal courts from 2009 to 2020, and in which at least one claim is related to misappropriation of trade secrets. We use the textual analysis of litigation records to study the types of information firms protect as trade secrets and the mechanisms for the misappropriation of these secrets. These mechanisms show how private knowledge can continue to flow between firms, even in the absence of public knowledge disclosure.

We identify two types of intellectual property: business information, such as customer lists, and technological information, such as chemical compounds. The technological information is more likely to be patent-eligible, but firms decided to keep it secret for strategical reasons. Within technological trade secrets we also classify information related to two technologies prominent in recent decades: software and biotech/pharmaceuticals. To identify the mechanisms of misappropriation, we read a random sample of cases and rely on Almeling *et al.* (2010) who also did the analysis of a sample of litigation cases. This analysis shows that in almost all lawsuits the misappropriation of trade secrets happened through employees moving to other companies or through business partners, for example, common input suppliers. We classify cases involving employees as the ones which have at least one person among defendants and which have keywords for a breach of contract in the legal complaint. The examples of such keywords are "noncompete", "solicitation", "poaching", and others.

We show that the litigation activity has increased after 2016 when the U.S. Congress signed the Defend Trade Secrets Act (DTSA). The increase in litigation activity was especially salient among cases involving business information and software technologies. Traditionally, trade secrets were governed by state laws, and the DTSA created for the first time a federal civil cause of action for misappropriation of trade secrets. In general, the passage of the DTSA has improved the protection of trade secrets.

We also show that there is a significant variation in trade secrets litigation across the US states, industries, and firms. Firms from manufacturing industries account for around 50% of litigation cases. These cases are more likely to be about technological trade secrets, and the misappropriation of trade secrets is more likely to happen through business partners rather than employees. We show that across all industries larger firms are more likely to file a suit against an employee who breached contractual obligations, such as non-disclosure, non-solicitation, or non-compete agreements. Larger firms are also more likely to file lawsuits for misappropriation of business information, are more likely to be defendants in misappropriation of software trade secrets, and are less likely to be defendants in biotech and pharmaceutical cases.

We develop a theory that clarifies the conditions under which the litigation of secrets occurs. The theory is based on the model of knowledge sharing in a supply chain from Fadeev (2022). In the theory, the incumbent decides whether to share its trade secrets with the supplier who can leak them to the incumbent's competitor. The incumbent can sign a non-disclosure agreement (NDA) with the supplier, but this agreement is not perfectly enforceable: if the supplier breaches the contract, and the incumbent sues the supplier, it will win the lawsuit with a probability less than one.

The litigation happens in a highly competitive environment and when the protection of trade secrets is neither too strong nor too weak. If the competition between the incumbent and its competitor is low, then then the incumbent does not sign NDA with the supplier. If the competition is high but contracts are too weak, the incumbent does not share its secrets with the supplier because in expectation it cannot recover the damages from knowledge leakages to the competitor. If the competition is high and contracts are too strong, then the incumbent signs NDA with the supplier, it has incentives to share knowledge with it, but the supplier has no incentives to leak the incumbent's knowledge. In both cases, there is no litigation. The litigation happens under the intermediate level of trade secret protection in which both the incumbent shares knowledge with the supplier, and the

supplier leaks these secrets to the competitor.

The theory also shows how to evaluate welfare consequences of changes in trade secret protection, for example, the passage of the DTSA. Stronger protection gives incentives for owners of secrets to share knowledge with business partners and employees, but discourages further knowledge diffusion through these partners and the movement of employees. We show that the net effect of greater protection depends on the degree of horizontal competition and on the initial level of trade secret protection. In our stylized framework, changes in the litigation activity following the DTSA serve as a sufficient statistic for welfare evaluation.

2.2 Facts on Trade Secret Litigation

2.2.1 Background

Historically, trade secret misappropriation was governed by the case law of each state. In 1979, the National Commission on Uniform State Laws promulgated the Uniform Trade Secrets Act (UTSA) which was drafted by the American Bar Association (Menell *et al.* (2021)). The goal of the UTSA was to unify trade secret law across states, but states did not interpret the statues in a uniform way (e.g., Parker III and Justice (2014)), and New York still did not adopt the UTSA at all.

In 2016, the U.S. government signed the Defend Trade Secrets Act (DTSA) which created for the first time a federal *civil* cause of action for misappropriation of trade secrets. Before DTSA, trade secret disputes were litigated in federal courts mostly under the Economic Espionage Act (EEA) which criminalized trade secret misappropriation or under the Computer Fraud & Abuse Act (CFAA) which provided both criminal and civil remedies. The goal of the DTSA was to unify trade secret law, to provide litigants with the access to resources of federal courts, and to handle extraterritorial and international

disputes. This act generated a lot of debates among legal scholars who argued that the federal law might be in conflict with state laws. However, the DTSA does not preempt the state law, and litigants can choose whether to file a suit in a state or a federal court.[3] Overall, the passage of the DTSA provided a greater enforcement of laws against trade secret misappropriation.

2.2.2 Data and Methodology

One of the problems with data on the litigation of trade secrets is that these lawsuits are not identified in the courts' internal tracking system. For example, the misappropriation of trade secrets is often recorded under contractual disputes. We use the proprietary data from Lex Machina who classified trade secret cases in federal courts using the natural language processing and manual checking of cases. We use the following information from litigation records: parties involved in litigation (plaintiffs, defendants, third parties), court venues, filing years, termination years, case resolutions, and texts of legal complaints.[4] A legal complaint describes the nature of a case and information that was allegedly misappropriated.

We match parties in litigation records to names in Compustat North America. From Compustat, we take data on industry affiliation and sales for publicly traded firms. To account for name changes, mergers and acquisitions, and subsidiary information we use the dynamic matching of patents to Compustat from Autor *et al.* (2020) and Fadeev (2022).

We use the texts of legal complaints to study types of information that firms protect as trade secrets. In particular, we distinguish between business and technological information. The examples of business information are client lists, pricing strategy, marketing plans, and others. Cases about business information are identified using the combination of keywords

[3]For more debates on the benefits and costs of the DTSA see Goldman (2014) and Levine and Seaman (2018).

[4]In cases where the complaints are unavailable we took the first file in legal dockets.

such as “client”, “pricing”, “insurance”, and others. For technological information, we use keywords such as “technology”, “patenting”, “invention”, and others. We classify cases involving business (technological) information as the ones in which complaints contain at least two business (technological) keywords. Within cases with technological trade secrets, we also identify information related to software and biotech/pharmaceutical technologies using keywords and two-word phrases (“technical bigrams”), such as “cloud computing” and “machine learning” for software, and “biotech” and “antibody” for biotech/pharmaceuticals. We take some of the technological phrases from Bloom *et al.* (2021) who used the full text of patents, job postings, and conference calls to identify bigrams associated with disruptive technologies. The full list of keywords is given in Table B.1 in Appendix B.

Next, we use the texts of legal dockets to study the mechanisms of misappropriation of trade secrets. In particular, we identify litigation cases where the misappropriation was likely to happen through an employee. We identify cases which have at least one person among defendants and which have keywords for the breach of contractual obligations, for example, “non-compete”, “solicitation”, “poaching”, and others. The full list is given in Table B.1 in Appendix B.

2.2.3 Empirical Results

The section describes the main empirical results. We document aggregate litigation activity and its composition over time, across US states, across industries, and across firms based on their size.

Figure 2.1 shows the total number of cases filed in each year from 2009 to 2020. The number of cases was stable until 2016, and then it jumped by 30% after 2016 when the Defend Trade Secrets Act was signed. Unfortunately, we do not have data on cases filed in state courts, so it is not clear to what extent firms shifted their litigation activity from state to federal courts.

Figure 2.2a shows the share of cases by their resolution. Around 81% of cases are settled, and in around 15% of cases plaintiffs win. There is a slight increase in the share of plaintiffs' victories after 2016, and a slight decline in the share of settled cases.

Figure 2.2b shows the share of cases involving the misappropriation of trade secrets by employees. The upper line shows that around 80% of cases have people among the defendants. The bottom line shows the share of cases which have at least one person among the defendants, and in which legal complaints contain words for a breach of contractual obligations. Overall, the share of such cases is around 50%, but it has increased since 2009, and the rise is especially pronounced from 2015 to 2018.

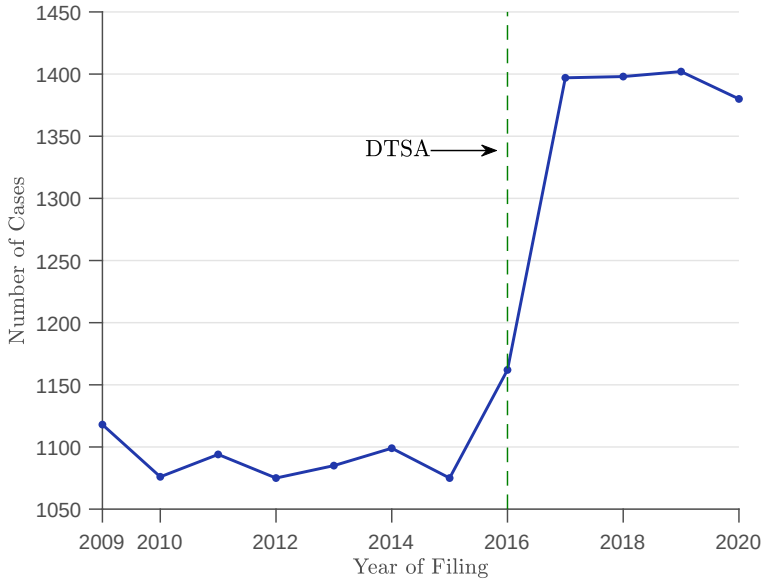
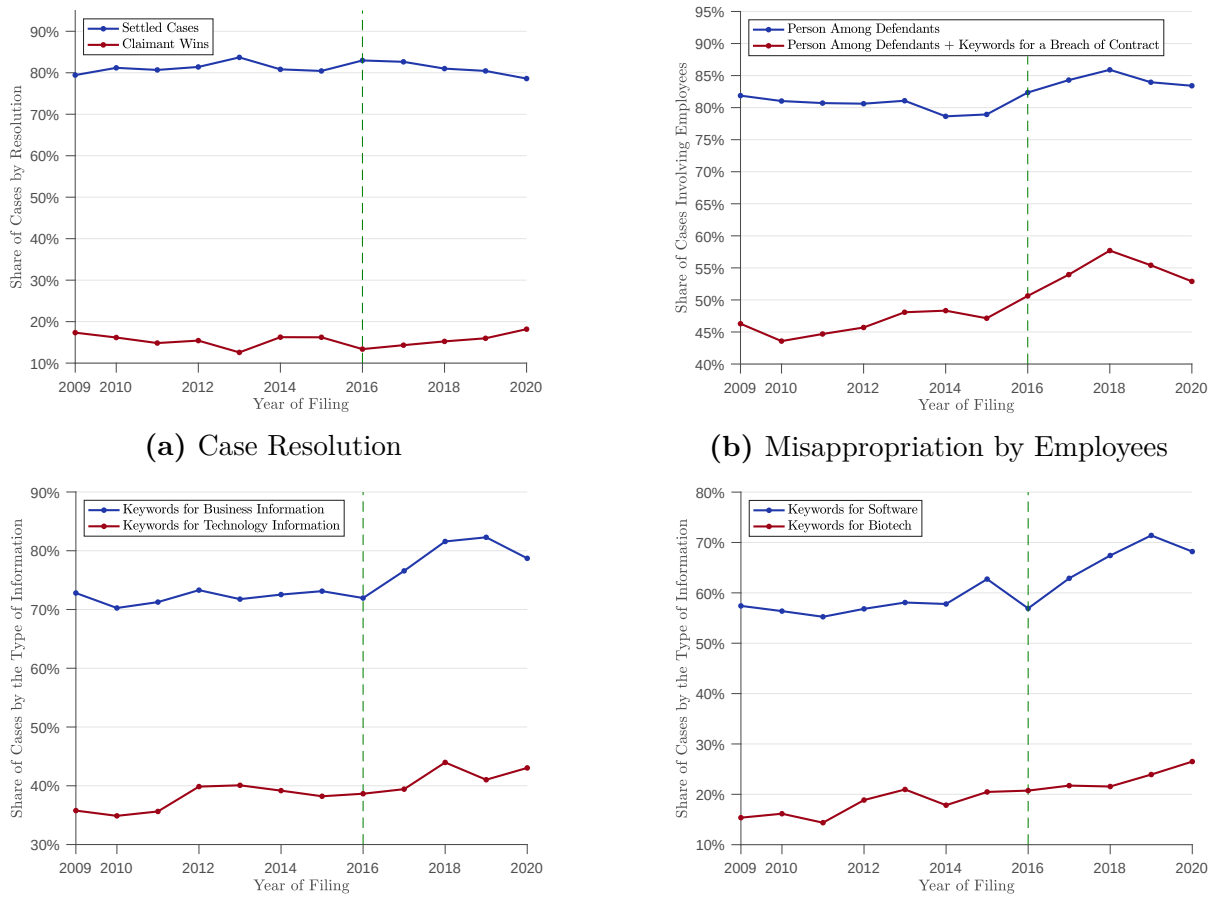


Figure 2.1: Number of Trade Secret Litigation Cases by Years

This figure shows for each year the number of lawsuits filed in the US federal courts and which have at least one claim related to the misappropriation of trade secrets.

Figure 2.2c shows the share of cases based on the type of protected trade secrets. The upper line shows the share of cases which contain keywords for business information, such as customer lists or pricing strategy. The bottom line shows the share of cases with keywords for technological information. These types of information are not mutually



(a) Case Resolution

(b) Misappropriation by Employees

(c) Type of Information

(d) Software and Biotech/Pharma

Figure 2.2: Trade Secret Litigation over Time

These figures shows for each year the composition of trade secret litigation based on the lawsuits' resolutions, mechanism for misappropriation of trade secrets, and type of misappropriated information. Figure 2.2a shows the share of cases in which parties reached a settlement and in which the plaintiffs win. Figure 2.2b shows the share of cases in which the misappropriation is likely to happen through employees. Figure 2.2c shows the share of cases by the type of the misappropriated information: business vs. technological. Figure 2.2d shows the share of cases among technological ones in which the misappropriated information is related to software technologies and (or) to biotech and pharmaceutical technologies.

exclusive, and cases can be about both business and technological information at the same time, for example, as described in Motorola's complaint against Hytera Communications:

> "Hytera's misappropriation was deliberate, wholesale, and systematic—not only did Hytera take and then copy Motorola's technical trade secrets, it even

> copied the marketing, configurations, and product manuals related to the misappropriated features as well ..."

Both types of trade secrets show an upward trend over time, but after 2016 the share of business trade secrets experienced a much more pronounced increase relative to the technical secrets.

Finally, Figure 2.2d shows the share of cases among technological ones which have keywords for software or (and) biotech/pharmaceutical technologies. The cases with software technologies experienced a more pronounced increased after the DTSA relative to biotech and pharmaceutical trade secrets.

Figures 2.3 – 2.6 show the distribution of cases across US states. Table B.2 in Appendix B also shows more detailed information for 15 most litigious states. Top 5 states in terms of the number of cases are California, Texas, Florida, New York, and Pennsylvania. There is a significant variation across states in the composition of cases. For example, California has a high share of cases involving technological trade secrets, but a low share of cases involving breaches of contracts by employees. This is probably caused by strict state laws regarding employees' protection, for instance, a ban on non-compete agreements. In contrast, Florida has one of the lowest shares of cases involving technological trade secrets, but it has a high share of cases about employees misappropriating trade secrets.

We also document the variation in the number and the composition of cases across 3-digit industries, according to Standard Industry Classification. The most litigious industry is "Computer Programming, Data Processing, And Other Computer Related Services" (SIC 737) which accounts for 12% of litigation cases. Table B.3 in Appendix B lists top 10 industries in terms of the number of cases. Among top 10, there are 6 manufacturing industries that can be divided into medical instruments, drugs, computers, and motor vehicles, and there are 4 non-manufacturing industries including insurance, finance, and wholesale trade. On average, manufacturing industries have around 65% of cases involving technical trade secrets, while non-manufacturing industries have only 36% of such cases.

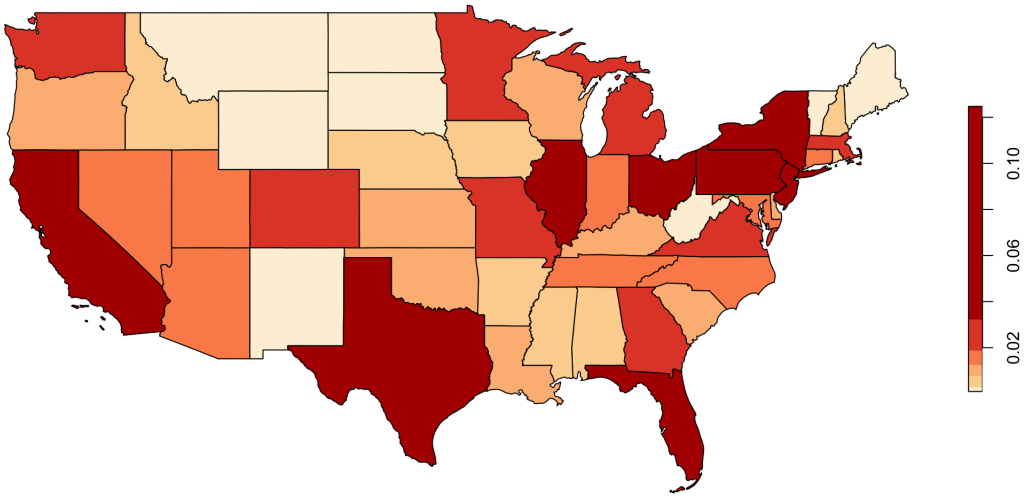


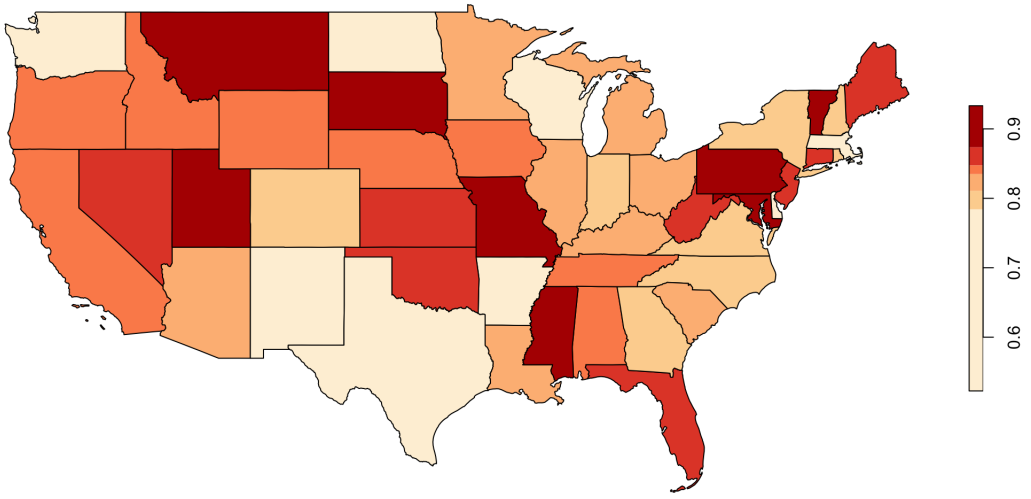
Figure 2.3: Share of Litigation Cases Across US States

This figure shows the share of federal litigation cases involving the misappropriation of trade secrets, across US states.

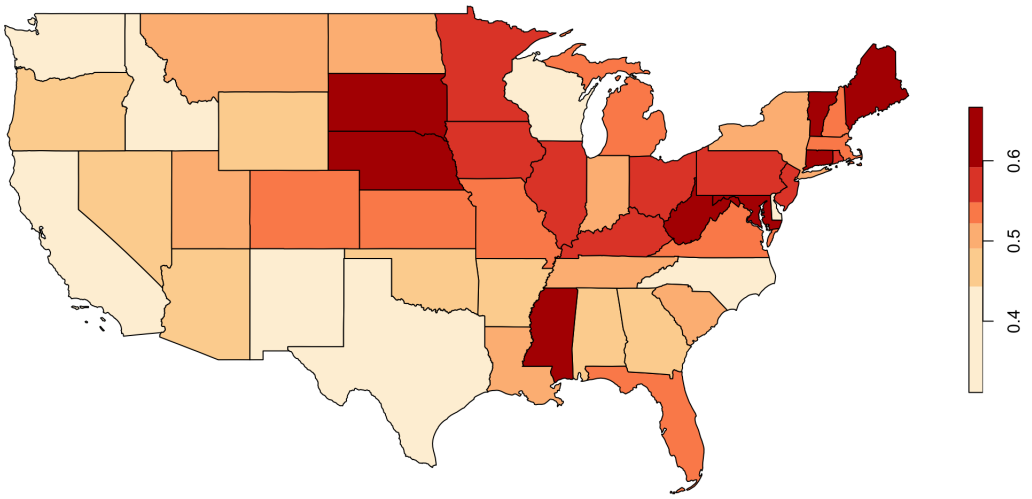
In terms of the source of misappropriation of secrets, manufacturing industries have only around 36% of cases against employees breaching their contractual obligations, while in non-manufacturing ones this share is almost 72%.

Finally, we consider the variation in the number and the composition of cases across firms based on their size. Within each 3-digit SIC industry, we divide firms into four equally-sized bins based on their average sales between 2009 and 2020. Figure 2.7 shows that the litigation activity is primarily concentrated among large firms. Top 25% of firms by sales account for around 70% of cases, both among plaintiffs and defendants.

Figure 2.8 shows the composition of trade secret cases across firms of different sizes. Figure 2.8a shows that large firms are more likely to file suits against employees for a breach of contract (non-compete, NDA, etc): firms in the bottom 25% have around 40% of such cases, while firms in the top quartile have more than 60% of cases involving employees. There are no significant differences between small and large firms in terms of being defendants in such cases. Figure 2.8b shows that small and large firms are not



(a) A Person Among Defendants

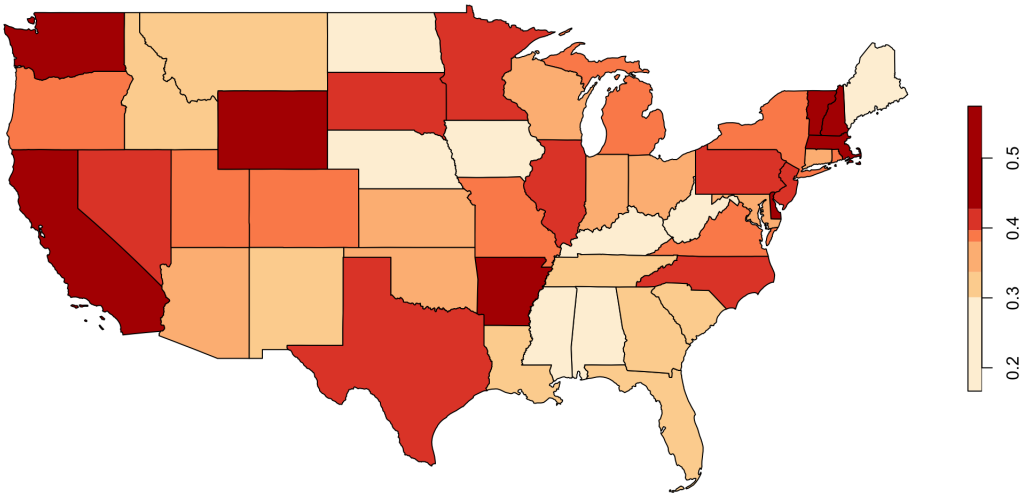


(b) A Person Among Defendants + Keywords for a Breach of Contract

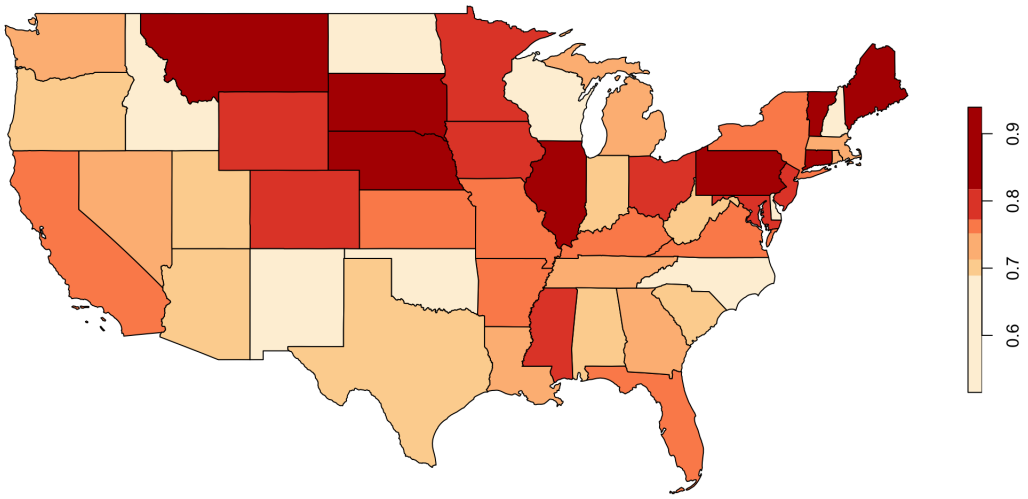
Figure 2.4: Distribution Across States of Cases Involving Employees

For each state we compute the share of cases involving in which the misappropriation of trade secrets was likely to happen through employees. Figure 2.4a shows the share of cases in which at least one of the defendants is a person. Figure 2.4b includes a more conservative definition in which we also require the complaint to contain keywords for a breach of contractual obligations by an employee.

different from each other in terms of litigation of technological trade secrets. However, figure 2.8c shows that large firms are more likely to file suits against the misappropriation of business information: 55% of cases for the bottom 25% versus more than 75% for the top



(a) Keywords for Technological Information

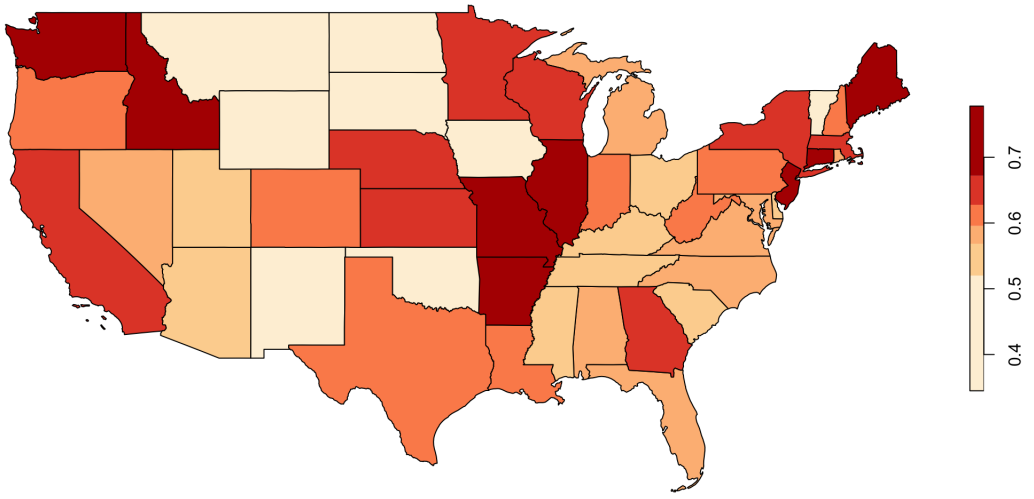


(b) Keywords for Business Information

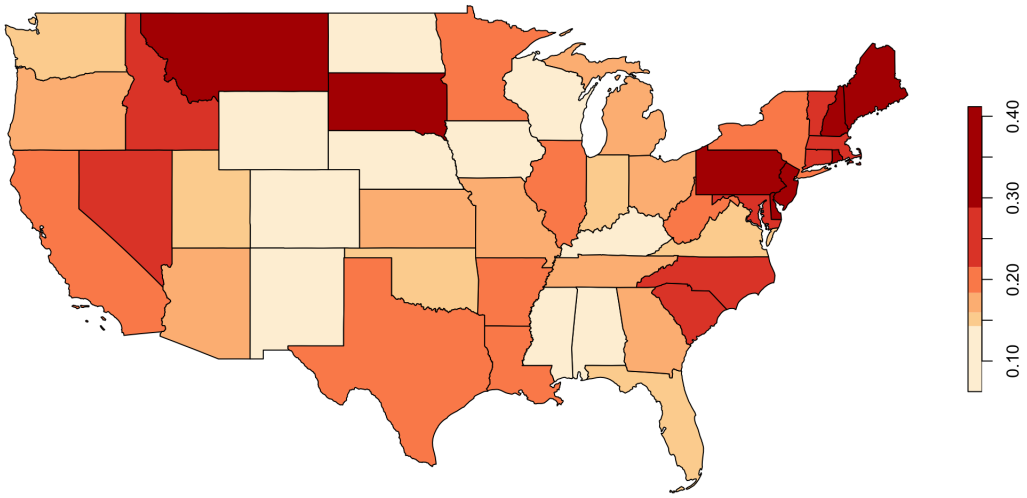
Figure 2.5: Distribution Across States of Cases based on the Types of Protected Information

For each state we compute the share of cases in which the misappropriation of trade secrets involved technical information (2.5a) and business information (Figure 2.5b).

25%. Large firms are also more likely to be defendants in such cases. Finally, large firms are more likely to be defendants in misappropriation of software trade secrets, and are less likely to be defendants in misappropriation of trade secrets related to biotech/pharmaceutical technologies. For biotech and pharmaceutical technologies, small firms are also more likely to be plaintiffs: 50% in the bottom quartile versus 30% in the upper one.



(a) Keywords for Software Technology



(b) Keywords for Bio/Pharmaceutical Technology

Figure 2.6: Distribution Across States of Cases based on the Types of Protected Information

For each state we compute the share of cases among technological ones in which the misappropriation of trade secrets involved software (2.6a) and biotech/pharmaceutical (Figure 2.6b) technologies.

Overall, the evidence in this section suggests that the trade secret litigation activity significantly varies across US states, industries, and firms. Therefore, the Defend Trade Secrets Act signed in 2016 had a differential impact on various industries and firms.

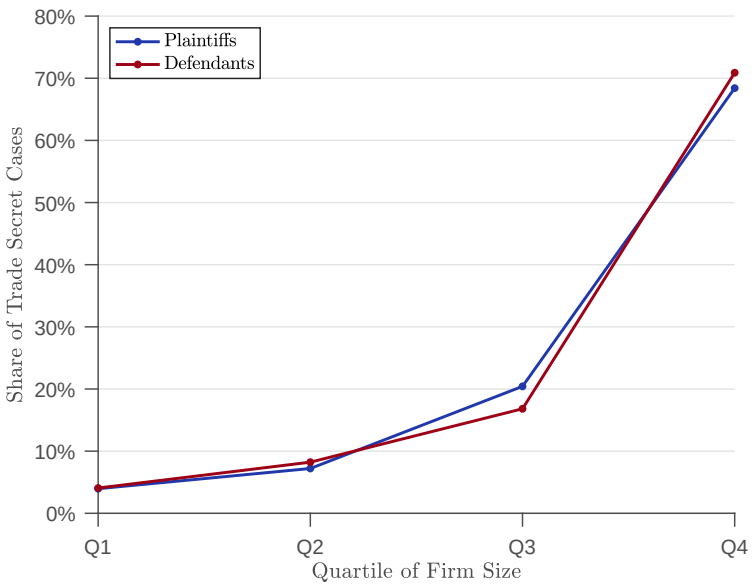


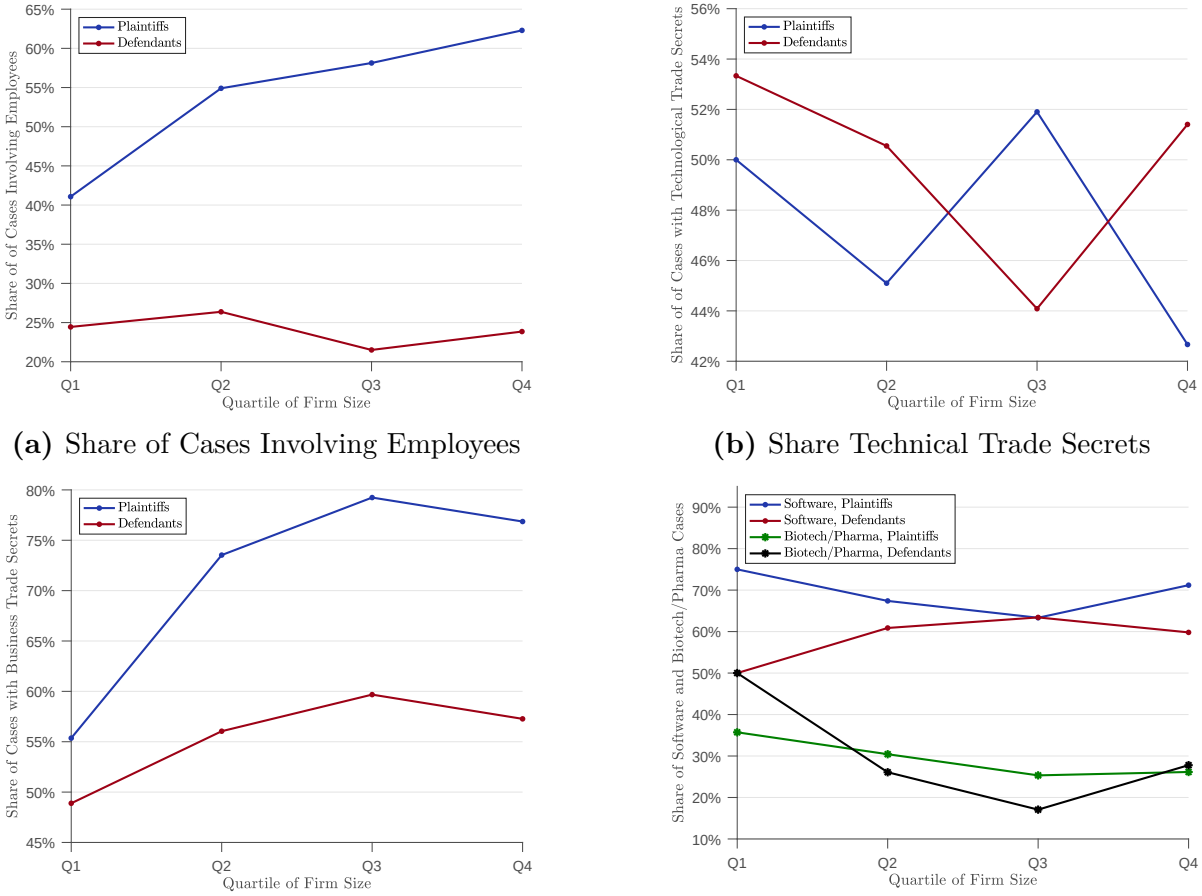
Figure 2.7: Share of Litigation Cases by Firm Size

This figure shows the share of cases by firm size. Within each 3-digit SIC industry we find quartiles (25-th, median, 75th) of the average sales between 2009 and 2020. Then we compute the share of cases, separately for plaintiffs and defendants, within each size bin.

2.3 Theory: Regulation of Trade Secrets

The previous section shows two main results. First, there is a significant variation in trade secrets litigation across US states, industries, and firms. Second, the litigation activity has increased after 2016 when the U.S. Congress signed the Defend Trade Secrets Act. The increase in litigation activity was especially salient among cases involving business information and software technologies.

In this section, we develop a theory to clarify two questions. First, when should we expect to see the litigation activity? Second, how should we analyze the welfare implications of the Defend Trade Secrets Act? The theory is the modification of the model of knowledge sharing between firms from 1.3.1. The main deviation is that we assume that contracts are not perfectly enforceable, and this will lead to breaches of contractual arrangements and litigation activity. For brevity, we focus on knowledge sharing between



(a) Share of Cases Involving Employees

(b) Share Technical Trade Secrets

(c) Share of Business Trade Secrets

(d) Software and Biotech/Pharma

Figure 2.8: The Composition of Trade Secrets By Firm Size

For each firm size bin, we compute the composition of trade secret litigation, separately for plaintiffs and defendants: share of cases involving employees (Figure 2.8a), technological trade secrets (Figure 2.8b), business trade secrets (Figure 2.8c). Figure 2.8d shows the share of cases among technological ones in which the misappropriated information is related to software technologies and (or) to biotech and pharmaceutical technologies.

firms in a supply chain where one supplier can serve two competitors. The theory can be easily modified for other types of business relationships, and for the case of knowledge sharing with workers and the movement of employees across firms.

Simplifying Assumptions. We consider the model from Section 1.3.1. To simplify the analysis we abstract from exclusive dealing and assume that firms have access to confidentiality agreements only. Second, we assume that the competition between the

incumbent ($\mathcal{I}$) and the entrant ($\mathcal{E}$) is not too high, $\alpha \leqslant \frac{1}{3}$. Under this assumption, I showed in Section 1.3.1 that the owner of trade secrets ($\mathcal{I}$) and the supplier ($\mathcal{S}$) have no incentives to sign exclusive dealing.

New Assumptions. We assume that contracts are not perfectly enforceable. Formally, if the owner of trade secrets ($\mathcal{I}$) and the supplier ($\mathcal{S}$) sign a confidentiality agreement, then in the case of a breach of this agreement by $\mathcal{S}$, firm $\mathcal{I}$ can sue $\mathcal{S}$ and win the case with probability q.[5] If $\mathcal{I}$ wins the case, then $\mathcal{S}$ has to pay damages to $\mathcal{I}$ which are proportional to the loss in profits for $\mathcal{I}$ in the final stage of the game:

$$F = \kappa \cdot \text{Loss of Profits for } \mathcal{I} = \kappa \cdot \Big[\alpha\left(\lambda - \lambda_0\right) - \alpha\left(\lambda - \lambda\right)\Big]$$

where $\kappa \geqslant 1$ is the coefficient of proportionality. Therefore, the protection of trade secrets is governed by two parameters: the probability a contractual enforcement, q, and the costs of breaching a contract, κ. The equilibrium will depend on the product $q \cdot \kappa$, so for notation simplicity and without loss of generality we assume that $\kappa = 1$.[6]

The timing of the game is the same as in Section 1.3.1. In stage $t3$, if $\mathcal{I}$ and $\mathcal{S}$ do not sign a confidentiality agreement, then $\mathcal{S}$ has incentives to leak $\mathcal{I}$'s secrets to $\mathcal{E}$. If $\mathcal{I}$ and $\mathcal{S}$ sign a confidentiality agreement, the supplier $\mathcal{S}$ will leak $\mathcal{I}$'s only if the gains from knowledge sharing for $\mathcal{S}$ exceed the expected damages that it needs to pay to $\mathcal{I}$:

$$\underbrace{\frac{1-\alpha}{2}\gamma\lambda - \frac{1-\alpha}{2}\gamma\lambda_0}_{\text{Gains from Knowledge Leakages}} \geqslant \underbrace{q \cdot F = q \cdot \alpha\left(\lambda - \lambda_0\right)}_{\text{Expected Costs}}$$

which can be rewritten as

$$q \leqslant \frac{1-\alpha}{2\alpha} \cdot \gamma \text{ or } \alpha \leqslant \frac{\gamma}{\gamma + 2q} \tag{2.3.1}$$

Inequalities in (2.3.1) show that $\mathcal{S}$'s incentives to leak $\mathcal{I}$'s knowledge to $\mathcal{E}$ depend on

[5] In Section 1.3.1, I assumed that contracts are perfectly enforceable, $q = 1$.

[6] The equilibrium with $q \cdot \kappa > 1$ is similar to the case $q \cdot \kappa = 1$.

the probability of losing a case in a court (q) and on the degree of the competition between $\mathcal{I}$ and $\mathcal{E}$. With higher competition, there are two effects. First, there are less gains from knowledge sharing, $\frac{1-\alpha}{2}\gamma(\lambda - \lambda_0)$, because profits dissipate under competitive pressure. Second, there are greater damages to be paid to $\mathcal{I}$ in a case of a loss in a court, $F = \alpha\,(\lambda - \lambda_0)$.

In stage $t2$, $\mathcal{I}$ decides whether to share its secrets with $\mathcal{S}$. The trade-off for $\mathcal{I}$ is similar to the one described in Section 1.3.1: knowledge sharing with $\mathcal{S}$ allows $\mathcal{S}$ to produce an input of higher quality, but $\mathcal{E}$ can also use the input of higher quality from $\mathcal{S}$, and $\mathcal{S}$ can leak $\mathcal{I}$'s secrets. With perfect enforcement of contracts, a confidentiality agreement eliminates the problem of knowledge leakages. However, with imperfect contracts the problem of knowledge leakages depends on $\mathcal{S}$'s incentives. If q is low enough, so that inequality (2.3.1) is satisfied, then $\mathcal{S}$ will leak $\mathcal{I}$'s secrets to $\mathcal{E}$ even if they signed an NDA. In this case, $\mathcal{I}$ still shares knowledge with $\mathcal{S}$ if

$$\underbrace{\frac{1-\alpha}{2}(1-\gamma)\lambda + \frac{1}{2}\cdot\frac{1-3\alpha}{2}}_{\mathcal{I}'s \text{ profits under knowledge sharing}} + q\cdot F \geqslant \underbrace{\frac{1-\alpha}{2}(1-\gamma)\lambda + \alpha(1-\gamma)(\lambda-\lambda_0)}_{\mathcal{I}'s \text{ profits under no knowledge sharing}}$$

which can be rewritten as

$$\alpha \leqslant \frac{\gamma}{(1-q)4\left(1-\frac{\lambda_0}{\lambda}\right)+\gamma\left(4\frac{\lambda_0}{\lambda}-1\right)} \tag{2.3.2}$$

Under $q = 0$, equation (2.3.2) is the same as the $\mathcal{I}$'s incentive-compatibility constraint for knowledge sharing with $\mathcal{S}$ under no contractual arrangement (see equation (1.3.6)).

The equilibrium of the game is summarized in Proposition 2.

Proposition 2. *Denote the following thresholds for the competition level as*

$$\alpha^\star = \min\left\{\frac{\gamma}{\gamma+2}, \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)}\right\}$$
$$\alpha^* = \min\left\{\frac{\gamma}{\gamma+2q}, \frac{\gamma}{(1-q)\cdot 4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)}\right\}$$

The subgame perfect equilibrium is the following:

$$\begin{cases} \text{CA Without Knowledge Leakages} & \text{if } \alpha > \frac{\gamma}{\gamma+2q} \\ \text{No Knowledge Sharing} & \text{if } \alpha^* < \alpha \leqslant \frac{\gamma}{\gamma+2q} \\ \text{CA With Knowledge Leakages} & \text{if } \alpha^\star < \alpha \leqslant \alpha^* \\ \text{Null Contract} & \text{if } \alpha \leqslant \alpha^\star \end{cases} \tag{2.3.3}$$

Proof. $\mathcal{I}$ and $\mathcal{S}$ have incentives to prevent knowledge leakages to $\mathcal{E}$ if $\alpha \geqslant \frac{\gamma}{\gamma+2}$ and allow leakages otherwise. Given that a confidentiality agreement is enforced with probability q, the incentive-compatibility (IC) condition for knowledge sharing in $t2$ by $\mathcal{I}$ is given in (2.3.2). Under $q = 0$, this condition also shows the IC constraint for $\mathcal{I}$ when it does not sign an confidentiality agreement with $\mathcal{S}$. The incentive-compatibility condition for knowledge leakages in $t3$ by $\mathcal{S}$ is given in (2.3.1). Combining all these conditions gives the equilibrium described in (2.3.3). □

There are two main differences of the equilibrium with imperfect contracts from the equilibrium in Section 1.3.1 with perfectly enforceable contracts. First, since contracts are not perfectly enforceable, there are situations under which firm $\mathcal{I}$ decides not to share its knowledge with $\mathcal{S}$. With perfectly enforceable contracts, $\mathcal{I}$ always shares knowledge $\mathcal{S}$ because they can solve the problem of knowledge leakages with a confidentiality agreement. Second, under $\alpha^* < \alpha \leqslant \alpha^\star$ firms $\mathcal{I}$ and $\mathcal{S}$ sign a confidentiality agreement, but $\mathcal{S}$ still breaches it, and $\mathcal{I}$ sues $\mathcal{S}$. With enforceable contracts, $\mathcal{S}$ never leaked $\mathcal{I}$'s secrets if it signed an NDA, and there was no litigation.

Imperfect enforceability of confidentiality agreements has implications for social welfare through its impact on the diffusion of $\mathcal{I}$'s knowledge. On the one hand, weak enforcement of NDAs decreases the incumbent's ($\mathcal{I}$) incentives to share knowledge with the supplier ($\mathcal{S}$) because the incumbent cannot protect its secrets from leakages to the competitor ($\mathcal{E}$). On the other hand, strong enforcement of NDAs might prevent the entrant ($\mathcal{E}$) from

getting access to $\mathcal{I}$'s secrets in situations in which it could have happened under weak enforcement.

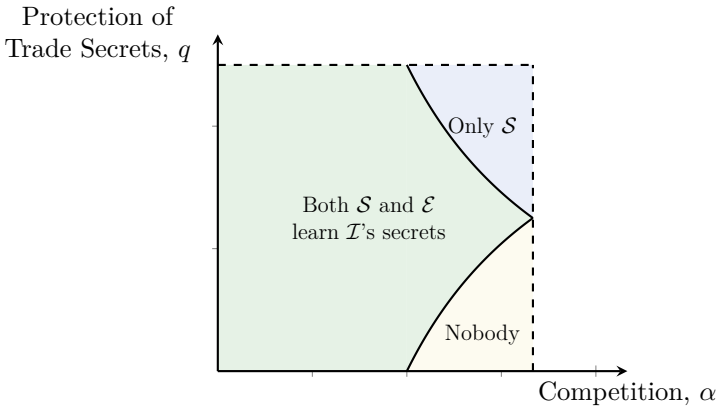


Figure 2.9: Equilibrium as a function of (α, q)

This figure illustrates patterns on knowledge flows based on the equilibrium from Proposition 2. Parameter α measures the degree of competition between $\mathcal{I}$ and $\mathcal{E}$, and parameter q measures the extent to which non-disclosure agreements are enforceable.

Figure 2.9 shows a graphical illustration of the equilibrium in Proposition 2 as a function of the competition (α) and trade secret protection (q). If competition between $\mathcal{I}$ and $\mathcal{E}$ is low, the enforcement of NDAs does not matter because $\mathcal{I}$ shares knowledge with $\mathcal{S}$ even without any contracts. If competition is high, then with low protection of trade secrets the incumbent $\mathcal{I}$ does not share its knowledge with $\mathcal{S}$, and this is the worst situation from the social point of view. With high protection, $\mathcal{I}$ shares its secrets with $\mathcal{S}$, but $\mathcal{S}$ does not spread it to the entrant $\mathcal{E}$. The intermediate level of protection in which $\mathcal{I}$ still shares knowledge $\mathcal{S}$ and $\mathcal{S}$ has incentives to diffuse it to $\mathcal{E}$ is the optimal one for the social welfare.

Summary. We consider a situation in which the incumbent firm ($\mathcal{I}$) decides whether to share its private knowledge with the supplier ($\mathcal{S}$). The problem is that the supplier might want to leak this knowledge to the incumbent's competitor ($\mathcal{E}$). $\mathcal{I}$ and $\mathcal{S}$ can sign a confidentiality agreement which has imperfect enforcement in courts. The litigation

happens under two conditions. First, the competition between $\mathcal{I}$ and $\mathcal{E}$ is high. Second, the enforcement of NDAs is neither strong nor weak, so $\mathcal{I}$ has incentives to share knowledge with $\mathcal{S}$, and $\mathcal{S}$ has incentives to leak it to $\mathcal{E}$.

The Defend Trade Secrets is likely to increase the protection of trade secrets (higher q). The welfare implications of this act depend on the degree of competition and the protection of trade secrets prior to the DTSA. If competition is high and the initial protection is weak, then the DTSA is likely to increase welfare by increasing firms' incentives to share knowledge. If competition is high and the initial protection is intermediate or strong, then the DTSA is likely to decrease welfare because it eliminates incentives to leak knowledge. In our framework, changes in the litigation activity following the DTSA provide a sufficient statistic for welfare evaluation: an increase (decrease) in the litigation activity indicates the positive (negative) net effect of the DTSA. These results should be interpreted with caution because the model misses many elements of the litigation activity, for example, the costs of lawsuits. However, the theory points out that changes in litigation activity might provide additional information about welfare consequences of intellectual property laws.

2.4 Conclusion

Secrecy is an important part of intellectual property protection toolkit. In this paper, we shed light on the prevalence of secrecy using data on the litigation of trade secrets' misappropriation. We identify different types of information protected by trade secrets, such as business and technological information, and mechanisms for the misappropriation of trade secrets, such as the mobility of employees. We document that the litigation of trade secrets has increased after the passage of the Defend Trade Secrets Act in 2016, and this increase was especially pronounced for business information and software technologies. We also document the variation in litigation activity across US states, industries, and firms based on their size. We develop a theory of knowledge sharing between firms and

show that the optimal social policy on trade secrets litigation depends on the degree of horizontal competition.

In the future, we plan to focus on the changes in litigation activity associated with the passage of the Defend Trade Secrets Act (DTSA). We will extend the theory and the empirical analysis of litigation activity to better understand the types of technologies and firms that were particular affected by the DTSA, and the welfare implications of this act.

Chapter 3

Global Sourcing and Multinational Activity: A Unified Approach[1]

3.1 Introduction

World trade flows are dominated by a small number of large firms – or 'superstar' exporters – that capture large market shares of their sector's exports (see Freund and Pierola, 2015; Gaubert and Itskhoki, 2018). For instance, in 2007, the top one percent of US exporters accounted for a staggering 79 percent of exports (Bernard *et al.*, 2018).[2] As significant

[1]Co-authored with Pol Antràs, Teresa Fort, and Felix Tintelnot. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the US Census Bureau or the Bureau of Economic Analysis. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this data product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release (CBDRB-FY21-CED006-0004). The results have also been reviewed by the Bureau of Economic Analysis to ensure that no confidential information has been disclosed. We thank Fabian Eckert and Natalia Ramondo for very constructive conference discussions, as well as seminar audiences at CESC in Barcelona, University of Minnesota, Stanford, Tokyo University, UC-San Diego, ASSA Meetings, Harvard-MIT, and Duke for useful questions and suggestions. We are also grateful to Emek Basker, Fariha Kamal and Jim Fetzer for valuable input on the data construction, Michael Freiman and Abdul Munasib for support with the disclosure process, and Hannah Rubinton for excellent research assistance.

[2]The analogous figure is 48 percent for Belgium, 44 percent for France, 59 percent for Germany, 77 percent for Hungary, 32 percent for Italy, 53 percent for Norway, and 42 percent for the United Kingdom (see Mayer and Ottaviano, 2007).

as their exporting activity is, the global involvement of these superstar firms goes well beyond the mere act of selling domestically produced goods to foreign consumers.

On the one hand, large exporters rely heavily on global sourcing (and thus foreign value added) in the production of their goods (Handley, Kamal and Monarch, 2020), and importing is also highly skewed, with the top one percent of importers accounting for 83.5 percent of US imports (Bernard *et al.*, 2018). In this paper, we calculate that in 2007 almost 95 percent of US goods exports were sold by firms that also import. On the other hand, and despite the fact that large firms dominate international trade flows, surveys conducted by various countries indicate that the very biggest firms often choose to produce their goods in foreign countries, rather than exporting them from their domestic base. This evidence is consistent with the well-known quantitative importance of multinational activity, some of which is explained by firms' incentives to save on transportation and trade costs when selling to foreign consumers. This form of horizontal or export-platform foreign direct investment (FDI) is a prominent feature of the data. We find that in 2007, exports by all US manufacturing firms accounted for slightly over one fourth of their foreign manufacturing sales.[3]

The incentives of firms, and especially large firms, to import intermediate inputs or to set up assembly plants abroad is well understood, but with very few exceptions that we review below, the literature tends to study each of these activities in isolation. Most work on horizontal or export platform FDI assumes that assembly only uses local factors of production, while most work on global sourcing or vertical FDI tends to assume that final goods are either nontradable or are perfectly tradable, with these extreme assumptions rendering export-platform strategies trivial. This pervasive dichotomy in the literature is

[3]Our aggregate matched data include 0.74 trillion USD exports of goods (Tables 3.1 and 3.2) and approximately $1.86 trillion USD foreign manufacturing sales by the majority-owned foreign affiliates (Tables 3.3 and 3.5) of all firms with US manufacturing establishments. The share of US exports in total foreign sales is even lower (about 24 percent) when including multinational enterprises outside manufacturing and when considering sales of services abroad (Antràs and Yeaple, 2014).

partly driven by theoretical considerations – the desire to isolate the determinants of either global sourcing or export-platform FDI – and partly by empirical considerations – the fact that available datasets have *not* contained detailed information on *both* the global sourcing strategies of firms as well as on their horizontal and export-platform FDI decisions abroad.

In this paper, we provide a unified framework to study firms' global sourcing and multinational production decisions. We develop a multi-country model in which firms choose not only the locations of their various assembly plants, as in the horizontal FDI and export-platform literatures, but also the countries from which all those plants import inputs, as in the global sourcing literature. Our framework identifies a natural complementarity between these two decisions and delivers novel implications for the role of geography and trade policy in shaping the global production strategies of firms. Empirically, we merge US Census domestic and trade data with the US Bureau of Economic Analysis (BEA) comprehensive surveys on multinational activity to document a series of novel facts regarding the global assembly and global sourcing strategies of US-based firms, and how they are interrelated.

We begin in section 3.2 by describing newly linked, comprehensive data on the domestic and foreign activities of all firms with US operations. We construct this new dataset by merging the Bureau of Economic Analysis (BEA) data on US-based firms' multinational activities to the universe of private, non-farm employer establishments in the United States provided by the US Census Bureau. Throughout the paper, we use the term "US-based MNEs" to describe all multinational firms with positive employment in the United States. We classify firms as multinationals if they are majority owners of affiliates in other countries (US MNEs), or if they are majority owned by a foreign firm (foreign-owned MNEs). We combine these data with detailed Customs trade transactions to construct a unified dataset on the domestic, export, import, and multinational activity of US-based firms. While numerous studies had exploited the BEA and Census data in isolation, ours is one of the

first projects to use both.[4]

The newly linked data provide the first detailed information on the full range of US-based MNEs' imports and exports. Access to the Customs trade transaction data is critical, since we find that 40 percent of US MNEs' imports are from arms' length suppliers, and BEA surveys only provide aggregated information on these trade flows. We use these data to analyze how MNEs' global sourcing decisions differ from those of firms that only sell domestically or that export but do not engage in multinational activity. A crucial feature of the data for this project is that they shed light on how US MNEs' global sourcing strategies are related to the geography of their foreign affiliate activity. Furthermore, since our dataset includes information from the BEA survey on US affiliates of foreign-based multinationals, we can also assess whether and how these firms' US affiliates orient their sourcing decisions towards their parent countries.

We first present a series of novel facts on the aggregate importance of both US and foreign MNEs in the US economy, especially in the manufacturing sector. In line with our emphasis on the dramatic size differences of MNEs, the data indicate that although US MNEs are fewer than 0.06 percent of all US firms, they employ 19 percent of workers, and account for 31 percent of aggregate sales, 36 percent of US imports and 48 percent of US exports. Similarly, foreign-owned MNEs' US affiliates account for only 0.17 percent of all firms, but employ 6 percent of the US labor force, and account for 13 percent of total sales, 33 percent of imports, and 24 percent of exports.

The disproportionate importance of MNEs is even more salient among manufacturers.[5] Firms engaged in multinational activity (either outward or inward) constitute only 1.5

[4]As described in section 3.2, merging the outward and inward MNE BEA datasets is itself challenging because a nontrivial number of firms appears in both datasets. To resolve this duplication issue, we use ownership and voting share information from the BEA data, along with the Census Bureau's Company Organization Survey (COS), to distinguish US MNEs from foreign-owned MNEs.

[5]We follow Fort, Pierce and Schott (2018) and classify all firms with one or more manufacturing plants as manufacturers. This ensures that our sample covers all US manufacturing activity.

percent of all US manufacturing firms, yet account for 41 percent of US manufacturing employment, as well as 74 percent of sales, 87 percent of imports, and 84 percent of exports among all manufacturers in the United States. We further show that these MNE manufacturers are not only large in the US, but also have very sizable operations abroad. More specifically, total foreign affiliate sales by US MNEs with foreign manufacturing are 74 percent of their total US establishments' sales. Furthermore, manufacturing sales to countries other than the US by these foreign affiliates are about four times larger than their corresponding US establishments' merchandise exports from the US, which highlights that foreign production is by far the most salient vehicle to make US-branded products available to foreign consumers. Taken together, these facts suggest that changes in the global sourcing and global assembly strategies of this small set of roughly 3,700 firms have the potential to cause sizable effects on the US manufacturing sector, and the US economy more broadly.

Moving beyond these descriptive statistics on the quantitative importance of multinational companies in the US manufacturing sector, in section 3.3 we present reduced-form evidence suggesting the existence of significant interdependencies in MNEs' global sourcing and global assembly strategies. The descriptive statistics already hint at the fact that global sourcing is more important for MNEs than for non-MNEs, and that among MNEs, foreign-owned firms import disproportionately more. While domestic importers' ratio of imports to sales is 6.0 percent, US MNEs import 11 percent of their total sales value, and foreign-owned firms import 14 percent. An interpretation of this finding is that global sourcing is more prevalent for firms whose center of gravity, in terms of their global assembly strategies, is further away from the United States.

We explore this hypothesis more formally via regressions that assess how the cross-country variation in the extensive and intensive margins of US global sourcing is related not only to standard gravity variables (such as distance, origin country GDP, common language and contiguity), but also to measures of the global operations (if any) of US

importers. To do so, we focus on the subset of manufacturing firms that import from two or more countries. These firms constitute just over half of manufacturing importers, but account for 99 percent of their imports. Even after controlling for firm and country fixed effects, we find that US MNEs are over 50 percentage points more likely to import from countries in which they have a majority-owned manufacturing affiliate, and foreign-owned MNEs are about 70 percentage points more likely to import from their headquarter country. These relationships are economically large: a doubling of GDP is associated only with a 0.02 percentage point increase in the probability of importing from a country. We similarly find an economically large and statistically significant relationship with the amount a firm imports from a country and its MNE activity. US MNEs imports are about 230 log points higher from countries in which they have affiliates, and foreign-owned firms import over 300 log points more from their home country. These strong relationships between firm-country-level imports and MNE activity are suggestive of significant complementarities between a firm's FDI and foreign sourcing decisions.

One interpretation of this novel finding is that performing one type of activity – global assembly or global sourcing – in a country lowers the additional fixed cost associated with the other activity. Alternatively, we hypothesize that activating one of these strategies increases the marginal benefit of activating the other one due to the benefits of locating assembly and suppliers in nearby locations. To provide additional evidence on the existence of this type of spatial correlation between global assembly and global sourcing, we estimate the relationship between a firm's country-level imports and the presence of MNE activity in the country's region, though *not* in the source country itself. We find that foreign-owned MNEs are 9 percentage points more likely to import from, and have 48 log points higher imports from, countries that are in the same region as their headquarters. We also find that the presence of an affiliate in the same region is associated with a 7 percentage point higher probability of importing from a country. By contrast, there is no statistically significant relationship between the amount a firm imports from the country and the presence of a

foreign affiliate in the region. These results thus provide new evidence that firms' global sourcing strategies are oriented towards those regions in which they have multinational activity, and that for US MNEs, this reorientation is driven solely by variation in their extensive margin import decisions.

Motivated by these empirical findings, in section 3.4 we develop a multi-country model in which a firm decides on both its global assembly strategy, namely the set of countries in which to set up assembly plants, as well as its unique global sourcing strategy, that is the set of countries from which all of its plants worldwide source inputs. The model constitutes a marriage of the Tintelnot (2017) model of export-platform of FDI and the global sourcing framework in Antràs, Fort and Tintelnot (2017). We model an economy in which consumers in J countries have nested constant-elasticity-of-substitution (CES) preferences over differentiated varieties produced by various firms in a single manufacturing sector. On the supply side, each firm produces a continuum of differentiated varieties in a sector characterized by monopolistic competition with free entry. Although firms face constant marginal costs, technology features increasing returns to scale due to the presence of several fixed costs of production: each firm pays an initial fixed cost of entry to obtain the blueprint for its bundle of differentiated consumer goods, and then incurs additional overhead costs to market its goods abroad, to set up assembly plants worldwide, and to activate specific countries as sources for its inputs. As in Melitz (2003), firms draw heterogeneous core productivity levels that are revealed upon entry, and which result in heterogeneous levels of participation in the global economy. More specifically, after entry, a firm faces three main decisions regarding the extensive margins of its operations. First, it decides whether to remain domestically oriented (at no further cost), or whether to pay a fixed cost to become a "global exporter," which allows the firm to sell goods in any market (we incorporate *destination-specific* selection into exporting in an extension). Second, it decides on the number and locations of the various assembly plants in which to manufacture goods (these locations serve as export platforms as in Tintelnot, 2017).

Third, it chooses the set of countries from which to source inputs (as in Antràs, Fort and Tintelnot, 2017).

Our theoretical framework shares many features with the model developed by Bernard *et al.* (2018) to organize their review of the export and import activities of firms producing in the US. A key novel feature of our framework is that a firm's sourcing strategy is a firm-level (rather than plant-level) 'asset', in the sense that all plants worldwide belonging to the same firm have access to the same sources of inputs. More specifically, if a US MNE has an affiliate in Japan that profitably sources inputs from China, the firm's US manufacturing plants will also have access to Chinese inputs at no additional fixed cost. We show that these economies of scale at the firm level in designing global sourcing strategies are crucial for the model to rationalize our empirical findings described above indicating that the global sourcing strategies of US importers are shaped by their global assembly operations.[6]

In order to simplify the characterization of the firm's intensive margin decisions, we follow Tintelnot (2017) and Antràs, Fort and Tintelnot (2017) in applying the probabilistic approach to productivity in Eaton and Kortum (2002). This produces simple gravity-style formulas for both firm-level bilateral shipments of consumer goods from any country where a firm assembles finished goods to all other countries in the world, as well as firm-level bilateral purchases of intermediate inputs from countries in a firm's sourcing strategy to each country in which that same firm assembles final goods.

The determination of a firm's extensive margin decisions is much more complex because it is shaped by various forces. Consider a firm that currently produces only in its home country, possibly exporting to other countries. Adding an assembly location entails an additional fixed cost, but also reduces the trade costs the firm will incur when selling not only to consumers in that new assembly location, but also to consumers in nearby

[6]Our baseline framework is admittedly less general than Bernard *et al.* (2018) in some dimensions (notably, we treat all firms as being non-atomistic).

locations. Furthermore, this new plant may impact the marginal benefit of setting up additional assembly plants in other countries via demand complementarities or via demand cannibalization effects. The extensive market decisions associated with the global assembly strategy of firms are thus interdependent across countries, but unlike in Tintelnot (2017), we allow for the possibility of the firm's profit function featuring not only decreasing but also increasing differences in these various extensive margin assembly decisions.

Next, consider the firm's global sourcing strategy. Adding a new source of inputs again entails a fixed cost, but it also reduces the marginal cost of obtaining inputs for the firm's assembly plant(s), thereby increasing the optimal scale of the firm. Whether extensive margin sourcing decisions are complements or substitutes depends on parameter values, much as the condition identified in Antràs, Fort and Tintelnot (2017). Finally, our model features an additional force that, in many (but not all cases) creates a complementarity between both types of extensive margin decisions. Intuitively, a richer sourcing strategy reduces marginal costs, increases optimal firm scale, and thus makes a richer assembly strategy more appealing (or its associated fixed costs easier to amortize). Similarly, richer assembly strategy increases overall firm sales and thus makes a more expansive sourcing strategy more appealing (or its associated fixed costs easier to amortize).

Overall, due to the various sources of substitutability and complementarity in the model, the problem of determining of a firm's extensive margin decisions constitutes a complex combinatorial optimization problem with $2^{J\times 2}$ possible choices, where J denotes the number of possible assembly and source countries. We identify a set of conditions under which the profit function features increasing differences in all extensive margin choices, which facilitates the derivation of sharp comparative statics and simple algorithms to solve the firm-level problem, as in Jia (2008) and Antràs, Fort and Tintelnot (2017). Outside that region of the parameter space, however, our framework does not feature the type of *single-crossing* properties that typically rationalize the use of iterative algorithms to reduce the dimensionality of the problem of solving for the firm's extensive margin (as

in Arkolakis, Eckert and Shi, 2021).

We conclude the paper with a specific case of our model that isolates the sources of interdependency that are novel to our framework, and we demonstrate that these forces provide a natural explanation for our reduced-form results. Furthermore, we provide two low-dimensional examples to illustrate the implications of the presence of firm-level economies of scale in global sourcing. In particular, we show that changes in tariffs on inputs and on final goods may lead to nonlinear responses in US manufacturing, in contrast to the results we obtain when global sourcing strategies are at the plant (rather than the firm) level.

Literature Review. Our paper contributes to the literature in two broad ways. First, we provide a unified framework to analyze global sourcing and multinational firm activity, while most previous studies examine each in isolation. Specifically, work on offshoring and global sourcing – such as Helpman (1984), Antràs and Helpman (2004), Grossman and Rossi-Hansberg (2008) and Antràs, Fort and Tintelnot (2017) – typically ignores the fact that firms face a nontrivial choice in their optimal location of assembly in the presence of trade costs in final goods. On the other hand, work on horizontal FDI and export-platform FDI – such Markusen (1984), Brainard (1997), Helpman, Melitz and Yeaple (2004), Arkolakis *et al.* (2018) or Tintelnot (2017) - typically ignores that the various assembly plants of multinational companies worldwide rely heavily on foreign inputs in their production processes. We are certainly not the first to incorporate intermediate input flows into models of horizontal and export-platform FDI (see, for instance, Helpman, 1985; Ramondo and Rodríguez-Clare, 2013), but previous frameworks did not incorporate an active extensive margin of global sourcing. While Yeaple (2003) and Grossman, Helpman and Szeidl (2006) explicitly discuss the existence of the type of complementarities between global sourcing and global assembly that feature prominently in our framework, they do so in models with at most three countries and two inputs, and do not study the empirical

relevance of their findings. In the survey on 'global firms', Bernard *et al.* (2018) develop a model featuring global assembly and global sourcing strategies, but their framework imposes assumptions that render assembly decisions independent across countries, and also focuses on the case in which global sourcing strategies are at the plant level.[7]

Our second main contribution is to advance the construction and analysis of a newly linked dataset that allows us to observe the global operations of firms based in the United States. We build on extensive work by Kamal, McCloskey and Ouyang (2021) to merge the BEA and Census datasets. Our focus on MNEs' trade behavior adds to a significant literature documenting the export and import decisions of US firms (e.g., as in Bernard and Jensen, 1999; Bernard *et al.*, 2007a, 2018; Antràs *et al.*, 2017) by showing how these decisions are shaped by companies' global operations. Prior work highlights the important role of multinational companies in US trade (see, for instance Bernard *et al.*, 2009), but our linked data allow us to provide the first complete portrait of how these decisions are related to MNEs' global assembly locations (*i.e.*, the countries and regions in which US MNEs locate foreign production), as well as the headquarter country and regions of foreign-owned firms with US operations.

The rest of the paper is structured as follows. In section 3.2, we introduce our novel, linked dataset and provide descriptive evidence supporting the quantitative relevance of global sourcing, global assembly, and their joint determination. In section 3.3, we provide reduced-form evidence suggestive of the interdependencies between the various margins of firm participation in global markets. In section 3.4, we present the assumptions of our model and solve for the equilibrium of the model for fixed global assembly and global sourcing strategies. The choice of these optimal strategies, and the challenges that such a

[7]The recent literature on global value chains (see Costinot, Vogel and Wang, 2013; Antràs and Chor, 2013; Antràs and De Gortari, 2020) also characterizes the optimal location of input production and assembly. Nevertheless, that work typically assumes environments with either no trade costs or with constant returns to scale, with these assumptions shutting down most of the interesting forces in our framework.

choice entails, are discussed in section 3.5. In section 3.6, we seek to isolate the key novel aspects of our theory and relate them to our empirical results in section 3.3. Section 3.7 concludes.

3.2 New Data and Facts on US Multinational Activity

In this section, we describe the newly linked, comprehensive data on US-based firms' domestic and foreign activities. We first describe all the data sources and how we combine them. We then present new facts on the aggregate importance of both US and foreign multinationals, in terms of total US sales, employment, and trade flows. We also quantify multinational firms' total imports and exports, as well as their share of these flows that take place within the firm. Finally, we describe the sample of firms on which we focus in this paper.

3.2.1 Data Description

An important contribution of this paper is to advance the construction and analysis of a new dataset that links the Bureau of Economic Analysis (BEA) data on US-based firms' multinational activities to the universe of private, non-farm employer establishments in the United States provided by the US Census Bureau. While numerous studies have used the BEA or the Census Bureau data separately, ours is one of the first projects to use both. A crucial benefit of the combined data is that they allow us to measure the full extent of US MNEs' import and export patterns, regardless of whether those transactions occur at arm's length or within the boundary of the firm. We are also able to measure the full range of all US-based MNEs' domestic activities, and to compare them to domestic-only firms. We focus on 2007 data in this project so that we can employ the most detailed

Census data, while avoiding contamination from the Great Recession.

We use the Census Bureau's 2007 Longitudinal Business Database (LBD), which is based on administrative tax records and provides employment and industry for every private, non-farm employer establishment in the United States. An establishment is a single physical location where business transactions take place and for which payroll and employment are recorded.[8] The LBD contains a firm identifier ($firmid$) that captures all of the establishments that are under common ownership or the control in a given year. We use $firmid$ to identify all establishments under the same firm ownership. This identifier is particularly important for US multinationals, that tend to have multiple establishments across a wide range of industries.[9]

We supplement the LBD with additional information from the 2007 Economic Censuses (ECs) of Manufacturing (CMF), Wholesale Trade (CWH), Retail Trade (CRT), Construction (CCN), Mining (CMI), Transportation, Communications, and Utilities (CUT), and Services (CSR). The censuses are conducted in years that end in "2" and "7" and provide information on establishment sales, value added, and input usage.

We also link the dataset to the 2007 Longitudinal Trade Transaction Database (LFTTD). The trade data are collected by US Customs and Border Protection and based on the universe of import and export transactions by US-based firms. They contain information on the products, values, and countries of firms' imports and exports. They also include an indicator for whether a transaction between the US importer or exporter takes place with a related party in the foreign country.[10] We match these data at the firm level to

[8]In practice, this information is reported at the employer identification number (EIN) level. The Census Bureau allocates this information across establishments using data for known multi-establishment firms from the Company Organization Survey (COS). New multi-unit firms are identified in the Economic Census data. See Jarmin and Miranda (2002) and Chow *et al.* (2021) for details on the LBD construction.

[9]The BEA data also contain firm identifiers. In practice, we find that the Census firmid includes more US activity than the BEA firm identifiers. This is likely due to the Census Bureau's access to the COS data that contain annual information on all multi-unit firms' establishments.

[10]For exports, related-party transactions are those in which one of the trading entities owns, directly

the LBD and EC data, and follow Antràs, Fort and Tintelnot (2017) in dropping mineral trade (HS2=27) so that we exclude trade in oil from the analysis. These data allow us to characterize the complete picture of MNEs' importing and exporting behavior.

We combine the Census data with the annual BEA outward and Benchmark BEA inward foreign direct investment survey data. The BE-11 survey provides annual information on US-based firms' outward foreign affiliate employment, local sales, sales back to the United States (and whether these are intra-firm), and sales to other markets, by affiliate country and industry. These data are collected for US-based firms that have 10 percent or greater ownership shares in foreign affiliates whose sales, assets, or net income are greater than $60 million. While these US-based firms are often US MNEs, foreign MNEs may locate their North American headquarters in the United States and thus report outward foreign affiliate activity.

We use data collected from the BE-12 survey to identify foreign-owned MNEs with inward activity in the United States. Since this is a benchmark survey, all foreign firms with a 10 percent or higher voting ownership interest in a US affiliate are required to file the BE-12 form. We build on extensive work by Kamal, McCloskey and Ouyang (2021) to match the BEA and Census datasets using Employer Identification Numbers (EINs), and by name and address. Additional details on the matching process are in Appendix Section C.2.1.

A novel contribution of our data construction effort is to distinguish US MNEs from foreign-owned firms. While most prior studies and published BEA data on outward FDI include both types of firms, our theoretical framework requires distinguishing these two forms of ownership. To do so, we use the ownership and voting share information from the BEA data, along with the Census Bureau's Company Organization Survey (COS),

---

or indirectly, 10 percent or more of the other entity. For imports, related-party transactions are those between members of the same family, shared officers or directors, partners, employers and employees, or a 5 percent controlling interest. See Bernard *et al.* (2007b) and Kamal and Ouyang (2020) for additional details on the LFTTD. The matched data cover about 80 percent of total exports and imports.

to separate US MNEs from foreign-owned MNEs. We provide details of our method in Appendix Section C.2.2.

3.2.2 Aggregate Importance of US-based MNEs

We first use the newly linked BEA-Census data to present statistics on the aggregate importance of globally-involved firms. To do so, we focus on four mutually-exclusive firm types: US multinational enterprises (US MNEs), foreign-owned multinationals (foreign-owned MNEs), US importers that are not multinationals (Importers), and all other domestic firms (Domestic). We define foreign-owned MNEs as those that are majority owned by a foreign entity. We define US MNEs as firms that have majority-owned foreign affiliate activity and are not foreign owned. We classify all remaining firms as domestic, and separate these into firms that do or do not import.

Table 3.1 presents aggregate information on the number of firms, workers, sales, and trade flows by firm type. The first column shows that although there are only about 2,800 US MNEs, which account for less than 0.06 percent of all US firms, they employ about 20 percent of workers and account for one third of aggregate sales. US MNEs are even more important in terms of trade, mediating one third of total imports and almost of half of all US exports.[11] Foreign-owned firms are also disproportionately involved in trade, accounting for six percent of aggregate employment, but one third of imports and a quarter of US exports. Finally, (non-MNE) US importers are also important in the aggregate. They account for about a quarter of US employment, sales, imports, and exports. In contrast, non-importers are about 50 percent of employment but account for only 5.0 percent of exports.

[11] We exclude trade in oil in these aggregates by dropping trade in HS2=27.

Table 3.1: Aggregate statistics for US-based firms in 2007, by firm import and MNE status

		Employment		Sales		Imports		Exports	
Firm Type	Firms	(000s)	Share	($B)	Share	($B)	Share	($B)	Share
Domestic	4,281,000	54,489	0.48	8,004	0.29	0	0.00	45	0.05
Importers	273,000	30,020	0.27	7,528	0.27	439	0.31	221	0.24
Foreign-Owned	7,600	6,964	0.06	3,764	0.13	478	0.33	224	0.24
US MNEs	2,800	21,666	0.19	8,655	0.31	518	0.36	446	0.48
Total	4,564,400	113,139	1.00	27,951	1.00	1,435	1.00	937	1.01

Notes: Table presents levels of firms, employment, sales, imports, and exports, for US firms in 2007. Sample includes all private, non-farm, employer establishments. "Domestic" firms are those that are not multinationals and do not import. "Importers" are firms that import but are not multinationals. "Foreign-owned" are firms that are majority-owned by a foreign firm. US MNEs are firms that are majority-owned by a US firm and that have majority-owned foreign affiliate activity. Observations rounded per Census disclosure rules.

3.2.3 Definition of Our Sample

In this paper, we focus on the subset of firms with one or more manufacturing establishments located in the United States so that we cover all US manufacturing activity. We provide context on this sample by separating the aggregate totals presented in Table 3.2 into firms with domestic manufacturing establishments (Panel A) and firms without them (Panel B). This distinction for multinationals is new to this paper, since the comprehensive domestic establishment-level data allow us to identify multinationals with some US manufacturing activity, even if that is not the firm's primary industry. For US MNEs, we further distinguish those that have majority-owned foreign manufacturing affiliates from those that do not. This allows us to analyze the extent and aggregate importance of firms that may have shed (or never had) domestic manufacturing and replaced it with foreign affiliates.[12]

Table 3.2 shows that the majority of US MNEs have domestic manufacturing activity. These MNEs account for 10 percent of total employment and 30 percent of manufacturing employment. US MNEs that manufacture domestically are vastly more important in terms

[12]The vast majority of foreign-owned US-based firms with US manufacturing activity also report foreign manufacturing activity, so we do not separate these for disclosure avoidance purposes.

of trade flows. They account for over 30 percent of US imports and 45 percent of exports, whereas US MNEs without domestic manufacturing account for only 5 and 2 percent, respectively. Foreign-owned firms with US manufacturing also trade disproportionately more. They account for 26 percent of imports and 21 percent of exports, compared to only 7 and 3 percent for foreign MNEs that do not manufacture. We can also infer from Table 3.2 that US-based manufacturing account for a striking 87 percent of US manufacturer's imports and 83 percent of US manufacturer's imports. An interesting difference is evident for domestic importers. While non-manufacturing importers account for 22 percent of imports, manufacturing importers mediate just 9 percent. These domestic importers' export shares are comparable, at just over 10 percent each of the aggregate.

We draw three main conclusions from Table 3.2. First, our sample of firms (Panel A), accounts for two-thirds of US imports and 80 percent of US exports. Second, the majority of US MNEs that manufacture domestically also have majority-owned foreign manufacturing affiliates. These 1,200 firms (last detailed row of Panel A) comprise our set of US MNEs. Third, it is quite rare for firms to produce goods in-house overseas without also having domestic manufacturing. The last detailed row of Panel B shows that there are only 150 US MNEs with foreign manufacturing affiliates but no US manufacturing plants, and these firms account for less than one percent of US employment and exports, and only 1.0 percent of imports. Our sample is thus particularly relevant for understanding global production and trade patterns, and the focus on US manufacturers does not appear to omit firms that account for significant production in their own foreign facilities.

We also calculate the aggregate importance of US-based MNEs' foreign affiliate activity, and the extent to which our sample of firms captures these flows. Table 3.3 presents total sales by US-based MNEs, which are equal to the sum of all sales by these firms' domestic establishments and majority-owned foreign affiliates. Columns 1 to 4 present the levels and share of the aggregate total of these firms' total worldwide and foreign

Table 3.2: Aggregate statistics for firms in 2007, by firm manufacturing, import, and MNE type

		Share of Total				
Firm Type	Firms	Emp	Man Emp	Sales	Imports	Exports
Panel A: Manufacturing Firms						
Domestic	182,000	0.02	0.19	0.02	0.00	0.01
Importers	60,000	0.07	0.40	0.08	0.09	0.12
Foreign-Owned	2,200	0.03	0.12	0.10	0.26	0.21
US MNEs						
w/o Manuf Aff	350	0.04	0.03	0.05	0.03	0.02
w/Manuf Aff	1,200	0.06	0.27	0.14	0.29	0.43
Manufacturers' Total	245,750	0.22	1.01	0.39	0.67	0.79
Panel B: Non-Manufacturing Firms						
Domestic	4,099,000	0.46		0.27	0.00	0.04
Importers	213,000	0.19		0.19	0.22	0.11
Foreign-Owned	5,400	0.03		0.04	0.07	0.03
US MNEs						
w/o Manuf Aff	1,100	0.09		0.11	0.04	0.02
w/Manuf Aff	150	0.00		0.01	0.01	0.00
Non-Manuf Total	4,318,650	0.77	0.00	0.62	0.34	0.20

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward, and BEA outward datasets. Table presents levels of firms and shares of employment, sales, imports, and exports, for all US private, non-farm, employer establishments in 2007. "Foreign-owned" are firms that are majority-owned by a foreign firm. US MNEs are non-foreign-owned firms with majority-owned foreign affiliates. "Domestic" firms are non-multinationals that do not import. "Importers" are non-multinationals that import. Top panel presents the number of firms and the shares of total US employment, manufacturing employment, sales, imports, and exports accounted for by firms with manufacturing establishments in the United States in 2007. Total US manufacturing employment for these firms is 13.1 million in 2007. Bottom panel presents comparable statistics for firms without US manufacturing establishments. Observations rounded per Census disclosure rules. Columns to 6 sum to 1.00 except due to rounding.

affiliate sales. Column 4 show that MNEs with US manufacturing establishments account for 80 percent of all US-based MNEs' foreign sales. Column 5 presents the ratio of MNEs' foreign affiliate sales to US establishment sales, by firm type. Total foreign affiliate sales by US MNEs with foreign manufacturing sales are equal to a striking 74 percent of their domestic establishment sales. This figure highlights the relevance of these firms' foreign assembly locations. This statistic is much smaller for foreign-owned MNEs, at only 0.31. It is important to note, however, that the BEA data only collects the foreign affiliate sales

of these firms' foreign affiliates that are majority-owned by their US operations. Finally, in column 6 we calculate the share foreign affiliate sales that is in manufacturing by each firm type. 60 percent of US MNEs' affiliate sales are of manufactured goods. For the remainder of the paper, we use these affiliates and their sales as our measure of US MNEs' foreign assembly locations and operations.

Table 3.3: US-Based MNEs' foreign affiliate activity, by firm manufacturing and MNE type

	Worldwide Sales		Affiliate Sales		$\frac{\text{Aff Sales}}{\text{US Estab Sales}}$	$\frac{\text{Manuf Aff Sales}}{\text{Aff Sales}}$
	($B)	Share	($B)	Share		
Panel A: Manufacturing Firms						
Foreign-Owned US MNEs	2,702	0.22	839	0.17	0.31	0.43
w/o Manuf Aff	1,446	0.12	249	0.05	0.17	0.00
w/Manuf Aff	3,853	0.31	2,857	0.58	0.74	0.60
MNE Manufacturers' Total	8,001	0.64	3,945	0.80	0.49	0.53
Panel B: Non-Manufacturing Firms						
Foreign-Owned US MNEs	1,062	0.09	24	0.00	0.02	0.04
w/o Manuf Aff	3,183	0.26	809	0.16	0.25	0.00
w/Manuf Aff	173	0.01	172	0.03	0.99	0.14
MNE Non-Manufacturers' Total	4,418	0.36	1,005	0.20	0.23	0.02

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward, and BEA outward datasets. Columns 1-4 present levels of firms' total worldwide sales (sales by US establishments and foreign affiliates) and affiliate sales, and the share of these aggregates accounted for by each MNE firm type. Column 5 presents $^{\text{Aff Sales}}/_{\text{US Estab Sales}}$, which is the ratio firms' foreign affiliate sales over their US establishments' sales. Column 6 presents $^{\text{Manuf Aff Sales}}/_{\text{Aff Sales}}$, which is the share of firms' total manufacturing affiliate sales over their total affiliate sales. Only sales by the firms' majority-owned foreign affiliates are included in these calculations. "Foreign-owned" are firms that are majority-owned by a foreign firm. US MNEs are non-foreign-owned firms with majority-owned foreign affiliates. Top panel statistics for firms with manufacturing establishments in the United States in 2007. Bottom panel presents comparable statistics for firms without US manufacturing establishments.

3.2.4 Aggregate Import and Export Statistics

We also exploit the newly linked data to provide aggregate statistics on the extensive and intensive margins of manufacturing firms' foreign trade decisions by firm type. Table 3.4 reveals the relative importance and propensity of firms' trade engagement, both for total imports and exports, and for related-party versus arms-length transactions, by firm

type. Column 1 in Panel A reports the share of total imports over sales and shows that domestic importers are the least import intensive, with a 6.0 percent value.[13] Imports are almost twice as important for US MNEs, whose imports-to-sales ratio is 11 percent, and foreign-owned MNEs are the most intensive with a ratio of 14 percent. Column 2 reports the share of each firm type that imports. This share is 1.0 for domestic importers (by definition), and equal to 25 percent for all non-MNEs. The overwhelming majority of multinational firms import. In terms of related-party trade imports, 29 percent of domestic firms' imports are with related parties, while US MNEs and foreign-owned MNEs source 61 and 79 percent of their imports from related parties, respectively.[14]

Panel B of Table 3.4 reports comparable statistics for exporting. The patterns are quite similar to those for imports, with two notable exceptions. First, US MNEs are the most export-intensive of all firm types, with 10 percent of their domestic establishments' sales shipped abroad. Second, related-party trade shares are considerably lower for exports than for imports: both US and foreign-owned MNEs sell about 40 of their exports to related parties. These shares are approximately 34 percent less than their related-party import shares for US MNEs and 47 percent less for foreign-owned MNEs. Finally, we note that exporting is more common than importing. Combining information from Tables 3.2 and 3.4, we calculate that there are 4,600 more exporters (about 2 percent of total US manufacturers) than importers.

[13]We exclude US MNEs without foreign manufacturing affiliates from this comparison since they are few in number and we will classify them as "Domestic Importers" for the remainder of this paper. As discussed in Section 3.2.3, this is because we define a US MNE as firm with majority-owned foreign manufacturing affiliate activity that is not majority foreign owned.

[14]The ownership threshold for related-party trade is only 5 percent, thus making it possible for domestic importers to have related-party transactions. Combining information from Tables 3.2 and 3.4, we note that non-MNE manufacturing firms account for only 6.2 percent of the total manufacturers' related-party imports and 7.0 percent of total related-party exports.

Table 3.4: Trade statistics for manufacturing firms in 2007, by import and MNE status

Panel A: Imports	$\frac{Imports}{Sales}$	$\frac{Importers}{Firms}$	$\frac{RPImporters}{Importers}$	$\frac{RPImports}{Imports}$
Importers	0.06	1.00	0.20	0.29
Foreign-Owned	0.14	0.91	0.85	0.79
US MNEs				
w/o Manuf Aff	0.03	1.00	0.86	0.17
w/Manuf Aff	0.11	0.92	0.91	0.61
Manufacturers' Total	0.09	0.26	0.24	0.62
Panel B: Exports	$\frac{Exports}{Sales}$	$\frac{Exporters}{Firms}$	$\frac{RPExporters}{Exporters}$	$\frac{RPExports}{Exports}$
Domestic	0.02	0.14	0.06	0.10
Importers	0.05	0.64	0.22	0.16
Foreign-Owned	0.07	0.91	0.70	0.42
US MNEs				
w/o Manuf Aff	0.01	1.00	0.86	0.24
w/Manuf Aff	0.10	1.00	0.92	0.40
Manufacturers' Total	0.07	0.28	0.19	0.36

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward, and BEA outward datasets. Table presents the ratios of firm imports to sales and exports to sales, the share of each firm type that imports and exports, the share of related-party importers and exporters, and the share of related-party imports and exports. "Foreign-owned" are firms that are majority-owned by a foreign firm. "US MNEs" are non-foreign-owned firms with majority-owned foreign affiliates. "Domestic" firms are non-multinationals that do not import. "Importers" are non-multinationals that import.

3.2.5 Foreign Manufacturing Affiliate Statistics

We conclude this section with statistics on the foreign affiliate sales in manufacturing by US MNEs with domestic manufacturing establishments. US MNEs have majority-owned foreign manufacturing affiliates in an average of 6.42 countries. In Table 3.5, we present the weighted average of these firms' total manufacturing affiliate sales, as well as their average sales across destinations. Column 2 shows that the majority of foreign affiliate sales are local, with 54 percent of US MNEs' sales remaining in the host country. Export

platform sales are also significant, with 35 percent of US MNEs' affiliate sales shipped to 'third' markets (i.e., markets other than the host country and the US). Only 11 percent of affiliate sales return to the United States, with the vast majority of these (over 80 percent) shipped to affiliated US parties. This is in line with results in Ramondo, Rappoport and Ruhl (2016), who find that sales back to the US are not a dominant feature of affiliates' activity.[15]

Table 3.5: Foreign affiliate manufacturing sales for US MNEs in 2007, by destination

	Total	Local	Third Markets	US Third Parties	US Intra-Firm
Firm Average	1,458	782.5	506.1	31.06	138
Share	1.00	0.54	0.35	0.02	0.09

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward, and BEA outward datasets. Table presents the weighted firm-level average of foreign affiliates' manufacturing sales in millions USD by destination for US MNEs with US manufacturing establishments. "US MNEs" are non-foreign-owned firms with majority-owned foreign manufacturing affiliates.

3.3 Global Sourcing and MNE Status: Reduced-Form Evidence

We use the newly linked data to analyze the foreign sourcing and exporting patterns of multinational enterprises (MNEs), and to assess how they differ from those of domestic importers and exporters. We first provide summary statistics on the number of import

[15]We note that these statistics include firms with zero reported flows. If some of these zeros include missing flows, our estimates may be biased. For instance, our results in Tables 3.1, 3.2, and 3.4 imply that US MNEs import 253.9 billion USD within their firm boundaries, which corresponds to an average of 211 million USD per firm. This figure clearly exceeds the 138 million USD figure reported in the last column of Table 3.5. Specifically, the (roughly) 1,200 US MNEs with domestic and foreign manufacturing account for 29 percent of US merchandise imports (see Table 3.2), and 61 percent of these imports are related-party imports (see Table 3.4). Thus, their total imports are 1,435 billion US $\times$ 0.29 $\times$ 0.61, or 253.9 billion USD, and each of these 1,200 MNEs imports an average of 211 million USD within their firm boundaries.

and export countries by firms' MNE status. We then analyze how US MNEs' foreign affiliate activity, and foreign-owned firms' headquarter locations, relate to their global sourcing decisions. Unlike previous studies that analyze the trading patterns of MNEs using their foreign affiliate shipments, we use the Census Bureau's LFTTD, which is based on US Customs transactions. These data provide the first complete portrait of the global sourcing strategies of US MNEs with majority-owned affiliates.

3.3.1 Import and Export Behavior by Firm Type

As described in Section 3.2.3, the sample of firms in this paper is based on the US Census Bureau's LBD and consists of all firms with one or more manufacturing establishments in the United States in 2007. For the remainder of the paper, we use the BEA survey data to classify manufacturing firms into three mutually-exclusive firm types: domestic firms, foreign-owned MNEs, and US MNEs. As before, we classify a firm as foreign-owned if its US establishments are majority owned by a foreign firm. We now classify a firm as a US MNE only if it has majority-owned *manufacturing* affiliate activity in a foreign country, and is not foreign-owned. US MNEs with majority-owned non-manufacturing affiliate activity, which we show in Tables 3.2 and 3.4 are few in number and account for small shares of aggregate activity, are therefore classified as domestic firms in the remainder of this paper.

A central goal of this paper is to document and analyze the foreign sourcing strategies of MNEs and how they compare to domestic importers. We therefore distinguish firms that import from just one country from those that import from multiple countries. The top panel in Table 3.6 presents information on the share of imports and importers by firm MNE status, and by single versus multiple country importers. Among domestic firms, it is common for firms to import from a single country, with half of domestic importers sourcing from a single location. In contrast, effectively all MNEs import from multiple

countries. Despite the large share of firms that imports from a single country, the second column in Table 3.6 shows that these single-country importers account for only 1.0 percent of manufacturers' imports.

The last two columns in Table 3.6 report the average and median number of countries from which multi-country importers source. Domestic manufacturers import from an average of 4 countries, while the median number is 3. Foreign-owed firms import from an average of 12 countries and median of 8 countries. Finally, US MNEs have the most expansive sourcing strategies, importing from an average of 21 and median of 17 foreign countries.[16]

We also analyze how firms' export patterns differ by the MNE status. The bottom panel of Table 3.6 presents comparable statistics for firms' export behavior. Multi-country exporters account for 99 percent of US manufacturers' exports. The extensive margin of exporting is generally larger than the import margin, though also more skewed. Domestic exporters sell to an average of 8 countries, twice their median of 4. Foreign MNEs export to an average of 19 countries and a median of 10. As for imports, US MNEs sell to the largest number of countries, with an average of 40 and a median of 35.

3.3.2 Interactions Between FDI and Importing

In this section, we analyze how firms' foreign ownership and manufacturing affiliate activity relate to their import behavior. To do so, we estimate the following linear probability model of a firm's extensive margin import decisions:

[16]Census disclosure avoidance rules preclude us from disclosing the true median. We therefore calculate a fuzzy median equal to the average number of countries for firms in the 49th to the 51st percentiles.

Table 3.6: Import and export patterns by MNE status for US manufacturing firms

Panel A: Import Patterns					
Firm Type	No. of Import Countries	Share of Importers	Share of Imports	Number of Import Countries Average	Median
Domestic	1	0.47	0.01	1	1
	2+	0.48	0.17	4	3
Foreign-Owned MNEs	1	0.00	0.00	1	1
	2+	0.03	0.40	12	8
US MNEs	1	0.00	0.00	1	1
	2+	0.02	0.43	21	17
Panel B: Export Patterns					
Firm Type	No. of Export Countries	Share of Exporters	Share of Exports	Number of Export Countries Average	Median
Domestic	1	0.44	0.01	1	1
	2+	0.52	0.18	8	4
Foreign-Owned MNEs	1	0.00	0.00	1	1
	2+	0.03	0.27	19	10
US MNEs	1	0.00	0.00	1	1
	2+	0.02	0.54	40	35

Source: 2007 Economic censuses, LBD, LFTTD, and BEA inward and outward datasets. Top panel presents the share of US importers and import value, and the average and median number of import countries for all firms with manufacturing establishments in the United States in 2007 by firm MNE status. Bottom panel presents comparable statistics for US exports. Foreign-owned consists of firms that are majority-owned by a foreign firm. US MNEs consist of firms that have majority-owned foreign affiliate manufacturing activity and are not a foreign-owned. Domestic consists of all other firms. The data in this table are limited to countries for which gravity variable data are available, and from which multiple US firms import and export. This ensures that the samples of firms in this table match the samples used in the regression analysis in the next section.

$$\begin{aligned} Pr(I_{frc} = 1|X) = \quad & \beta_D log(distance_c) + log(GDP_c) + \beta_L Language_c + \beta_C Contiguous_c + \\ & \beta_A Affiliate_{fc} + \beta_{AR} AffiliateRegion_{fr} + \\ & \beta_F Foreign_{fc} + \beta_{FR} ForeignRegion_{fr} + \gamma_f + \gamma_r, \end{aligned} \tag{3.3.1}$$

where I_{frc} is an indicator equal to one if firm f imports from country c, that belongs to region r. The first row of equation (3.3.1) includes the following standard gravity variables: $log(distance_c)$, the distance from the United States to country c, $Language_c$, an indicator

for whether country c speaks English, and $Contiguous_c$, an indicator for whether country c is contiguous to the United States (i.e., Mexico or Canada).[17] The variables in the second row capture new information on firms' MNE status. $Affiliate_{fc}$ is an indicator for whether the firm has a majority-owned manufacturing affiliate in country c. $AffiliateRegion_{fr}$ is an indicator for whether the firm has a majority-owned manufacturing affiliate in the same region as country c, though not country c itself.[18] $Foreign_{fc}$ is an indicator for whether the firm is majority owned by a firm in country c, and $ForeignRegion_{fr}$ is an indicator for whether it is owned by a firm in the same region as country c, though not country c itself.

A primary goal of this analysis is to document whether and how the geography of firms' MNE activity influences their foreign sourcing behavior. We therefore include firm fixed effects and limit the regression sample to firms that import from two or more countries to avoid incorrect inference (e.g., see Correia, 2015). The firm fixed effects control for all unobservable firm characteristics that are constant across countries, and the limitation to multi-country importers makes the comparison to domestic importers more similar. As noted above, this sample covers about 99 percent of the value of US-based manufacturers' imports. We also estimate equation (3.3.1) either with region fixed effects, or with country fixed effects, in which case we focus exclusively on the firm-by-country variation from the affiliate and foreign-owned indicators. We cluster the standard errors by country and by firm.

Table 3.7 presents the results from estimating equation (3.3.1) via ordinary least squares (OLS) on all firms with US manufacturing establishments that import from multiple countries. Columns 1-4 include firm and region fixed effects, while columns 5 and

[17] We download these gravity variables from the CEPII (Centre d'Etudes Prospectives et d'Informations Internationales) gravity dataset. See Head *et al.* (2010) for details on the construction of these variables.

[18] We define the following regions: Africa, Central Asia, East Asia, Europe (excluding the New Member States), Middle East, New Member States of the European Union, North America, OWH, Oceania, South and Central America, South Asia, Southeast Asia, and Western Asia.

6 include firm and country fixed effects. The first two columns present estimates from specifications with standard gravity variables, where we add an indicator for contiguity in column 2.[19] As expected, firms are more likely to import from richer countries and from contiguous ones. In column 3, we add indicators to capture firms' multinational activity in a particular country. The estimates suggest that firms are 55 percentage points more likely to import from a country in which they have a majority-owned foreign manufacturing affiliate, and 73 percentage points more likely to import from their home country. In column 4, we also include regional indicators, which show that firms are 7 percentage points more likely to import from a country if they have an affiliate in the region, and 8.6 points more likely to import from the same region as their home country. These estimates are economically large, since the average multi-country importers sources from only 2.8 percent of the 182 countries in the sample. Columns 5 and 6 show that these relationships persist even when controlling for firm and country fixed effects.

We also examine how the intensive margin of imports relates to firms' MNE activity across countries. To do so, we estimate:

$$\begin{aligned} y_{frc} = \quad & \beta_D log(distance_c) + log(GDP_c) + \beta_L Language_c + \beta_C Contiguous_c + \\ & \beta_A Affiliate_{fc} + \beta_{AR} AffiliateRegion_{fr} + \\ & \beta_F Foreign_{fc} + \beta_{FR} ForeignRegion_{fr} + \gamma_f + \gamma_r + \varepsilon_{frc}, \end{aligned} \tag{3.3.2}$$

where y_{frc} is the log of firm f imports from country c, that belongs to region r.

Table 3.8 reports the results from estimating equation (3.3.2) via OLS on all firms with US manufacturing establishments that import from multiple countries. As before,

[19] Our region definitions designate Canada to North America and Mexico to South and Central America. The contiguity dummy with region fixed effects thus identifies Mexico. In additional specifications available upon request, we note that the coefficient on the contiguity dummy is approximately double when we exclude the region dummies. We focus on specifications with the region controls since the intensive margin of imports is strongly increasing in $log(distance_c)$ without them once we control for contiguity.

Table 3.7: Extensive margin import regressions

Dependent variable is $importer_{frc} = 1$ if firm imports from country c

	(1)	(2)	(3)	(4)	(5)	(6)
Common Language$_c$	0.002	0.001	0.001	0.001		
	(0.008)	(0.008)	(0.008)	(0.008)		
$log(distance_c)$	-0.017	-0.006	-0.006	-0.006		
	(0.013)	(0.009)	(0.009)	(0.009)		
$log(GDP_c)$	0.015***	0.014***	0.014***	0.014***		
	(0.003)	(0.003)	(0.002)	(0.002)		
Contiguous$_c$		0.133***	0.128***	0.129***		
		(0.013)	(0.013)	(0.013)		
Affilliate$_{fc}$			0.550***	0.582***	0.501***	0.536***
			(0.028)	(0.031)	(0.025)	(0.028)
Foreign-Owned$_{fc}$			0.726***	0.735***	0.669***	0.678***
			(0.046)	(0.047)	(0.047)	(0.047)
Affiliate in Region$_{fr}$				0.069***		0.074***
				(0.015)		(0.015)
Foreign in Region$_{fr}$				0.086***		0.090***
				(0.020)		(0.021)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Region FEs	Yes	Yes	Yes	Yes	No	No
Country FEs	No	No	No	No	Yes	Yes
Adj. R2	0.194	0.197	0.215	0.216	0.278	0.28
Observations (000s)	6,330	6,330	6,330	6,330	6,330	6,330

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward and outward datasets, and CEPII. Dependent variable is an indicator for whether firm f imports from country c in region r. Sample is all firms with manufacturing establishments in the United States in 2007 that import from multiple countries. Observations in 1000s and rounded per Census disclosure rules. The mean and standard deviation of the dependent variable are 0.028 and 0.165, respectively. There are 182 countries in this sample. Standard errors two-way clustered by firm and by country. *,**, *** denote p<0.10, p<0.05, and p<0.01, respectively.

the specifications reported in columns 1 to 4 include firm and region fixed effects, while columns 5 and 6 include firm and country fixed effects. Columns 1 and 2 report coefficient estimates for standard gravity variables. While the column 1 suggests that firm-level imports decrease in distance, the estimates in column 2 suggest that this is driven by trade with Mexico, since contiguity has a large and statistically significant coefficient, and

its inclusion in the specification reduces the size of the distance coefficient dramatically and renders it statistically insignificant. Column 3 shows that firms' imports are 227 log points higher from countries in which they have a foreign affiliate, and 340 log points higher from their home country. The starkest difference between firms' extensive and intensive margin import decisions is evident in column 5: while foreign-owned MNEs import about 47 percent more from countries in the same region as their home country, there is no statistically significant relationship between the amount US MNEs import from a country and the presence of a foreign affiliate in the same region. Columns 5 and 6 present comparable estimates when controlling for firm fixed effects.

The results in this section provide new evidence that firms' MNE activity is strongly related to their import behavior. The estimates indicate that although the amount a US MNE sources from a country is not higher if it has an affiliate in the same region, the likelihood that it will import from a country is about 9 percentage points higher if it has an affiliate elsewhere in the region. This estimate is over three times the size of the average share of countries from which a firm in the sample imports. These results are consistent with the premise that a firm's global sourcing strategy – *i.e.*, the set of countries from which it purchases inputs – is influenced by the geography of its foreign production locations.[20] In the next section, we develop a new theoretical framework that accommodates the potential for firms' foreign production and sourcing location decisions to depend upon each other.

[20]In Appendix Section C.2.6 we show that both the intensive *and* extensive margins of firms' exports are larger from both countries and regions in which US MNEs have foreign affiliates. The responsiveness of the intensive margin of exports but not imports to an affiliate in a country is thus a unique feature of firms' sourcing decisions.

Table 3.8: Intensive margin import regressions

Dependent variable is $log(imports_{frc})$

	(1)	(2)	(3)	(4)	(5)	(6)
Common Language$_c$	-0.264***	-0.252**	-0.272**	-0.269**		
	(0.101)	(0.110)	(0.113)	(0.113)		
$log(distance_c)$	-0.719***	-0.157	-0.105	-0.107		
	(0.191)	(0.347)	(0.386)	(0.385)		
$log(GDP_c)$	0.392***	0.377***	0.326***	0.331***		
	(0.050)	(0.054)	(0.058)	(0.058)		
Contiguous$_c$		0.874**	0.898**	0.885**		
		(0.378)	(0.411)	(0.411)		
Affilliate$_{fc}$			2.265***	2.363***	2.224***	2.331***
			(0.127)	(0.112)	(0.123)	(0.110)
Foreign-Owned$_{fc}$			3.399***	3.545***	3.617***	3.765***
			(0.165)	(0.177)	(0.227)	(0.223)
Affiliate in Region$_{fr}$				0.162		0.181
				(0.115)		(0.113)
Foreign in Region$_{fr}$				0.468***		0.480***
				(0.156)		(0.160)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Region FEs	Yes	Yes	Yes	Yes	No	No
Country FEs	No	No	No	No	Yes	Yes
Adj. R2	0.233	0.234	0.268	0.269	0.282	0.283
Observations (000s)	177	177	177	177	177	177

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward and outward datasets, and CEPII. Dependent variable is the log of imports by firm f, from country c, in region r. Sample is all firms with manufacturing establishments in the United States in 2007 that import from multiple countries. Observations in 1000s and rounded per Census disclosure rules. Standard errors two-way clustered by firm and by country. The mean and standard deviation of the dependent variable are 4.198 and 2.606, respectively. There are 182 countries in this sample. *,**, *** denote p<0.10, p<0.05, and p<0.01, respectively.

3.4 Theoretical Framework

Motivated by our empirical results, in this section we develop a new framework to analyze the effects of firms' joint decisions on the countries in which to locate their production establishments and the countries from which to source their inputs.

3.4.1 Environment

Consider a world in which individuals in J countries value the consumption of differentiated varieties of manufacturing goods as well as the consumption of the output of a non-manufacturing sector. Consumers worldwide spend a constant share η of their income on manufacturing goods. Individuals supply one unit of labor inelastically, with L_i denoting the total labor force in country $i \in J$ (with some abuse of notation, we denote by J both the number as well as the set of countries). There are no other factors of production, so labor should be interpreted as "equipped" labor. The non-manufacturing sector is perfectly competitive and operates under a constant-returns-to scale technology in labor. We assume that the non-manufacturing sector is freely tradable and large enough to pin down wages (denoted by w_i in country i) in the manufacturing sector.

There is an endogenous measure Ω_i of manufacturing firms selling goods in country i. As in Tintelnot (2017), each of these firms produces and sells a continuum of measure one of varieties of manufactured goods. We index firms by φ and varieties within firms by ω. We assume a nested-CES structure in which the degree of substitutability σ across varieties produced by different firms and the degree of substitutability σ_w across varieties produced by the same firm may differ from each other, or

$$U_{Mi} = \left(\int\limits_{\varphi \in \Omega_i} \left(\int_0^1 q_i\left(\varphi, \omega\right)^{(\sigma_w - 1)/\sigma_w} \mathrm{d}\omega \right)^{\frac{\sigma_w}{\sigma_w - 1}\frac{(\sigma-1)}{\sigma}} \mathrm{d}\varphi \right)^{\sigma/(\sigma-1)}, \quad \sigma_w, \sigma > 1. \tag{3.4.1}$$

These preferences imply that consumers in country i spend a share

$$s_i(\varphi) = \left(\frac{p_i(\varphi)}{P_i} \right)^{1-\sigma} E_j \tag{3.4.2}$$

of their income on firm φ. In this expression, E_i is total spending on manufacturing goods

in country $i \in J$,

$$p_i(\varphi) = \left(\int_0^1 p_i(\varphi, \omega)^{1-\rho} \mathrm{d}v \right)^{\frac{1}{1-\rho}} \tag{3.4.3}$$

is the overall price index for varieties sold by firm φ, and

$$P_i = \left(\int_{\varphi \in \Omega_i} p_i(\varphi)^{1-\sigma} \mathrm{d}\varphi \right)^{\frac{1}{1-\sigma}} \tag{3.4.4}$$

is the economy-wide ideal price index in country i.

Demand in country i for each individual variety ω produced by firm φ is given by

$$q_i(\varphi, \omega) = \left(\frac{p_i(\varphi, \omega)}{p_i(\varphi)} \right)^{1-\sigma_w} s_i(\varphi), \tag{3.4.5}$$

where $s_i(\varphi)$ is given in equation (3.4.2). Note that we can thus write

$$q_i(\varphi, \omega) = (p_i(\varphi, \omega))^{-(\sigma_w - 1)} (p_i(\varphi))^{\sigma_w - \sigma} E_j P_i^{\sigma - 1},$$

which illustrates that whether demand for individual varieties produced by firm φ, increases or decreases with the price of other varieties produced by the same firm depends on the relative size of σ_w and σ. When varieties are more substitutable within firms than across firms ($\sigma_w > \sigma$), the lower the firm-level price index $p_i(\varphi)$, the lower the demand for an individual variety, thus capturing a *demand cannibalization* effect. Conversely, when varieties are more substitutable across firms than within firms ($\sigma_w < \sigma$), a lower firm-level price index $p_i(\varphi)$ will disproportionately redirect demand towards all of firm φ's varieties, thus creating a form of *demand complementarity* across a firm's varieties.

3.4.2 Manufacturing Production

Manufactured varieties are produced under increasing returns to scale, and market structure in this final-good production sector is characterized by monopolistic competition with free entry. As mentioned above, each firm owns a blueprint to produce a unit measure

of differentiated varieties of goods. Production of final-good varieties requires labor and a bundle of intermediate inputs. We index final-good firms by their 'core productivity', which we denote by φ, and following Melitz (2003), we assume that firms only learn their productivity φ after incurring an entry cost equal to f_h^e units of labor in their country of incorporation h (i.e., in the $h$eadquarter country). This core productivity is drawn from a country-specific distribution $g_h(\varphi)$, with support in $[\underline{\varphi}_h, \infty)$, and with an associated continuous cumulative distribution $G_h(\varphi)$.

The mapping between final-good production and the bundle of intermediate inputs is similar to that in Antràs, Fort and Tintelnot (2017). The bundle of intermediates contains a continuum of measure one of firm-specific inputs, assumed to be imperfectly substitutable with each other, with a constant and symmetric elasticity of substitution equal to ρ. Although intermediates are produced worldwide, a final-good producer based in country h only acquires the capability to offshore to j after incurring a fixed cost equal to f_{hj}^s units of labor in country h. We denote by $\mathcal{J}_h(\varphi) \subseteq J$ the set of countries for which a firm headquartered in h with productivity φ has paid the associated fixed cost of offshoring (denoted by $w_h f_{hj}^s$ for $j \in \mathcal{J}_h(\varphi)$). For brevity, we will often refer to $\mathcal{J}_h(\varphi)$ as the *global sourcing strategy* of that firm.

Intermediates are produced by a competitive fringe of suppliers who sell their products at marginal cost.[21] All intermediates are produced with labor under constant-returns-to-scale technologies. We denote by $a_j(v, \varphi)$ the unit labor requirement associated with the production of firm φ's intermediate $v \in [0, 1]$ in country $j \in J$. Shipping intermediates from country j to country k entails iceberg trade costs τ_{jk}^s. As a result, the cost at which firms producing in k can procure input v from country j is given by $\tau_{jk}^s a_j(v, \varphi) w_j$.

Note that we are using four different subindices to denote countries: h denotes the

[21] As in Antràs, Fort and Tintelnot (2017), we implicitly assume that contracts between final-good producers and suppliers are perfectly enforceable, so that the firm-specificity of inputs is irrelevant for the prices at which inputs are transacted.

country in which a firm is headquartered (i.e., the country of entry); k denotes a country in which assembly takes place; j denotes a country from which inputs are sourced; and i denotes the country in which a final good is sold and consumed.

The overall marginal cost for firm φ headquartered in h to produce units of final-good variety ω in country k is given by

$$c_k\left(\{j(v)\}_{v=0}^{1}, \varphi, \omega\right) = \frac{1}{\varphi}\frac{1}{z_k(\varphi,\omega)}(w_k)^{1-\alpha}\left(\int_0^1\left(\tau_{jk(v)}^{s}a_{j(v)}(v,\varphi)w_{j(v)}\right)^{1-\rho}dv\right)^{\alpha/(1-\rho)}, \tag{3.4.6}$$

where $\{j(v)\}_{v=0}^{1}$ corresponds to the infinitely dimensional vector of locations of intermediate input production, $\tau_{j(v)k}$ denotes the iceberg trade costs between the input production location $j(v)$ and the assembly country k, $1-\alpha$ is the value-added (labor) share in final-good production, and $z_k(\varphi,\omega)$ is a firm- and location-specific productivity level associated with assembling product ω in location k.[22]

As in Antràs, Fort and Tintelnot (2017), we treat the (infinite-dimensional) vectors of firm-specific intermediate input efficiencies $1/a_j(v,\varphi)$ as the realization of a Fréchet distribution

$$\Pr(a_j(v,\varphi) \geqslant a) = e^{-T_j^s a^{\theta^s}}, \quad \text{with } T_j^s > 0. \tag{3.4.7}$$

These draws are assumed to be independent across locations and inputs. As in Eaton and Kortum (2002), T_j^s governs the state of input production technology in country j, while θ^s determines the dispersion of productivity draws across inputs, with a lower θ^s fostering the emergence of comparative advantage *within* the range of intermediates across countries. For technical reasons described below, for our equilibrium to be well-behaved, we need to impose a lower bound on the dispersion in the input productivity draws $a_j(v,\varphi)$:

[22]It is worth stressing that we do not introduce any *direct* dependence of the cost function in (3.4.6) on the country h in which the headquarters are located. It might be natural to let $z_k(\varphi,\omega)$ include a shifter that decreases in the distance between countries h and k, thereby generating "headquarter gravity" (see Head and Mayer (2019) and Wang (2019)). We choose not do so to emphasize the role of endogenous sourcing strategies in inducing such headquarter gravity.

Assumption 1. $\rho - 1 < \theta^s$.

Apart from satisfying this restriction, the value of ρ does not matter for any outcomes of interest and will be absorbed into a constant hereafter.

The main substantive deviation from Antràs, Fort and Tintelnot (2017) is that we relax the assumption that final goods are nontradable and allow firms to produce and market their goods in *any* country in the world. Selling goods abroad involves additional fixed costs. First, we introduce an initial fixed cost $w_h f_h^g$ required for a firm to become a "global firm"Ṁore specifically, firms that do not incur this fixed cost cannot market goods in countries other than their home base h and cannot import inputs from countries other than h; once that fixed cost is paid, a firm can market goods in any country, but to import inputs from particular countries, it needs to pay the additional costs described above. Second, we assume that setting up an assembly plant in a given country $k \in J$ is associated with fixed overhead costs, so in equilibrium, firms will only find it optimal to set up a limited number of assembly plants (possibly a single one).

Following Tintelnot (2017), we consider the problem of solving for the optimal set of assembly locations from which to service consumers worldwide. We denote by $\mathcal{K}_h(\varphi) \subseteq J$ the set of countries $k \in J$ for which a firm headquartered in h with productivity φ has paid the associated fixed cost of assembly $w_h f_{hk}^a$. For brevity, we will often refer to $\mathcal{K}_h(\varphi)$ as the *global assembly strategy* of that firm. Shipping final goods from country k to country i entails variable (iceberg) trade costs τ_{ki}^a, which in principle may differ from those associated with shipping intermediate inputs. For simplicity, we abstract from modeling destination-specific fixed costs of exporting (see Tintelnot (2017) for a discussion of the complications that arise when introducing such fixed costs), though we note that the fixed cost of becoming global *will* create selection into exporting in our model (see Proposition 7 below).

Following Tintelnot (2017), we assume that the firm- and location-specific assembly

productivity shifters are drawn from the following Fréchet distribution:

$$\Pr(1/z_k(\varphi,\omega) \geqslant a) = e^{-T_k^a a^{\theta^a}}, \quad \text{with } T_k^a > 0. \tag{3.4.8}$$

Analogously to (3.4.7), T_k^a governs the state of assembly technology in country k, while θ^a determines the dispersion of productivity draws across final-good varieties, with a lower θ^a being associated with a higher variance, and thus with a higher benefit from producing final-good varieties in various locations. Again for technical reasons described below, a well-behaved equilibrium requires that we impose a lower bound on the dispersion in the final-good productivity draws $z_k(\varphi,\omega)$:

Assumption 2. $\sigma_w - 1 < \theta^a$.

Beyond ensuring that this restriction is satisfied, the parameter σ_w will not play an important role in the results below, so we will hereafter absorb into a constant (as we did with ρ above).

As in Antràs, Fort and Tintelnot (2017) and Tintelnot (2017), we assume that, at the time at which they decide on their sourcing and assembly strategies, firms know the distributions (3.4.7) and (3.4.8) but not the actual realizations of these various random variables (across the different goods ω they produce and the various active locations in their sourcing and assembly strategies).

Isomorphism with Armington Model. In our model, firms have an incentive to activate assembly locations or sourcing locations to reduce the cost at which they can satisfy the foreign demand for their products. Such cost minimization partly takes the shape of trade cost reductions (e.g., activating an export-platform close to specific countries), but it is also related to the fact that new assembly plants or sourcing locations allow the firm to learn about alternative technologies with which to produce final goods or procure inputs. The Fréchet formulation of these technologies coupled with the assumption that

the firm produces (sources) a continuum of final-good (input) varieties leads to a simple expression for the benefits of activating new assembly or sourcing locations (see section for more on this). It should be noted, however, that an entirely isomorphic set of equilibrium conditions arises in an Armington-like world in which final-good and input varieties are differentiated by country of origin. In that world, activating assembly locations would confer the firm the benefit to produce a new differentiated final-good variety, and activating a source of inputs would allow the firm to get access to a new differentiated input. To derive a set of isomorphic conditions to those of our Fréchet formulation, one needs to set the elasticity of substitution within firms across assembly locations as $\sigma_w = 1 + \theta^a$, and the elasticity of substitution across inputs produced in different countries as $\rho = 1 + \theta^s$.[23]

This completes the discussion of the assumptions of our model. Before describing the equilibrium of the model, it is worth pausing to spell out the precise timing of events in the model (focusing on the manufacturing sector).

1. Firms worldwide decide whether to pay a fixed cost $w_h f_h^e$ to set up headquarters in any country $h \in J$.

2. Upon observing their realized core productivity level φ, firms decide whether to exit, remain inward oriented, or pay a fixed cost $w_h f_h^g$ to be global and thus be able to import from and sell in foreign markets.

3. Global firms decide on their assembly strategy $\mathcal{K}_h(\varphi)$ and their sourcing strategy $\mathcal{J}_h(\varphi)$, paying the associated fixed costs $w_h f_{hk}^a$ and $w_h f_{hj}^s$. For inward-oriented firms, $\mathcal{K}_h(\varphi) = \mathcal{J}_h(\varphi) = \{h\}$.

4. Firms observe the realization of the productivity levels $a_j(v, \varphi)$ and $z_k(\varphi, \omega)$ for all

[23]Strictly speaking, the isomorphism also requires that the unit labor requirement for producing intermediate inputs in country $j \in J$ be proportional to $\left(T_j^s\right)^{-1/\theta^s}$, and the unit labor requirement for producing final goods in country $k \in J$ be proportional to $\left(T_k^a\right)^{-1/\theta^a}$.

$j \in \mathcal{J}_h(\varphi)$ and all $k \in \mathcal{K}_h(\varphi)$.

5. All assembly plants source inputs from their cheapest location within the firm's global sourcing strategy and consumers worldwide purchase manufacturing good varieties from the assembly plants that offer the minimum price for those varieties.

6. Production and consumption take place.

We solve for the equilibrium of the model in three steps. In the remainder of this section, we describe optimal firm behavior, industry equilibrium, and general equilibrium for given assembly and sourcing strategies $\mathcal{K}_h(\varphi)$ and $\mathcal{J}_h(\varphi)$, and we also describe some comparative statics related to how final-good and intermediate-input trade flows respond to changes in the parameters of the model. In the next section, we focus on characterizing the choice of these assembly and sourcing strategies, Finally, in section 3.6, we relate our results to our empirical evidence on interdependencies in section 3.3 and numerically evaluate the model.

3.4.3 Firm Behavior for Fixed Assembly and Sourcing Strategies

Consider a firm headquartered in country h with productivity φ that has incurred all necessary fixed costs associated with a given assembly strategy $\mathcal{K}_h(\varphi)$ and a given sourcing strategy $\mathcal{J}_h(\varphi)$. In light of the cost function in (3.4.6), it is clear that after learning the vector of unit labor requirements for each country $j \in \mathcal{J}_h(\varphi)$ and for each location of assembly $k \in \mathcal{K}_h(\varphi)$, the firm will choose the location of production for each input v that solves $\min_{j(v)\in\mathcal{J}_h(\varphi)} \left\{ \tau^s_{jk(v)} a_{j(v)}(v,\varphi)\, w_{j(v)} \right\}$. Notice that this is true independently of the particular realization of the productivity of the firm- and location-specific assembly productivity shifter $z_k(\varphi,\omega)$.

Using the properties of the Fréchet distribution in (3.4.7), one can then show that each assembly plant of the firm will source a positive measure of intermediates from each country in the firm's sourcing strategy set $\mathcal{J}_h(\varphi)$. Furthermore, the share of intermediate input purchases sourced by an assembly plant in $k \in \mathcal{K}_h(\varphi)$ from any country j is simply given by

$$\chi_{hjk}(\varphi) = \frac{T_j^s \left(\tau_{jk}^s w_j\right)^{-\theta^s}}{\Theta_{hk}(\varphi)} \quad \text{if } j \in \mathcal{J}_h(\varphi) \tag{3.4.9}$$

and $\chi_{hjk}(\varphi) = 0$ otherwise, where

$$\Theta_{hk}(\varphi) \equiv \sum_{j' \in \mathcal{J}_h(\varphi)} T_{j'}^s \left(\tau_{j'k}^s w_{j'}\right)^{-\theta^s}. \tag{3.4.10}$$

The term $\Theta_{hk}(\varphi)$ summarizes the *sourcing capability* of an assembly plant located in country k producing goods for a firm φ headquartered in country h. Note that, in equation (3.4.9), each sourcing country j's market share in country k's assembly plant input purchases corresponds to this sourcing country's contribution to its sourcing capability $\Theta_{hk}(\varphi)$. Countries in the set $\mathcal{J}_h(\varphi)$ with lower wages w_j, more advanced input technologies T_j^s, or lower trade costs when selling to country k will have higher market shares in the intermediate input purchases of firms based in country i. We shall refer to the term $T_j^s \left(\tau_{jk}^s w_j\right)^{-\theta^s}$ as the *sourcing potential* of country j from the point of view of assembly plants in k.[24]

After choosing the lowest-cost country for each input v, the overall marginal cost faced by firm φ, based in h, assembling final goods in country k can be expressed as

$$c_{hk}(\varphi, \omega) = \frac{1}{\varphi} \frac{1}{z_k(\varphi, \omega)} (w_k)^{1-\alpha} \left(\gamma \Theta_{hk}(\varphi)\right)^{-\alpha/\theta^s}, \tag{3.4.11}$$

[24]Notice that these sourcing potentials are independent of the headquarter country h. This prediction suggests that firms will source more from their headquarter country only due to variation in their extensive margin of sourcing. While the reduced-form evidence in section 3.3 shows that foreign-owned MNEs extensive and intensive margins of imports are larger from their home country, some those trade flows may represent imports of final goods rather than inputs.

where $\gamma = \left[\Gamma\left(\frac{\theta^s+1-\rho}{\theta^s}\right)\right]^{\theta^s/(1-\rho)}$ and Γ is the gamma function.[25] To ensure a well-defined marginal cost index, we assume that $\theta^s > \rho - 1$. Apart from satisfying this restriction, the value of ρ does not matter for any outcomes of interest and will be absorbed into a constant.

With this cost function in hand, we now consider the firm's choice of the optimal assembly plant from which to ship final-good varieties to a given destination i. Because the firm has already incurred all requisite fixed costs, this amounts to solving: $\min_{k(\omega)\in\mathcal{K}_h(\varphi)}\left\{\tau^a_{k(\omega)i}c_{hk(\omega)}(\varphi,\omega)\right\}$ for each variety ω. Using the properties of the Fréchet distribution, the share of firm φ's sales in market i originating from assembly plants in country k is given by:

$$\mu_{hki} = \frac{T^a_k\left(\tau^a_{ki}\right)^{-\theta^a}\left(w_k\right)^{-(1-\alpha)\theta^a}\left(\Theta_{hk}(\varphi)\right)^{\alpha\theta^a/\theta^s}}{\Psi_{hi}} \tag{3.4.12}$$

with

$$\Psi_{hi}(\varphi) = \sum_{k'\in\mathcal{K}_h(\varphi)} T^a_{k'}\left(\tau^a_{k'i}\right)^{-\theta^a}\left(w_{k'}\right)^{-(1-\alpha)\theta^a}\left(\Theta_{hk'}(\varphi)\right)^{\alpha\theta^a/\theta^s}. \tag{3.4.13}$$

Henceforth, we refer to the term $T^a_k\left(\tau^a_{ki}\right)^{-\theta^a}\left(w_k\right)^{-(1-\alpha)\theta^a}$ as the *assembly potential* of country k when selling to country i, while we refer to the term $\Psi_{hi}(\varphi)$ in equation (3.4.13) as the *global production capability* of a firm φ headquartered in country h when selling in i.

This global production capability turns out to be a sufficient statistic for the price index at which firm φ based in h sells its unit measure of varieties in market i, as defined in equation (3.4.3). In particular, a cumbersome set of derivations demonstrate that this price index is given by:

$$p_{hi}(\varphi) = \frac{\sigma_w}{\sigma_w - 1}\frac{1}{\varphi}\left(\zeta\Psi_{hi}(\varphi)\right)^{-1/\theta^a}, \tag{3.4.14}$$

where $\zeta = \left[\Gamma\left(\frac{\theta^a+1-\sigma}{\theta^a}\right)\right]^{\theta^a/(1-\sigma_w)}$ and Γ is again the gamma function. This formula

[25] These derivations are analogous to those performed by Eaton and Kortum (2002) to solve for the aggregate price index in their model.

illustrates that the benefits a firm obtains from building a global production capability by either selecting into global sourcing or assembly from more and more countries is crucially shaped by the parameter θ^a: the lower is θ^a, the higher is the achieved firm-level price index reduction. To ensure a well-defined price index, we need to assume that $\sigma_w - 1 < \theta^a$. Beyond ensuring that this restriction is satisfied, the parameter σ_w will not play an important role in the results below, so we will absorb into a constant (as we did with ρ above).

Finally, we can express the firm's profits conditional on a sourcing strategy $\mathcal{J}_h(\varphi)$ and an assembly strategy $\mathcal{K}_h(\varphi)$ as

$$\begin{aligned}\pi_h(\varphi, \mathcal{J}_h(\varphi), \mathcal{K}_h(\varphi)) = \kappa\varphi^{\sigma-1} \sum_{i \in J} \left(\Psi_{hi}(\varphi)\right)^{(\sigma-1)/\theta^a} E_i P_i^{\sigma-1} - \\ - w_h \sum_{j \in \mathcal{J}_h(\varphi)} f_{hj}^s - w_h \sum_{k \in \mathcal{K}_h(\varphi)} f_{hk}^a - w_h f_h^g,\end{aligned} \tag{3.4.15}$$

where κ is a constant, P_i is the standard ideal price index associated with (3.4.1) and defined in (3.4.4), and E_i is aggregate spending on manufactured goods in country i. As is clear from equation (3.4.15), the manner in which the global production capability $\Psi_{hi}(\varphi)$ shapes operating profits is crucially shaped by the exponent $(\sigma-1)/\theta^a$. As we discussed above, a lower θ^a enhances the scope for the within-firm exploitation of comparative advantage across the firm's assembly plants, thereby leading to a disproportionately lower firm-level price index $p_{hi}(\varphi)$ for firms choosing more complex global production strategies. A higher σ, in turn, leads to a larger impact of a reduction in $p_{hi}(\varphi)$ on firm-level sales (the often-called 'scale effect') and operating profits.

As we shall discuss in sections 3.4.9 and 3.5, whether $(\sigma-1)/\theta^a$ is higher or lower than 1 will have important implications for certain comparative statics results as well as for the determination of the optimal determination of the global sourcing and global assembly strategies of firms. In Tintelnot (2017), varieties are equally substitutable across and within firms, or $\sigma_w = \sigma$ in our notation, so invoking the Technical Condition 2 ($\sigma_w - 1 < \theta_a$),

his analysis follows on the case in which $(\sigma-1)/\theta^a < 1$. Our more general formulation of preferences demonstrates that the same condition would apply in the presence of demand cannibalization effects, which is associated with cases in which $\sigma_w < \sigma$ (see our discussion at the end of section 3.4.1). Nevertheless, in the presence of demand complementarities (i.e., $\sigma_w > \sigma$), it is perfectly possible for the case $(\sigma-1)/\theta^a \geqslant 1$ to apply.

In the derivations above, we assume that the firm has incurred the fixed cost f_h^g of becoming global, so that its global sourcing and assembly strategies are not trivially $\mathcal{K}_h(\varphi) = \mathcal{J}_h(\varphi) = \{h\}$. When the firm decides not to "go global", its profits above reduce to

$$\pi_h(\varphi, \{h\}, \{h\}) = \kappa\varphi^{\sigma-1}\left(\left(T_h^a\right)^{1/\theta^a}\left(T_h^s\right)^{\alpha/\theta^s}\left(\tau_{hh}^a\right)^{-1}\left(\tau_{hh}^s\right)^{-\alpha}\left(w_h\right)^{-(1-\alpha)}\right)^{\sigma-1} E_i P_i^{\sigma-1} - \\ - w_h f_{hh}^s - w_h f_{hh}^a.$$

If a firm's core productivity is such that $\pi_h(\varphi, \{h\}, \{h\}) < 0$, this firm will exit upon observing its core productivity level.

3.4.4 Intensive Margin Analysis: Complementarities and Cannibalization Effects

Before discussing the determination of the global sourcing and assembly strategies of firms, we briefly describe how conditional on these strategies, bilateral final-good and intermediate input flows are shaped by the parameters of the model.

Given CES preferences in (3.4.1), it is well understood that final-good sales of firm φ (based in h) in market i are proportional to the operating profits of selling in that market, so from equation (3.4.15), we have that the total final-good sales of firm φ are

$$S_{hi}(\varphi) = \widetilde{\kappa}\varphi^{\sigma-1}\sum_{i\in J}\left(\Psi_{hi}(\varphi)\right)^{(\sigma-1)/\theta^a} E_i P_i^{\sigma-1}, \tag{3.4.16}$$

where $\widetilde{\kappa}$ is a constant, and where remember that $\Psi_{hi}(\varphi)$ is the global production capability

of firm φ when selling in i. We also established above that a share μ_{hki} – defined in equation (3.4.12) – of this firm's sales in country i originate in a plant located in k, and thus total sales of the firm's plant in k (aggregated across destinations i) are given by

$$S_{hk}(\varphi) = \widetilde{\kappa}\varphi^{\sigma-1} T_k^a (w_k)^{-(1-\alpha)\theta^a} (\Theta_{hk}(\varphi))^{\alpha\theta^a/\theta^s} \sum_{i\in J} (\tau_{ki}^a)^{-\theta^a} (\Psi_{hi}(\varphi))^{\frac{(\sigma-1)}{\theta^a}-1} E_i P_i^{\sigma-1} \tag{3.4.17}$$

where remember that $\Theta_{hk}(\varphi)$ is the sourcing capability of that assembly plant.

We can now prove the following result:

Proposition 3. *Holding constant the market demand level $E_i P_i^{\sigma-1}$ and a firm's global sourcing $\mathcal{J}_h(\varphi)$ and assembly $\mathcal{K}_h(\varphi)$ strategies, an increase in plant k's assembly potential $T_k^a (\tau_{ki}^a)^{-\theta^a} (w_k)^{-(1-\alpha)\theta^a}$ or a decrease in a bilateral input trade cost τ_{jk}^s (for any j):*

i) increases sales of plants based in k in country i;

ii) increases sales of plants based in $k' \neq k$ in country i if $(\sigma-1)/\theta^a > 1$, and reduces them if $(\sigma-1)/\theta^a < 1$.

Proof. See Appendix C.1. □

In words, whether increases in the profitability of a given assembly plant k (driven by improvements in assembly potential or reductions in sourcing costs) increase or decrease the sales of other plants selling the firm's goods in market i crucially depends on the relative size of $\sigma - 1$ and θ^a. When $\sigma - 1 < \theta^a$, the dispersion in productivity across plants is relatively low, and our model delivers cannibalization effects. Conversely, when $\sigma - 1 > \theta^a$, complementarities in demand are high and productivity dispersion across plants is high, so increases in the efficiency of particular plants generate a positive effect on the sales of other plants of the same firm (holding constant market demand). Remember that to obtain a bounded price index, we imposed above $\sigma_w - 1 < \theta^a$. Thus $\sigma - 1 > \theta^a$ if and only if $\sigma_w < \sigma$, which corresponds to the case with demand complementarities

described at the end of section 3.4.1. Tintelnot (2017) focused on the case $\sigma_w = \sigma$ (i.e., symmetric CES preferences, so Technical Assumption 1 led him to focus on the case $\sigma_w - 1 < \theta^a$.

We next turn to study how intermediate input flows are shaped by the fundamental parameters of our model. We first note that given our functional form assumptions (particularly the Cobb-Douglas technology in (3.4.11)), total intermediate input purchases of a given plant (say in country k) are a constant share of this plant's total sales $S_{hk}(\varphi)$ in equation (3.4.17). Furthermore, imports from each source j correspond to $\chi_{hjk}(\varphi)$ in equation (3.4.9), and are thus given by:

$$M_{hkj}(\varphi) = \widehat{\kappa}\varphi^{\sigma-1} T_k^a (w_k)^{-(1-\alpha)\theta^a} T_j^s \left(\tau_{jk}^s w_j\right)^{-\theta^s} \left(\Theta_{hk}(\varphi)\right)^{\frac{\alpha\theta^a}{\theta^s}-1} \times \\ \times \sum_{i \in J} \left(\tau_{ki}^a\right)^{-\theta^a} \left(\Psi_{hi}(\varphi)\right)^{\frac{(\sigma-1)}{\theta^a}-1} E_i P_i^{\sigma-1}. \tag{3.4.18}$$

Two results follow from this expression. First, it is an immediate corollary of Proposition 3 that the complementarities or cannibalization effects associated with changes in assembly potential identified in that Proposition carry over to the input purchases of those plants. More formally,

Proposition 4. *Holding constant the market demand level $E_i P_i^{\sigma-1}$ and a firm's global sourcing $\mathcal{J}_h(\varphi)$ and assembly $\mathcal{K}_h(\varphi)$ strategies, an increase in plant k's assembly potential $T_k^a \left(\tau_{ki}^a\right)^{-\theta^a} (w_k)^{-(1-\alpha)\theta^a}$:*

i) increases input purchases from any source country $j \in \mathcal{J}_h(\varphi)$ of the plant based in k;

ii) increases input purchases from any source country $j \in \mathcal{J}_h(\varphi)$ for plants based in $k' \neq k$ in country i if $(\sigma-1)/\theta^a > 1$, and reduces them if $(\sigma-1)/\theta^a < 1$.

Proof. See Appendix C.1. □

Our second result relates to the impact of changes in input trade costs on bilateral input purchases. Such an analysis is a bit more cumbersome than the one related to

changes in assembly potentials because it is shaped by interdependencies in both final-good sales across assembly plants, as well as by interdependencies in importing across sourcing countries j. We are however able to prove the following result:

Proposition 5. *Holding constant the market demand level $E_i P_i^{\sigma-1}$ and a firm's global sourcing $\mathcal{J}_h(\varphi)$ and assembly $\mathcal{K}_h(\varphi)$ strategies, a decrease in a bilateral input trade cost τ_{jk}^s:*

i) increases input purchases from country j by plants based in k;

ii) increases input purchases from any country $j \neq j$ by plants based in k if $\sigma - 1 \geqslant \theta^a > \theta^s/\alpha$, and reduces them if $\sigma - 1 < \theta^a < \theta^s/\alpha$.

Proof. See Appendix C.1. □

This result is related to Proposition 3 in Antràs, Fort and Tintelnot (2017). Intuitively, when demand is sufficiently elastic (i.e., σ is high enough) and the strength of comparative advantage in the intermediate-good sector across countries is sufficiently high (i.e., θ^s is low enough), the scale effect caused by the marginal-cost reduction induced by the reduction in the sourcing cost τ_{jk}^s dominates the direct substitution effect related to market shares shifting towards the sourcing location whose cost of sourcing has been reduced. As a result, the reduction in τ_{jk}^s leads to an increase in sourcing by the plant in k from country j but also from all others sources of inputs. Conversely, when $\sigma - 1 < \theta^a < \theta^s/\alpha$, the reduction in τ_{jk}^s reduces sourcing from other sources. A small departure from the result in Antràs, Fort and Tintelnot (2017) is that whether input sources are complements or substitutes does not only depend on the relative size of $\sigma - 1$ and θ^s/α, but also on the relative size of $\sigma - 1$ and θ^a, because the ratio $(\sigma - 1)/\theta^a$ is a key determinant of the plant-level scale response to a change in marginal costs.

In section 3.6, we return to equation (3.4.18) and we leverage it to attempt to interpret the reduced-form results we obtained in section 3.3.

Although perhaps a more natural next step would be to solve for all the firm's global assembly $\mathcal{K}_h(\varphi) \subseteq J$ and global sourcing $\mathcal{J}_h(\varphi) \subseteq J$ sets, we first briefly outline the industry and general equilibrium of the model given those sets.

3.4.5 Industry and General Equilibrium

Once the problem of the firm is solved, the industry and general equilibrium is straightforward to characterize. As mentioned before, we simplify matters by assuming that consumers spend a constant share (which we denote by η) of their income on manufacturing. The remaining share $1-\eta$ of income is spent on a perfectly competitive non-manufacturing sector that competes for labor with manufacturing firms. Technology in that sector is linear in labor, and we assume that $1-\eta$ is large enough to guarantee that the wage rate w_h in each country h is pinned down by labor productivity in that sector. For simplicity, we also assume that this 'outside' sector's output is homogeneous, freely tradable across countries, and serves as a numéraire in the model. We thus can treat wages as exogenous in solving for the equilibrium in each country's manufacturing sector.

We next turn to describing the equilibrium in the manufacturing sector. Given our assumption that final-good producers only observe their productivity after paying the fixed cost of entry, we can use equation (3.4.15) to express the free-entry condition in manufacturing as

$$\int_{\tilde{\varphi}_h}^{\infty} \pi_h\left(\varphi, \mathcal{J}_h(\varphi), \mathcal{K}_h(\varphi)\right) dG_h(\varphi) = w_h f_{eh}. \tag{3.4.19}$$

In the lower bound of the integral, $\tilde{\varphi}_h$ denotes the productivity of the least productive active firm in country h. Firms with productivity $\varphi < \tilde{\varphi}_h$ cannot profitably carry out any strategy and thus exit upon observing their productivity level. Because expected profits are zero, all income is wage income, so $E_h = \eta w_h L_h$, and equation (3.4.19) constitutes a system of J equations from which the manufacturing price indices P_h can be solved for.

This completes the description of the model for given assembly and sourcing strategies.

3.5 Optimal Sourcing and Assembly Strategies

Having solved for the model's equilibrium conditional on the sets $\mathcal{K}_h(\varphi) \subseteq J$ and $\mathcal{J}_h(\varphi) \subseteq J$, we turn to characterizing these sets. To develop intuition, we first outline this problem as in Antràs, Fort and Tintelnot (2017) and Tintelnot (2017), where firms have direct control over these sets and decide whether to activate particular countries as sources of inputs or as platforms from which to assemble and export final-good varieties. Since finding a solution to this problem appears to be computationally infeasible in some potentially empirically relevant cases, we then outline an alternative approach that circumvents the obstacles that arise from our initial approach.

3.5.1 The Problem and General Results

Each firm's optimal assembly and sourcing strategy is a combinatorial optimization problem in which two sets of locations are chosen to maximize the firm's profits $\pi_h(\varphi)$ in equation (3.4.15).[26] Plugging in equations (3.4.10) and (3.4.13), and defining $\xi^a_{ki} = T^a_k \left(\tau^a_{ki}\right)^{-\theta^a} \left(w_k\right)^{-(1-\alpha)\theta^a}$ and $\xi^s_{jk} = T^s_j \left(\tau^s_{jk} w_j\right)^{-\theta^s}$ we can express this problem as

$$\begin{aligned}\max_{\substack{\mathcal{I}^a_k \in \{0,1\} \\ \mathcal{I}^s_j \in \{0,1\}}} \pi_h\left(\varphi, \mathcal{J}_h(\varphi), \mathcal{K}_h(\varphi)\right) \quad = \quad & \kappa \varphi^{\sigma-1} \sum_{i \in J} B_i \left(\sum_{k \in J} \mathcal{I}^a_k \cdot \xi^a_{ki} \left(\sum_{j \in J} \mathcal{I}^s_j \cdot \xi^s_{jk} \right)^{\frac{\alpha\theta^a}{\theta^s}} \right)^{\frac{(\sigma-1)}{\theta^a}} \\ & - \sum_{j \in J} \mathcal{I}^s_j \cdot w_h f^s_{hj} - \sum_{k \in J} \mathcal{I}^a_k \cdot w_h f^a_{hk} - w_h f^g_h, \end{aligned} \tag{3.5.1}$$

[26] More formally, let $\pi_h(\varphi) : \{0,1\}^{2J} \to \mathbb{R}_+$ be a variable profit function defined over the boolean hypercube. Let $\mathcal{I} = (\mathcal{I}^a, \mathcal{I}^s) \in \{0,1\}^{2J}$ with $\mathcal{I}^a \in \{0,1\}^J$ and $\mathcal{I}^s \in \{0,1\}^J$. If the firm builds an assembly plant in location k then $\mathcal{I}^a_k = 1$ and $\mathcal{I}^a_k = 0$ otherwise; if the firm builds a sourcing plant in location j then $\mathcal{I}^s_j = 1$ and $\mathcal{I}^s_j = 0$ otherwise. The corresponding assembly and sourcing strategies are defined as $\mathcal{K}_h(\varphi) = \left\{k \in J : \mathcal{I}^a_k = 1\right\}$ and $\mathcal{J}_h(\varphi) = \left\{j \in J : \mathcal{I}^s_j = 1\right\}$, respectively.

where $B_i = E_i P_i^{\sigma-1}$, and where the indicator variables $\mathcal{I}_k^a$ (respectively, $\mathcal{I}_j^s$) takes a value of 1 when $k \in \mathcal{K}_h(\varphi)$ (respectively, $j \in \mathcal{J}_h(\varphi)$), and 0 otherwise. The problem in (3.5.1) is an NP-complex combinatorial problem that is infeasible to solve computationally by brute force when the number of countries J is sufficiently large. A similar issue arises when focusing on studying the assembly and sourcing strategies separately, as in Antràs, Fort and Tintelnot (2017) and Tintelnot (2017), but the joint determination of these two strategies renders this problem even more formidable. The reason for this is that the monotone comparative statics techniques that can be applied in the frameworks of Antràs, Fort and Tintelnot (2017) and Tintelnot (2017) are much less powerful when these strategies are studied jointly.

To elucidate the added complexity of the current setting, we begin by noting some key properties of the profit function in (3.5.1).

Lemma 1. *The profit function* $\pi_h(\varphi, \mathcal{J}_h(\varphi), \mathcal{K}_h(\varphi))$ *in (3.5.1) features:*

i) increasing differences in $\left(\mathcal{I}_k^a, \mathcal{I}_{k'}^a\right)$ *for* $k, k' \in \{1, ..., J\}$ *and* $k \neq k'$ *when* $\sigma - 1 > \theta^a$, *and decreasing differences in* $\left(\mathcal{I}_k^a, \mathcal{I}_{k'}^a\right)$ *for* $k, k' \in \{1, ..., J\}$ *and* $k \neq k'$ *when* $\sigma - 1 < \theta^a$;

ii) increasing differences in $\left(\mathcal{I}_j^s, \mathcal{I}_{j'}^s\right)$ *for* $j, j' \in \{1, ..., J\}$ *when* $\sigma - 1 \geqslant \theta^a > \theta^s/\alpha$, *and decreasing differences in* $\left(\mathcal{I}_j^s, \mathcal{I}_{j'}^s\right)$ *for* $j, j' \in \{1, ..., J\}$ *when* $\sigma - 1 < \theta^a < \theta^s/\alpha$;

iii) increasing differences in $\left(\mathcal{I}_k^a, \mathcal{I}_j^s\right)$ *for* $k \in \{1, ..., J\}$ *and* $j \in \{1, ..., J\}$ *when* $\sigma - 1 \geqslant \theta^a$.

Proof. See Appendix C.1. □

In words, part i) of Lemma 1 implies that adding an assembly plant in any given country k may increase or decrease the profitability of doing so in any other country k' depending on the relative size of $\sigma - 1$ and θ^a. As mentioned before, Tintelnot (2017) focused on the case with no demand cannibalization or complementarity effects, which

naturally led him, given Technical Assumption 1, to focus on the case $\sigma - 1 < \theta^a$, in which assembly extensive margin decisions are substitutes. Nevertheless, in the presence of demand complementarities ($\sigma_w < \sigma$), it is theoretically perfectly possible for $\sigma - 1 > \theta^a$, in which case assembly extensive margin decisions are complements. Intuitively, in such a case, the combination of the scale effect and demand complementarity effects are strong enough to counterbalance the natural substitutability emanating from different plants within a firm competing to serve the same set of consumers worldwide.

Part ii) of Lemma 1 is closely related to one of the main results in Antràs, Fort and Tintelnot (2017): it identifies a key condition under which the extensive margin of global sourcing features substitutability. When demand is relatively inelastic (low σ) or the intermediate input share (α) is low, firm scale is not particularly responsive to variable cost reductions in sourcing. Under these circumstances, the addition of a country to a firm's sourcing strategy necessarily decreases the marginal benefit of investing in activating alternative sourcing locations, particularly when supplier productivity dispersion is low (high θ^s). A notable difference relative to the setting in Antràs, Fort and Tintelnot (2017) is that the condition $\sigma - 1 > \theta^s/\alpha$ or $\sigma - 1 < \theta^s/\alpha$ is not sufficient to determine whether profits feature increasing or decreasing differences in these extensive margin sourcing decision; it is also necessary for θ^a to fall between these two values. For instance, when $\sigma - 1 < \theta^a$, the profit function in equation (3.5.1) need not feature increasing differences in the extensive margin of firm sourcing when $\sigma - 1 > \theta^s/\alpha$ due to the substitutability of sales across assembly locations. The reason for this is that if there is enough heterogeneity in input trade costs, the addition of a new input source may lower production costs at plants that use those inputs intensively so much that it reduces the revenues of other assembly locations that do not. This potential for cannibalization across assembly locations may in turn reduce the marginal benefit of sourcing from countries that are proximate to the assembly locations for which revenues shrink. As a result, even when the extensive margin of assembly locations is fixed, the addition of one sourcing location has the potential to

reduce the marginal benefit of adding another sourcing location.

For similar reasons, even though scale effects and the complementarity between inputs and assembly in technology would suggest the existence of complementarities in the global assembly and global sourcing strategies of firms, part iii) of Lemma 1 indicates that for the profit function $\pi_h(\varphi)$ in (3.5.1) to necessarily feature complementarity (or increasing differences) between $\left(\mathcal{I}_k^a, \mathcal{I}_j^s\right)$ for $k, j \in \{1, ..., J\}$, it is necessary that $\sigma - 1 \geqslant \theta^a$. When assembly decisions are substitutes ($\sigma - 1 < \theta^a$), the marginal benefit of activating specific sourcing countries may actually be diminished by the activation of specific assembly locations. Furthermore, even when pairs of assembly locations and pairs of sourcing locations are substitutes (i.e., $\sigma - 1 < \theta^a < \theta^s/\alpha$), pairs of assembly and sourcing locations may actually prove to be complements, due to the scale effects associated with the marginal-cost reducing effects of global sourcing.

The fact that the profit function in (3.5.1) features various sources of complementarity and substitutability between the extensive margins of sourcing and assembly limits the analytical characterization of these optimal firm strategies. Nevertheless, the fact that the profit function (3.4.15) is supermodular in φ and the sum $\sum_{i \in J} \left(\mathbf{\Psi}_{hi}(\varphi)\right)^{(\sigma-1)/\theta^a} E_i P_i^{\sigma-1}$ implies that:

Proposition 6. *The optimal assembly and sourcing strategies that solve problem (3.5.1) imply that the vector of a firm's global production capabilities is such that* $\sum_{i \in J} \left(\mathbf{\Psi}_{hi}(\varphi)\right)^{(\sigma-1)/\theta^a} E_i P_i^{\sigma-1}$ *is nondecreasing in* φ.

Proof. See Appendix C.1. □

The result implies that more productive firms choose a vector of global production capabilities that translates into differences in world sales across firms that are magnified relative to the differences that would arise in a world without global assembly and global sourcing. An immediate corollary of this result is that the marginal benefit of paying the

fixed cost $w_h f_h^g$ of 'going global' is necessarily higher for more productive firms, which implies that:

Proposition 7. *There exists a threshold productivity φ_h^*, such that only firms headquartered in h with $\varphi > \varphi_h^*$ find it optimal to become global firms.*

Proof. See Appendix C.1. □

In sum, although characterizing the specific global sourcing and assembly strategies of firms and how they correlated with productivity φ is complicated, our model necessarily features selection into exporting and FDI of the type produced by the canonical frameworks in Melitz (2003) and Helpman, Melitz and Yeaple (2004). While our empirical evidence in section 3.2 suggests that firms with higher core productivity φ source inputs from, and locate assembly plants in, more countries, our model produces these results only when imposing additional parametric restrictions. In the next section, we illustrate this for a specific region of the parameter space.

We conclude by noting that beyond complicating the *characterization* of the extensive margins of global sourcing and global assembly, the coexistence of various sources of complementarity and substitutability also complicates the *computation* of these margins in quantitative analyses. Remember that the problem faced by the firm is a complex combinatorial optimization problem with $2^{J\times 2}$ possible choices, so it is infeasible to solve by brute force when the number of countries J is large. Furthermore, our framework does not generally feature the type of "single-crossing" properties that typically rationalize the use of iterative algorithms to reduce the dimensionality of the problem of solving for the firm's extensive margin, as in Jia (2008), Antràs, Fort and Tintelnot (2017) or Arkolakis, Eckert and Shi (2021). A special case where these algorithms continue to be applicable is offered in the next section.

3.5.2 The Case with Pervasive Complementarities

In this section, we study the determination of the extensive margins of global sourcing and global assembly for the special case in which $\sigma - 1 \geqslant \theta^a > \theta^s/\alpha$, which we refer to as the case with *pervasive complementarities*. From Lemma 1, this corresponds to the case in which the profit function features increasing differences in any two extensive margin decisions, regardless of whether they entail the addition of a source of inputs or of a platform of final-good production. Although the literature on export-platform FDI has generally focused on the case in which assembly decisions are substitutes (or $\sigma - 1 < \theta^a$), a few recent papers have considered environments with either independence or increasing differences in these strategies.[27]

In such a case, we can establish the following result:

Proposition 8. *Whenever* $\sigma - 1 \geqslant \theta^a \geqslant \theta^s/\alpha$, *we necessarily have that* $\mathcal{J}_h(\varphi_L) \subseteq \mathcal{J}_h(\varphi_H)$ *and* $\mathcal{K}_h(\varphi_L) \subseteq \mathcal{K}_h(\varphi_H)$ *for* $\varphi_H \geqslant \varphi_L$, *where* $\mathcal{J}_i(\varphi) = \left\{j : \mathcal{I}^s_{hj} = 1\right\}$ *and* $\mathcal{K}_i(\varphi) = \{k : \mathcal{I}^a_{hk}(\varphi) = 1\}$.

Proof. See Appendix C.1. □

Proposition 8 states that in the case with pervasive complementarities, our model delivers a strict hierarchical order in the extensive margin of global sourcing and of global assembly. More productive firms source from the same countries and possibly from additional ones relative to less productive firms, and they also produce in the same countries as less productive firms, and possibly in additional ones. Obviously, this strict 'pecking order' in the extensive margin of firms is violated in the data, but using our merged Census-NBER data, in section 3.3 we have validated a weaker implication of this prediction, namely that the more productive a firm is, the more countries it will source from and the more foreign affiliates it will set up to produce.

[27] See, in particular, Bernard *et al.* (2018), Arkolakis *et al.* (2021), and Garetto *et al.* (2019), with the latter paper offering suggestive evidence of independence in assembly decisions.

As in Antràs, Fort and Tintelnot (2017), the presence of (pervasive) complementarities in the extensive margin is also helpful for computational purposes when estimating the model, as it opens the door for the use of iterative algorithms that have the potential to dramatically decrease the dimensionality of the firm problem, even when fixed costs of sourcing and assembly are heterogeneous across firms. To illustrate this, consider the following result:

Proposition 9. *Define the mappings (i) $V_{h,j}(\varphi, \mathcal{J}, \mathcal{K})$ to take a value of one whenever including country j in the sourcing strategy $\mathcal{J}$ raises firm-level profits $\pi_h(\varphi, \mathcal{J}, \mathcal{K})$, and to take a value of zero otherwise, and (ii) $V_{h,k}(\varphi, \mathcal{J}, \mathcal{K})$ to take a value of one whenever including country k in the assembly strategy $\mathcal{K}$ raises firm-level profits $\pi_h(\varphi, \mathcal{J}, \mathcal{K})$, and to take a value of zero otherwise. Then, whenever $\sigma - 1 \geqslant \theta^a > \theta^s/\alpha$, $V_{h,j}(\varphi, \mathcal{J}', \mathcal{K}) \geqslant V_{h,j}(\varphi, \mathcal{J}, \mathcal{K})$ for $\mathcal{J} \subseteq \mathcal{J}'$ and $V_{h,k}(\varphi, \mathcal{J}, \mathcal{K}') \geqslant V_{h,k}(\varphi, \mathcal{J}, \mathcal{K})$ for $\mathcal{K} \subseteq \mathcal{K}'$.*

Proof. See Appendix C.1. □

The usefulness of this result is best demonstrated with an example. Suppose that one is trying to assess whether a given country k belongs in the firm's optimal assembly strategy $\mathcal{J}_h(\varphi)$. Without guidance from the theory, one would need to compute all $2^{J\times 2}$ possible candidate combinations of sourcing and assembly strategies to answer that question. Proposition 9 implies, however, that if for country j, $V_{h,j}(\varphi, j, \varnothing) = 1$ (so the initial sets $\mathcal{J}$ and $\mathcal{K}$ are the null sets), then j is necessarily in $\mathcal{J}_h(\varphi)$, while if $V_{i,j}(\varphi, \mathcal{J}) = 0$ when $\mathcal{J}$ includes all countries except for j and $\mathcal{K}$ includes all countries, then j cannot possibly be in $\mathcal{J}_h(\varphi)$. Following Jia (2008), Antràs, Fort and Tintelnot (2017) or Arkolakis, Eckert and Shi (2021), it is then straightforward to implement an iterative application of the V-operator that gradually tightens both the lower bound (i.e., the set of surely activated locations of sourcing and assembly) and the upper bound (i.e., the set of surely discarded locations for sourcing or for assembly) of the firm's sourcing and assembly strategies, thereby reducing the set of combinations that one needs to evaluate by brute force.

3.6 Zooming in on the Assembly-Sourcing Complementarity

Our framework features many interdependencies in both the extensive and intensive margins of global sourcing and global assembly. Some of these interdependencies are inherited from the baseline models we build on (i.e., Tintelnot, 2017 and Antràs, Fort and Tintelnot, 2017), but in this section we seek to highlight that some interdependencies are novel to our framework, and we also argue that these forces can be invoked to explain the reduced-form evidence we presented in section 3.3.2.

3.6.1 Isolating the Assembly-Sourcing Complementarity

A natural way to isolate the new forces in our paper is to focus on the case in which the extensive margin of assembly does not feature interdependencies across pairs of assembly locations and in which the extensive margin of sourcing does not feature complementarity or substitutability across pairs of sourcing locations. In terms of the parameters of our model, this corresponds to the case in which $\sigma - 1 = \theta^a = \theta^s/\alpha$, and from equation (3.5.1), this results in operating profits for firm φ based in h equal to

$$\kappa\varphi^{\sigma-1}\sum_{i\in J}B_i\sum_{k\in\mathcal{K}_h(\varphi)}T_k^a\left(\tau_{ki}^a\right)^{-\theta^a}\left(w_k\right)^{-(1-\alpha)\theta^a}\sum_{j\in\mathcal{J}_h(\varphi)}T_j^s\left(\tau_{jk}^s w_j\right)^{-\theta^s}.$$

Given these operating profits, consider then the impact on overall profits of adding some candidate sourcing location k' to the firm's sourcing strategy $\mathcal{J}_h(\varphi)$. Given the above equation, the firm will find the addition of j' profitable whenever

$$\kappa\varphi^{\sigma-1}\sum_{i\in J}B_i\sum_{k\in\mathcal{K}_h(\varphi)}T_k^a\left(\tau_{ki}^a\right)^{-\theta^a}\left(w_k\right)^{-(1-\alpha)\theta^a}T_{j'}^s\left(\tau_{j'k}^s w_{j'}\right)^{-\theta^s} > w_h f_{hj'}^s, \tag{3.6.1}$$

where the right-hand side is the fixed cost of activating that sourcing location j'. Clearly, the decision is shaped by both the sourcing potential $T_{j'}^s\left(\tau_{j'k}^s w_{j'}\right)^{-\theta^s}$ of location j' vis à vis all

assembly locations $k \in \mathcal{K}_h(\varphi)$, as well as by the assembly potential $T_k^a (\tau_{ki}^a)^{-\theta^a} (w_k)^{-(1-\alpha)\theta^a}$ of these activated assembly locations. Furthermore, the impact of each of these effects is increasing in the size of the other one. In words, the larger are the number of assembly locations and their assembly potential, the larger will be the benefit of activating location j'.

To further isolate the role of geography – or trade costs more broadly – on this complementarity, it is useful to consider the special case in which $T_k^a (w_k)^{-(1-\alpha)\theta^a} = G^a$, $T_j^s (w_j)^{-\theta^s} = G^s$, so variation in assembly and sourcing potentials is purely shaped by trade costs. Assume also that fixed costs of sourcing are common for all countries. In that case, equation (3.6.1) reduces

$$\kappa \varphi^{\sigma-1} G^a G^s \sum_{k \in \mathcal{K}_h(\varphi)} \sum_{i \in J} B_i (\tau_{ki}^a)^{-\theta^a} \left(\tau_{j'k}^s\right)^{-\theta^s} > w_h f_h^s. \tag{3.6.2}$$

It is then clear that whether an input source j' is activated or not depends on a market-access-weighted 'distance' of this source market j' from all the firm's assembly plants, where $B_i (\tau_{ki}^a)^{-\theta^a}$ are the weights and where 'distance' is captured by the trade costs $\tau_{j'k}^s$. In more plain words, activating an input source will be particularly profitable when this source can ship to the firm's various assembly plants at a relatively low cost.

This prediction of the model is in line with our empirical findings in section 3.3.2, and more specifically with our reduced-form regression results in Table 3.7, which show that multinational firms with manufacturing operations in the US are more likely to import from countries and regions in which the firm has operations (i.e., foreign affiliates of US-based MNEs or headquarters of foreign MNEs).

To further illustrate the source of this complementarity between sourcing and assembly, we consider the case in which the global sourcing strategy of firms is instead at the plant level, rather than at the firm level, as in the work of Bernard *et al.* (2018). More precisely, assume that assembly plants of the same company design their own independent sourcing

strategies, which we denote by $\mathcal{J}_{hk}(\varphi)$. In that case, a given plant k will choose to add location j' to their plant-specific sourcing strategy whenever

$$\kappa\varphi^{\sigma-1}T_k^a\left(w_k\right)^{-(1-\alpha)\theta^a}T_{j'}^s\left(\tau_{j'k}^s w_{j'}\right)^{-\theta^s}\sum_{i\in J}B_i\left(\tau_{ki}^a\right)^{-\theta^a} > w_h f_{hj'}^s.$$

This expression still features complementarity between j's sourcing potential and k's assembly potential, but notice that the location of the other assembly plants of firm φ become irrelevant. In regressions, such as those in Table 3.7 that fix the assembly country k (i.e., the US in our case) and that feature source country fixed effects, this variant of the model would thus predict a null effect of the presence of foreign production plants on the sourcing patterns of assembly plants in k. In other words, the patterns in Table 3.7 are strongly suggestive of the existence of firm-level (rather than plant-level) global sourcing strategies, which is a key novel feature of our framework.

So far we have focused on the extensive margin of global sourcing, but it is also worth discussing the implications of our framework for the intensive margin of global sourcing, and how these predictions relate to the findings in Table 3.8. For that purpose, note that we can write equation (3.4.18) as

$$\log M_{hkj}(\varphi) = \alpha + d_k + d_j + d_{hk\varphi} - \theta^s \ln\left(\tau_{jk}^s\right) \tag{3.6.3}$$

where $d_k = \log T_k^a\left(w_k\right)^{-(1-\alpha)\theta^a}$, $d_j = \log T_j^s\left(w_j\right)^{-\theta^s}$, and

$$d_{hk\varphi} = \log\left(\Theta_{hk}(\varphi)\right)^{\frac{\alpha\theta^a}{\theta^s}-1}\sum_{i\in J}\left(\tau_{ki}^a\right)^{-\theta^a}\left(\Psi_{hi}(\varphi)\right)^{\frac{(\sigma-1)}{\theta^a}-1}E_i P_i^{\sigma-1}.$$

Notice that in our empirical application with import data from US-based plants, the inclusion of firm fixed effects controls for d_{hk} and $d_{hk\varphi}$, so that once source country fixed d_j effects are included, our model predicts that the intensive margin of imports should *only* be shaped by bilateral trade frictions between j and k, and not by the presence of affiliates in other countries, as in the determination of the extensive margin of trade.

How does this prediction line up with our empirical results in Table 3.8? Consistent with our model, remember that our estimates indicated that imports by US-based MNE from a given country j are indeed unaffected by the presence of an affiliate in the same region to which j belongs (excluding country j itself).[28] Conversely, we have documented that imports by US affiliates of foreign-owned MNEs import significantly more from countries in the same region as their home country, which is inconsistent with the predictions of equation (3.6.3). A potential explanation for this inconsistency is that our framework has abstracted from the inclusion of 'headquarter gravity' forces, as in the work of Head and Mayer (2019) and Wang (2019). More specifically, if the trade costs associated with shipping inputs from country j into country k (or the efficiency with which they can be used in production) were to be a function of the distance between the headquarter country h and countries k or j, we would expect the emergence of 'extended gravity' only when $h \neq k$, as documented in Table 3.8.

In this section, we have focused on studying the impact of the geography of assembly on the global sourcing decisions of firms. Naturally, the same considerations would arise when studying how the addition of a new assembly plant is shaped by the global sourcing strategy of a firm. More specifically, when $\sigma - 1 = \theta^a = \theta^s/\alpha$, the addition of plant k' is profitable whenever

$$\kappa\varphi^{\sigma-1} \sum_{j\in\mathcal{J}_h(\varphi)} \sum_{i\in J} B_i T^a_{k'} \left(w_{k'}\right)^{-(1-\alpha)\theta^a} \left(\tau^a_{k'i}\right)^{-\theta^a} T^s_j \left(\tau^s_{jk'} w_j\right)^{-\theta^s} > w_h f^a_{hk'},$$

which depends not only on the assembly potential of k' vis à vis all markets i, but also on the overall global sourcing strategy of the firm. Focusing on the role of geography (i.e.,

[28] Although in Table 3.8 we document that firms import more from countries in which they have affiliates this is perhaps less surprising and driven by factors that are outside our model (such as vertical links between affiliates).

setting $T_k^a \left(w_k\right)^{-(1-\alpha)\theta^a} = G^a$, $T_j^s \left(w_j\right)^{-\theta^s} = G^s$), this condition reduces to

$$\kappa\varphi^{\sigma-1} G^a G^s \sum_{j \in \mathcal{J}_h(\varphi)} \sum_{i \in J} B_i T \left(\tau_{k'i}^a\right)^{-\theta^a} \left(\tau_{jk'}^s\right)^{-\theta^s} > w_h f_{hk'}^a,$$

which indicates that the addition of plant in country k' is more likely to be profitable if the firm has activated sourcing locations j that are relative close to country k'.

3.6.2 The Complementarity in Action: Two Illustrative Firm-Level Examples

We close our analysis with two examples to illustrate the implications of the key novel feature of our framework, namely the presence of firm-level economies of scale in global sourcing, in the sense that the global sourcing strategy is a firm-level one. This is in contrast to the approach in Bernard *et al.* (2018), in which each assembly plant is assumed to have to build its own set of plant-specific sourcing locations. As we have demonstrated in the previous section, this assumption is key for generating a complementarity between sourcing and assembly decisions. The goal of this section is to illustrate some consequences of this novel force. To do so, we will analyze how changes in tariffs on inputs and on final goods affect the optimal strategies of a specific firm, and will contrast the results we obtain under a firm-level global sourcing strategy versus a collection of plant-level global sourcing strategies.

Nonlinear Effects of Tariffs on Inputs

Consider a scaled down version of our model with only three countries: USA (us), China (ch), and Mexico (mex). Without loss of generality we normalize the US assembly and sourcing potentials to 1, so that $\xi_{us,us}^a = \xi_{us,us}^s = 1$. We also set $\kappa\varphi^{\sigma-1} = 1$ and $E_{us} P_{us}^{\sigma-1} = 1$. We focus on the optimal global production strategy of a firm headquartered in the US. We further dramatically simplify matters by assuming that (see also Figure

3.1):

1. The firm's goods are only demanded in the US, so $E_{mex}P_{mex}^{\sigma-1} = E_{ch}P_{ch}^{\sigma-1} = 0$.

2. The fixed costs of assembly and sourcing in the US are 0, so $\{US\} \in \mathcal{J}_h(\varphi)$ and $\{US\} \in \mathcal{K}_h(\varphi)$.

3. Mexico only has the capability of producing final goods, so $\xi_{mex,us}^a > 0$, but $\xi_{mex,us}^s = \xi_{mex,ch}^s = \xi_{mex,mex}^s = 0$).

4. China only has the capability of producing intermediate inputs, so $\xi_{ch,us}^s > 0$ and $\xi_{ch,mex}^s > 0$, but $\xi_{ch,us}^a = 0$.

5. The US does not export intermediate inputs to Mexico, or $\xi_{us,mex}^s = 0$.

6. Pairs of assembly locations and pairs of sourcing locations are independent, or $\sigma - 1 = \theta^a = \theta^s/\alpha$.

The last assumption serves to abstract from any interdependencies other than those generated by the complementarity between global sourcing and global assembly. In terms of this specific example, a key feature will be that the profitability of assembly in Mexico will be partly shaped by its access to inputs sourced from China, which is partly shaped by the sourcing potential of China vis à vis the US. To see this more formally, let us write the extensive margin problem of the firm as:

$$\max_{\left(\mathcal{I}_{ch}^s, \mathcal{I}_{mex}^a\right) \in \{0,1\}^2} \quad 1 + \mathcal{I}_{ch}^s \cdot \xi_{ch,us}^s + \mathcal{I}_{mex}^a \cdot \mathcal{I}_{ch}^s \cdot \xi_{mex,us}^a \cdot \xi_{ch,mex}^s - \mathcal{I}_{ch}^s \cdot f_{ch}^s - \mathcal{I}_{mex}^a \cdot f_{mex}^a.$$

Clearly, the firm only has three possible choices: activate both assembly in Mexico and sourcing from China, activate only sourcing from China, or activate neither.[29]

[29]Only activating assembly in Mexico is never optimal because without inputs from China, the Mexican plant cannot produce.

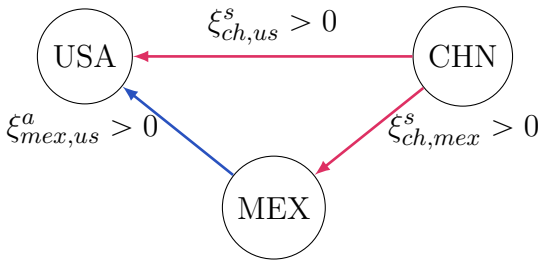


Figure 3.1: Trade Structure

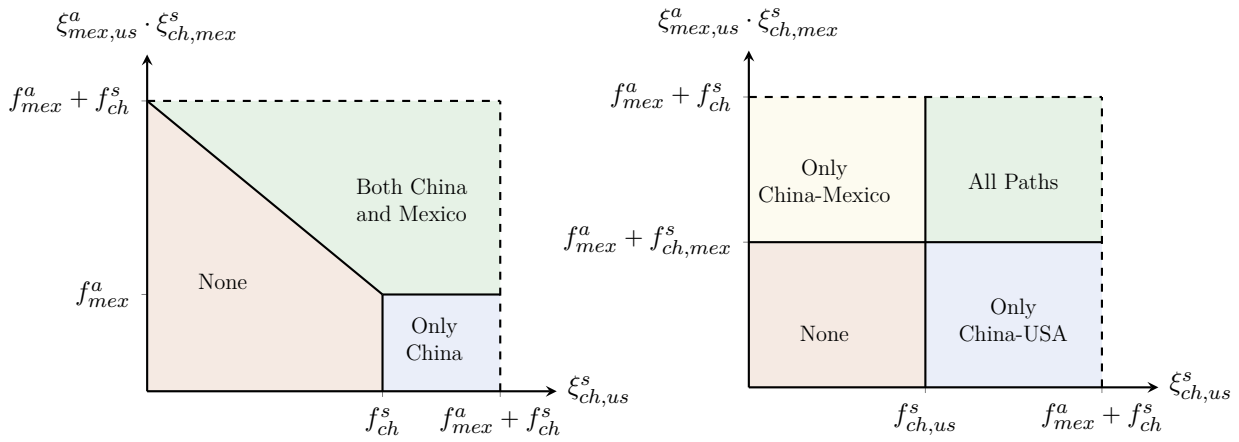
The solution to this problem is the following:

$$\begin{cases} \text{Activate Both if} & \xi^s_{ch,us} + \xi^a_{mex,us} \cdot \xi^s_{ch,mex} \geqslant f^s_{ch} + f^a_{mex} \\ & \text{and } \xi^a_{mex,us} \cdot \xi^s_{ch,mex} \geqslant f^a_{mex} \\ \text{Activate China Sourcing only if} & \xi^s_{ch,us} \geqslant f^s_{ch} \text{ and } \xi^a_{mex,us} \cdot \xi^s_{ch,mex} < f^a_{mex} \\ \text{Activate Neither if} & \xi^s_{ch,us} + \xi^a_{mex,us} \cdot \xi^s_{ch,mex} < f^s_{ch} + f^a_{mex} \\ & \text{and } \xi^s_{ch,us} < f^s_{ch}. \end{cases}$$

The solution is also illustrated in panel (a) of Figure 3.2a which shows that the assembly decision in Mexico ($\mathcal{I}^a_{mex}$) depends not only on the Mexican assembly potential ($\xi^a_{mex,us}$) and sourcing potential of Chinese inputs vis à vis Mexico ($\xi^s_{ch,mex}$), but also on the sourcing potential of Chinese inputs vis à vis the US ($\xi^s_{ch,us}$). In contrast, panel (b) of Figure 3.2b shows the equilibrium under the alternative assumption of plant-level sourcing strategies. Specifically, suppose all assembly plants have to pay plant-specific fixed costs ($f^s_{ch,us}$ and $f^s_{ch,mex}$) to activate those sourcing locations for each plant. In this case, the decision to activate an assembly plant in Mexico is independent of the value of inputs to the American plant ($\xi^s_{ch,us}$).[30]

The differences between firm-level and plant-level global sourcing strategies can be further illustrated in terms of their differential implications for the response of the firm

[30] More specifically, the firm chooses $\left(\mathcal{I}^s_{ch,us}, \mathcal{I}^s_{ch,mex}, \mathcal{I}^a_{mex}\right)$ to solve the problem $\max 1 + \mathcal{I}^s_{ch,us} \cdot \xi^s_{ch,us} + \mathcal{I}^a_{mex} \cdot \mathcal{I}^s_{ch,mex} \cdot \xi^a_{mex,us} \cdot \xi^s_{ch,mex} - \mathcal{I}^s_{ch,us} \cdot f^s_{ch} - \mathcal{I}^s_{ch,us} \cdot f^s_{ch} - \mathcal{I}^a_{mex} \cdot f^a_{mex}$.



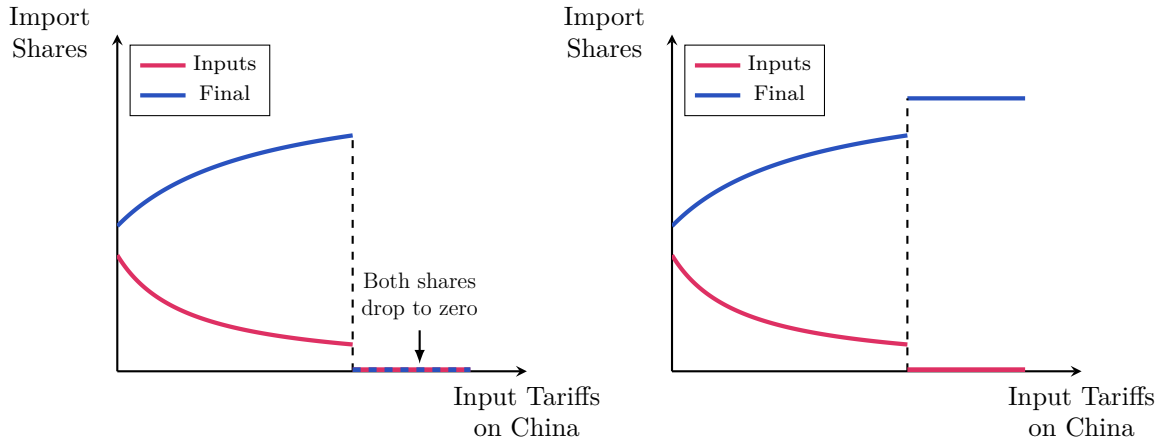
(a) Firm-level Sourcing Strategy **(b)** Plant-level Sourcing Strategy

Figure 3.2: Equilibrium Assembly and Sourcing Decisions

global production strategy to changes in tariffs. In particular, Figure 3.3 shows how unilateral tariffs on Chinese inputs set by the US affect the import share for inputs from China ($\frac{\xi^s_{ch,us}}{1+\xi^s_{ch,us}}$) and the import share for final goods from Mexico ($\frac{\xi^a_{mex,us} \cdot \xi^s_{ch,mex}}{1+\xi^s_{ch,usa}+\xi^a_{mex,us} \cdot \xi^s_{ch,mex}}$).

Under both firm- and plant-level global sourcing strategies, the share of Chinese inputs in the US gradually falls, and then after some point it drops to zero. However, for the Mexican assembly shares the response is quite different. The shares increase at first because the American plant becomes less productive as a result of the higher cost of Chinese inputs. Nevertheless, when tariffs on Chinese input are furthered raised, the firm may decide to shut down the Chinese sourcing plant (of course, the example assumes that the fixed cost f^s_{ch} is not sunk). Under firm-level global sourcing strategy, the loss of access to Chinese inputs leads to the closure of the Mexican assembly plant, and thus the import share of final goods drops to zero discontinuously. Under plant-level global sourcing strategies, the Mexican assembly plant continues to choose to source from China (since the cost of those inputs is unaffected by US tariffs), so in that case the Mexican import share goes *up* discontinuously.

In sum, the above example illustrates that firm-level economies of scale in sourcing



(a) Firm-level Sourcing Strategy **(b)** Plant-level Sourcing Strategy

Figure 3.3: The Effect of Tariffs

might well generate an endogenous complementarity between assembly plants, in the sense that a deterioration in the productivity of US final-good production leads to the closure of the foreign assembly plant. Conversely, under plant-level economies of scale we instead obtain an endogenous substitutability between assembly plants, in the sense that a deterioration in the productivity of US assembly translates into an increasing role of foreign assembly in US consumption.

Nonlinear Effects of Tariffs on Final Goods

We next consider an even simpler example with just two countries: USA (us) and China (ch). Without loss of generality we again normalize the US assembly and sourcing potentials to 1, so that $\xi^a_{us,us} = \xi^s_{us,us} = 1$, and we also set $\kappa\varphi^{\sigma-1} = 1$ and $E_{us}P^{\sigma-1}_{us} = 1$. We focus on the optimal global production strategy of a firm headquartered in the US. We further dramatically simplify matters by assuming that (see also Figure 3.4):

1. The firm's goods are only demanded in the US, so $E_{ch}P^{\sigma-1}_{ch} = 0$.

2. The fixed costs of assembly are zero in both countries, so $\mathcal{K}_h(\varphi) = \{US, Ch\}$.

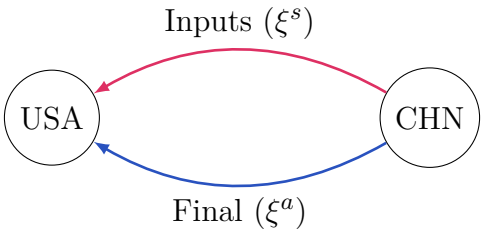


Figure 3.4: Trade Structure

3. The fixed costs of sourcing in the US are 0, so $\{US\} \in \mathcal{J}_h(\varphi)$.

4. The US does not export intermediate inputs to China, or $\xi^s_{us,ch} = 0$.

5. Pairs of assembly locations are substitutes $\sigma - 1 < \theta^a$, while pairs of sourcing locations are independent $\theta^a = \theta^s/\alpha$.

Under these assumptions, if the firm sets its global sourcing strategy at the firm level, then it solves the very simple problem

$$\max_{\mathcal{I}^s_{ch} \in \{0,1\}} \quad \left(1 + \mathcal{I}^s_{ch} \cdot \left[\xi^s_{ch,us} + \xi^a_{ch,us}\xi^s_{ch,ch}\right]\right)^{\frac{\sigma-1}{\theta^a}} - \mathcal{I}^s_{ch} \cdot f^s_{ch}, \tag{3.6.4}$$

and it generate a volume of sales in the US assembly plant that is proportional to :

$$\text{Sales of US Assembly Plant} \sim \left(1 + \mathcal{I}^s_{ch} \cdot \xi^s_{ch,us}\right) \cdot \left(1 + \mathcal{I}^s_{ch} \cdot \left[\xi^s_{ch,us} + \xi^a_{ch,us}\xi^s_{ch,ch}\right]\right)^{\frac{\sigma-1}{\theta^a}-1}. \tag{3.6.5}$$

The wage bill paid by the firm to US workers is also proportional to (3.6.5).

Assume the following initial conditions on parameters

$$\left(1 + \xi^s_{ch,us} + \xi^a_{ch,us}\xi^s_{ch,ch}\right)^{\frac{\sigma-1}{\theta^a}} > 1 + f^s_{ch} > \left(1 + \xi^s_{ch,us}\right)^{\frac{\sigma-1}{\theta^a}}, \tag{3.6.6}$$

so that activating China as a source of inputs is initially profitable, but if the assembly potential of China (vis à vis the US) is sufficiently deteriorated ($\xi^a_{ch,us} \to 0$), then activating China as a source of inputs is no longer profitable.

Consider now the implications of a unilateral increase in tariffs applied by the US on

imports of final goods from China. The immediate effect of this policy is to reduce the assembly potential $\xi^a_{ch,us}$ of China vis à vis the US. Figure 3.5a shows the response of US final-good sales (as well as the wage bill paid by the firm to US workers). Sales initially increase because tariffs make the Chinese assembly plant less competitive, and the US plant gains additional market share due to the substitutability implied by $\sigma - 1 < \theta^a$. However, after increasing tariffs beyond a certain threshold, at the resulting lower value of $\xi^a_{ch,us}$, the firm does not find it valuable to continue to activate China as a sourcing location. The discontinuation of that plant increases the marginal cost of the US plant on impact, leading to a discontinuous drop in its sales, profitability, and wage bill.[31]

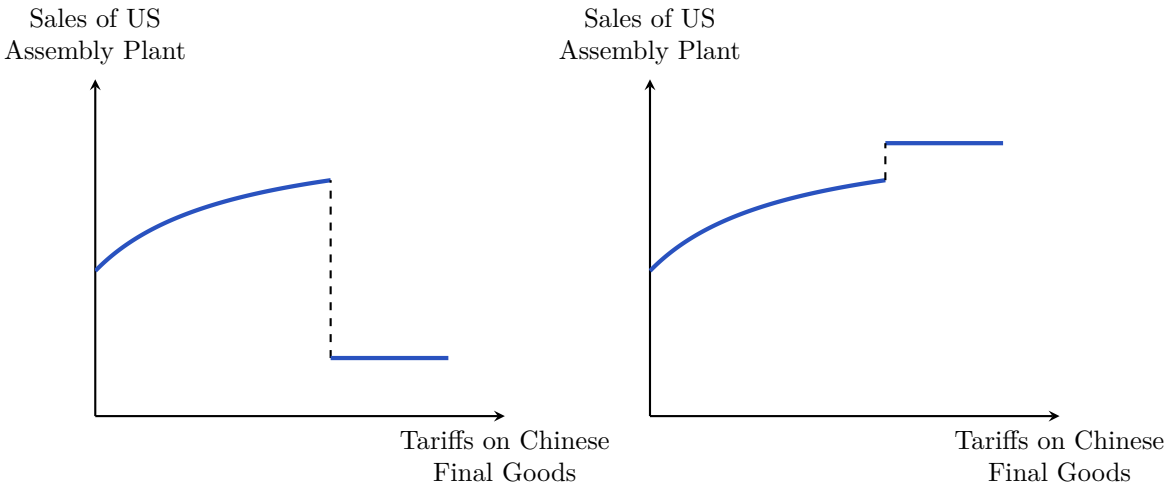
We can compare the above results to the case with plant-level global sourcing strategies. In that case, the firm can activate a sourcing plant in China specifically designed to sell inputs to the US plant, and another one designed to sell inputs to the Chinese assembly plant. The problem of the firm is then

$$\max_{(\mathcal{I}^s_{us,ch},\mathcal{I}^s_{ch,ch})\in\{0,1\}^2} \left(1 + \mathcal{I}^s_{us,ch}\cdot\xi^s_{ch,us} + \mathcal{I}^s_{ch,ch}\cdot\xi^a_{ch,us}\xi^s_{ch,ch}\right)^{\frac{\sigma-1}{\theta^a}} - \mathcal{I}^s_{ch,us}\cdot f^s_{ch} - \mathcal{I}^s_{ch,ch}\cdot f^s_{ch}. \tag{3.6.7}$$

Given the solution of this problem, it is then straightforward to show that an increase in tariffs on the imports of Chinese final goods (and associated fall in $\xi^a_{ch,us}$) can never decrease the sale revenue, operating profits and wage bill paid the US assembly plant.

Panel (b) of Figure 3.5b depicts the response of the size of the assembly plant in the US (and wage bill paid by the firm to US workers) to increases in final-good tariffs. Initially, sales goes up due to the tariffs faced by the Chinese plant and the substitution implied by $\sigma - 1 < \theta^a$. After a certain threshold, the firm does not find it valuable to assemble in China, and US sales discontinuously jump up (since they continue to source from China

[31]Figure 3.5a is generated under condition (3.6.6). We model the effect of tariffs on $\xi^a_{ch,us}$ as $\tilde{\xi}^a_{ch,us} = \xi^a_{ch,us}(1+t)^{-\theta^a}$. For a drop to be below the initial (pre-tariffs, $t = 0$) sales of the US assembly plant the following condition should be satisfied $\left(1 + \xi^s_{ch,us} + \xi^a_{ch,us}\xi^s_{ch,ch}\right)^{1-\frac{\sigma-1}{\theta^a}} < 1 + \xi^s_{ch,us}$.



(a) Firm-level Sourcing Strategy **(b)** Plant-level Sourcing Strategy

Figure 3.5: The Effect of Tariffs

but now face lower within-firm competition).[32]

3.7 Conclusion

Multinational firms are dominant players in domestic employment, output, and trade. Leveraging newly linked Bureau of Economic Analysis and US Census data, we have confirmed the quantitative importance of MNEs for the US economy, and have also unveiled the existence of a strong relationship between the importing decisions of firms operating in the US and their overall worldwide operations. More specifically, global sourcing is more prevalent for firms whose center of gravity, in terms of their global assembly strategies, is further away from the United States. Furthermore, even after controlling for firm and country fixed effects, US MNEs are significantly more likely to import from countries in which they have a majority-owned manufacturing affiliate (and

[32]The following numerical parameters can generate patterns in Figures 3.5a and 3.5b: $\xi^s_{ch,us} = \xi^s_{ch,us} = \xi^a_{ch,us} = 1$, $f^s_{ch} = 0.5$, $f^a_{us} = f^a_{ch} = 0.25$, $\sigma = \theta^a = 2$. Then, the discontinuity in Figure 3.5a happens at $t = 1$, and in Figure 3.5b at $t \approx 0.14$.

also from other countries in the same region), and foreign-owned MNEs are much more likely to import from their headquarter country (or other countries in the headquarters' region).

We have then developed a multi-country model in which firms decide on the location of their assembly plants (i.e., their assembly strategy) as well as the source of the inputs used in their plants worldwide (i.e., their global sourcing strategy). A key novel feature of our framework is the existence of firm-level economies of scale in the global sourcing strategies of firms. This delivers rich complementarities between the global sourcing and global assembly choices of firms, and constitutes a plausible mechanism to explain the observed complementarity between these two strategies we have documented in our reduced-form evidence. In addition to providing a useful lens through which to interpret the data, our framework delivers a rich set of comparative statics, and we have demonstrated via simple numerical analyses that it can produce surprising nonlinear responses to changes in tariffs on final goods and on inputs.

References

Acemoglu, D., Akcigit, U. and Kerr, W. (2016). Innovation Network. *PNAS*, **113** (41), 11483–11488.

Aghion, P., Akcigit, U. and Howitt, P. (2014). What Do We Learn from Schumpeterian Growth Theory,. *Handbook of Economic Growth*, **2B**, 515–563.

—, Bergeaud, A. and Van Reenen, J. (2021). The impact of regulation on innovation. *NBER Working Paper No. 28381*.

— and Bolton, P. (1987). Contracts as a Barrier to Entry. *American Economic Review*, **77** (4), 388–401.

— and Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, **60** (2), 323–51.

Aitken, B. J. and Harrison, A. E. (1999). Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela. *American Economic Reivew*, **89** (3), 605–618.

Akcigit, U., Hanley, D. and Stantcheva, S. (2021). Optimal Taxation and R&D Policies. *Econometrica*, **Fortcoming**.

— and Kerr, W. (2018). Growth Through Heterogenous Innovations. *Journal of Political Economy*, **126** (4), 1374–1443.

Alcácer, J., Gittelman, M. and Sampat, B. (2009). Applicant and Examier Citations in U.S. Patents: An Overview and Analysis. *Research Policy*, **38**, 415–427.

Alfaro-Ureña, A., Manelici, I. and Vasquez, J. P. (2021). The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages. *Working Paper*.

Almeling, D. S. (2012). Seven Reasons Why Trade Secrets Are Increasingly Important. *Berkeley Technology Law Journal*, **27** (1091), 1092–1118.

—, Snyder, D. W., Sapoznikow, M., McCollum, W. E. and Weader, J. (2010). A Statistical Analysis of Trade Secret litigation in Federal Courts. *Gonzaga Law Review*, **45** (2), 291–334.

—, —, —, — and — (2011). A Statistical Analysis of Trade Secret litigation in State Courts. *Gonzaga Law Review*, **46** (1), 57–101.

Amara, N., Landry, R. and Traoré, N. (2008). Managing the Protection of Innovations in Knowledge-Intensive Business Services. *Research Policy*, **37**, 1530–1547.

Anton, J. J., Greene, H. and Yao, D. A. (2006). Policy Implications of Weak Patent Rights. *Innovation Policy and the Economy, Volume 6, edited by Adam B. Jaffe, Josh Lerner, and Scott Stern*, pp. 1–26.

Antràs, P. and Chor, D. (2013). Organizing the global value chain. *Econometrica*, **81** (6), 2127–2204.

— and De Gortari, A. (2020). On the geography of global value chains. *Econometrica*, **88** (4), 1553–1598.

—, Fort, T. and Tintelnot, F. (2017). The Margins of Global Sourcing: Theory and Evidence from US Firms. *American Economic Review*, **107** (9), 2514–2564.

— and Helpman, E. (2004). Global sourcing. *Journal of political Economy*, **112** (3), 552–580.

— and Yeaple, S. R. (2014). Multinational firms and the structure of international trade. **4**, 55–130, from January 2013.

Arkolakis, C., Eckert, F. and Shi, R. (2021). Combinatorial Discrete Choice. *Working Paper*.

—, Ramondo, N., Rodríguez-Clare, A. and Yeaple, S. (2018). Innovation and Production in the Global Economy. *American Economic Review*, **108** (8), 2128–73.

Arora, A. (1997). Patents, Licensing and Market Structure in the Chemical Industry. *Research Policy*, **26** (4–5), 391–403.

—, Belenzon, S., Patacconi, A. and Suh, J. (2019). The Changing Structure of American Innovation: Some Cautionary Remarks for Economic Growth. *Innovation Policy and the Economy, Volume 20, edited by Josh Lerner, and Scott Stern*, pp. 39–93.

—, — and Sheer, L. (2021). Knowledge Spillovers and Corporate Investment in Scientific Research. *American Economic Reivew*, **111** (3), 871–898.

—, Fosfuri, A. and Gambardella, A. (2001). *Markets for Technology. The Economics of Innovation and Corporate Strategy*. The MIT Press.

Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Invention. *in R. Nelson, eds., The Rate and Direction of Inventive Activity. Princeton University Press, Princeton, NJ*, pp. 609–625.

Arts, S., Cassiman, B. and Gomez, J. C. (2018). Text Matching to Measure Patent Similarity. *Strategic Management Journal*, **39**, 62–84.

Autor, D., Dorn, D., Hanson, G. H., Pisano, G. and Shu, P. (2020). Foreign Competition and Domestic Innovation: Evidence from US Patents. *American Economic Review: Insights*, **2** (3), 357–374.

Autor, D. H., Dorn, D. and Hanson, G. H. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, **103** (6), 2121–2168.

Bai, J., Barwick, P., Cao, S. and Li, S. (2021). Quid Pro Quo, Knowledge Spillover, and Industrial Quality Upgrades: Evidence from the Chinese Auto Industry. *Working Paper*.

Barrot, J.-N. and Sauvagnat, J. (2016). Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks. *Quarterly Journal of Economics*, **131** (3), 1543–1592.

Bell, A., Chetty, R., Jaravel, X., Petkova, N. and van Reenen, J. (2019). Who Becomes an Inventor in America? The Importance of Exposure to Innovation. *Quarterly Journal of Economics*, **134** (2), 647–713.

Bernard, A. B. and Jensen, J. B. (1999). Exceptional exporter performance: cause, effect, or both? *Journal of international economics*, **47** (1), 1–25.

—, —, Redding, S. J. and Schott, P. K. (2007a). Firms in international trade. *Journal of Economic Perspectives*, **21** (3), 105–130.

—, —, — and — (2007b). Firms in international trade. *Journal of Economic Perspectives*, **21** (3), 105–130.

—, —, — and — (2009). The margins of us trade. *American Economic Review*, **99** (2), 487–93.

—, —, — and — (2018). Global firms. *Journal of Economic Literature*, **56** (2), 565–619.

Blalock, G. and Gertler, P. (2008). Welfare Gains from Foreign Direct Investment Through Technology Transfer to Local Suppliers. *Journal of International Economics*, **74** (2), 402–421.

Bloom, N., Hassan, T. A., Kalyani, A., Lerner, J. and Tahoun, A. (2021). The Diffusion of Disruptive Technologies. *NBER Working Paper No. 28999*.

—, Van Reenen, J. and Williams, H. (2019). A Toolkit of Policies to Promote Innovation. *Journal of Economic Perspectives*, **33** (3), 163–184.

Brainard, S. L. (1997). An empirical assessment of the proximity-concentration trade-off between multinational sales and trade. *The American Economic Review*, **87** (4), 520–544.

Buera, F. and Lucas, R. (2018). Idea Flows and Economic Growth. *Annual Review of Economics*, **10**, 315–45.

Caballero, R. and Jaffe, A. B. (1993). How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth. *NBER Working Paper No. 4370*.

Cassiman, B. and Veughelers, R. (2002). R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium. *American Economic Reivew*, **92** (4), 1169–1184.

Chattergoon, B. and Kerr, W. R. (2021). Winner Takes All? Tech Clusters, Population Centers, and the Spatial Transformation of U.S. Invention. *Research Policy*, **Forthcoming**.

Chodorow-Reich, G. (2014). The Employment Effects of Credit Market Disruptions: Firm-Level Evidence from the 2008-9 Financial Crisis. *Quarterly Journal of Economics*, **129** (1), 1–59.

Chow, M. C., Fort, T. C., Goetz, C., Goldschlag, N., Lawrence, J., Perlman, E. R., Stinson, M. and White, T. K. (2021). *Redesigning the Longitudinal Business Database*. Working Paper 28839, National Bureau of Economic Research.

Cohen, L. and Frazzini, A. (2008). Economic Links and Predictable Returns. *The Journal of Finance*, **63** (4), 1977–2011.

Cohen, W. M., Nelson, R. R. and Walsh, J. P. (2000). Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). *National Bureau of Economic Research Working Paper 7552*.

Correia, S. (2015). *Singletons, Cluster-Robust Standard Errors and Fixed Effects*. mimeo, Duke University.

Costinot, A., Vogel, J. and Wang, S. (2013). An elementary theory of global supply chains. *Review of Economic studies*, **80** (1), 109–144.

Duranton, G. and Puga, D. (2004). Micro-Foundations of Urban Agglomeration Economies. *Handbook of Regional and Urban Economics*, **4**, 2063–2117.

Eaton, J. and Kortum, S. (2002). Technology, geography and trade. *Econometrica*, **70** (5), 1741–1779.

Eeckhout, J. and Jovanovic, B. (2002). Knowledge Spillovers and Inequality. *American Economic Review*, **92** (5), 1290–1307.

Fadeev, E. (2022). Creative Construction: Knowledge Sharing in Production Networks. *Working Paper*.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, Inc., Third edn.

Fort, T. C., Pierce, J. R. and Schott, P. K. (2018). New perspectives on the decline of us manufacturing employment. *Journal of Economic Perspectives*, **32** (2), 47–72.

Freund, C. and Pierola, M. D. (2015). Export superstars. *The Review of Economics and Statistics*, **97** (5), 1023–1032.

Garetto, S., Oldenski, L. and Ramondo, N. (2019). *Multinational Expansion in Time and Space*. Working Paper 25804, National Bureau of Economic Research.

Gaubert, C. and Itskhoki, O. (2018). *Granular comparative advantage*. Tech. rep., National Bureau of Economic Research.

Goldman, E. (2014). Do We Need a New Federal Trade Secret Law. *Working Paper*.

Gomes-Casseres, B., Hagedoorn, J. and Jaffe, A. B. (2006). Do Alliances Promote Knowledge Flows? *Journal of Financial Economics*, **80** (1), 5–33.

Graham, S. and Vishnubhakat, S. (2013). Of Smart Phone Wars and Software Patents. *Journal of Economic Perspectives*, **27** (1), 67–86.

Grossman, G. M. and Helpman, E. (1991). Quality Ladders in the Theory of Growth. *Review of Economic Studies*, **58** (1), 43–61.

—, — and Szeidl, A. (2006). Optimal integration strategies for the multinational firm. *Journal of International Economics*, **70** (1), 216–238.

— and Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, **98** (5), 1978–97.

Gruber, J. and Johnson, S. (2019). *Jump-Starting America: How Breakthrough Science Can Revive Economic Growth and the American Dream*. PublicAffairs.

Hall, B. H., Helmers, C., Rogers, M. and Sena, V. (2014). The Choice between Formal and Informal Intellectual Property: A Review. *Journal of Economic Literature*, **52** (2), 375–423.

—, Jaffe, A. B. and Trajtenberg, M. (2005). Market Value and Patent Citations. *RAND Journal of Economics*, **36** (1), 16–38.

Handley, K., Kamal, F. and Monarch, R. (2020). *Rising Import Tariffs, Falling Export Growth: When Modern Supply Chains Meet Old-Style Protection*. Working Paper 26611, NBER.

Head, K. and Mayer, T. (2019). Brands in motion: How frictions shape multinational production. *American Economic Review*, **109** (9), 3073–3124.

—, — and Ries, J. (2010). The erosion of colonial trade linkages after independence. *Journal of International Economics*, **81** (1), 1–14.

Helpman, E. (1984). A simple theory of international trade with multinational corporations. *Journal of political economy*, **92** (3), 451–471.

— (1985). Multinational corporations and trade structure. *The Review of Economic Studies*, **52** (3), 443–457.

—, Melitz, M. J. and Yeaple, S. R. (2004). Export versus fdi with heterogeneous firms. *American economic review*, **94** (1), 300–316.

Henderson, R., Jaffe, A. B. and Trajtenberg, M. (2005). Patent Citations and the Geography of Knowledge Spillovers: A Reassessment: Comment. *American Economic Review*, **95** (1), 461–464.

Higham, K. W., de Rassenfosse, G. and Jaffe, A. B. (2020). Patent Quality: Towards a Systematic Framework for Analysis and Measurement. *NBER Working Paper No. 27598*.

Jaffe, A. B. and de Rassenfosse, G. (2019). Patent Citation Data in Social Science Research: Overview and Best Practices. *Research Handbook on the Economics of intellectual Property Law*.

— and Trajtenberg, M. (2002). *Patents, citations, and innovations: A window on the knowledge economy*. MIT Press.

—, — and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics*, **108** (3), 577–598.

Jarmin, R. S. and Miranda, J. (2002). *The Longitudinal Business Database*. CES Working Paper 02-17.

Jia, P. (2008). What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry. *Econometrica*, **76** (6), 1263–1316.

Jones, C. I. (2005). Growth and Ideas. *Handbook of Economic Growth, Volume 1B, edited by Philippe Aghion and Steven N. Durlauf*.

Jorda, K. F. (2008). Patent and Trade Secret Complementariness: An Unsuspected Synergy. *Washburn Law Journal*, **48** (1), 1–32.

Kamal, F., McCloskey, J. and Ouyang, W. (2021). *Multinational Firms in the US Economy: Insights from Newly Integrated Microdata*. mimeo, US Census Bureau.

— and Ouyang, W. (2020). *Identifying US Merchandise Traders: Integrating Customs Transactions with Business Administrative Data*. CES Working Paper 20-28.

Kogan, L., Papanikolaou, D., Seru, A. and Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth. *Quarterly Journal of Economics*, **132** (2), 665–712.

Krugman, P. (1991). *Geography and Trade*. MIT Press.

Kuhn, J., Younge, K. and Marco, A. (2020). Patent Citations Reexamined. *RAND Journal of Economics*, **51** (1), 109–132.

Lerner, J. and Seru, A. (2021). The Use and Misuse of Patent Data: Issues for Corporate Finance and Beyond. *Review of Financial Studies*, **Fortcoming**.

Levine, D. S. and Seaman, C. B. (2018). The DTSA at One: En Empirical Study of the First Year of litigation Under the Defend Trade Secrets Act. *Wake Forest Law Review*, **53**, 105–156.

Lim, K. (2009). The Many Faces of Absorptive Capacity: Spillovers of Copper Interconnect Technology for Semiconductor Chips. *Industrial and Corporate Change*, **18** (6), 1249–1284.

— (2017). Endogenous Production Networks and the Business Cycle. *Working Paper*.

Manski, C. F. and Lerman, S. R. (1977). The Estimation of Choice Probabilities From Choice Based Samples. *Econometrica*, **48** (8), 1977–1988.

Markusen, J. R. (1984). Multinationals, multi-plant economies, and the gains from trade. *Journal of international economics*, **16** (3-4), 205–226.

Marx, M. and Fuegi, A. (2020). Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles. *Strategic Management Journal*, **41** (9), 1572–1594.

Mayer, T. and Ottaviano, G. (2007). The happy few: new facts on the internationalisation of european firms. *Bruegel-CEPR EFIM2007 Report, Bruegel Blueprint Series*, (3).

Mazzoleni, R. and Nelson, R. R. (1998). The Benefits and Costs of Strong Patent Protection: a Contribution to the Current Debate. *Research Policy*, **27** (3), 273–284.

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

Menell, P. S., Lemley, M. A., Merges, R. P. and Balganesh, S. (2021). *Intellectual Property in the New Technological Age: 2021*. Caluse 8 Publishing.

Moretti, E. (2012). *The New Geography of Jobs*. Houghton Mifflin Harcourt.

— (2021). The Effect of High-Tech Clusters on the Productivity of Top Inventors. *American Economic Review*, **111** (10), 3328–75.

Ouellette, L. L. (2012). Do Patents Disclose Useful Information. *Harvard Journal of Law and Technology*, **25** (2), 545–607.

Parker III, W. S. and Justice, D. D. (2014). The Differing Approaches To Preemption Under The Uniform Trade Secrets Act. *Tort Trial & Insurance Practice Law Journal*, **49** (2), 645–665.

Perla, J. and Tonetti, C. (2014). Equilibrium Imitation and Growth. *Journal of Political Economy*, **122** (1), 52–76.

Price II, N., Rai, A. K. and Minssen, T. (2020). Knowledge transfer for large-scale vaccine manufacturing. *Science*, **369** (6506), 912–914.

Ramondo, N., Rappoport, V. and Ruhl, K. J. (2016). Intrafirm trade and vertical fragmentation in us multinational corporations. *Journal of International Economics*, **98**, 51–59.

— and Rodríguez-Clare, A. (2013). Trade, multinational production, and the gains from openness. *Journal of Political Economy*, **121** (2), 273–322.

Rey, P. and Tirole, J. (2007). A Primer on Foreclosure. *in M. Armstrong and R. Porter, eds., Handbook of Industrial Organization*, **3**, 2145–2220.

Roach, M. and Cohen, W. M. (2013). Lens or Prism? Patent Citations as a Measure of Knowledge Flows from Public Research. *Mangement Schience*, **59** (2), 504–525.

Roin, B. N. (2005). The Disclosure Function of the Patent System (Or Lack Thereof). *Harvard Law Review*, **118** (6), 2007–2028.

Romer, P. (1990). Endogenous Technical Change. *Journal of Political Economy*, **98** (5), S71–S102.

Rosenberg, N. (1963). Technological Change in the Machine Tool Industry, 1840–1940. *The Journal of Economic History*, **23** (4), 414–443.

Scotchmer, S. (2004). *Innovation and Incentives*. The MIT Press.

Segal, I. (1999). Contracting with Externalities. *The Quarterly Journal of Economics*, **114** (2), 337–388.

Singh, J. and Marx, M. (2013). Geographical Constraints on Knowledge Spillovers: Political Borders vs. Spatial Proximity. *Management Science*, **59** (9), 2056–2078.

Thompson, P. and Fox-Kean, M. (2005). Patent Citations and the Geography of Knowledge Spillovers: A Reassessment. *American Economic Review*, **95** (1), 450–460.

Tintelnot, F. (2017). Global Production with Export Platforms. *Quarterly Journal of Economics*, **132** (1), 157–209.

Trajtenberg, M. and Jaffe, A. B. (2002). *Patents, Citations, and Innovations. A Window on the Knowledge Economy*. MIT Press.

Wang, Z. (2019). *Headquarters Gravity: How Multinationals Shape International Trade*. Tech. rep., FREIT working paper.

Whinston, M. D. (2006). *Lectures on Antitrust Economics*. The MIT Press.

Yeaple, S. R. (2003). The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment. *Journal of International Economics*, **60** (2), 293–314.

Appendix A

Appendix to Chapter 1

A.1 Case Studies: Combining Secrecy and Patenting

On May 5th 2021, the U.S. administration announced that it would support the temporary waiver of IP rights on messenger RNA technology for Covid-19 vaccines. The announcement generated a lot of debates whether this policy can increase the production of vaccines. Many IP lawyers and scholars argue that one of the problems is that patents do not disclose enough information for the replication of the mRNA technology.[1]

> " 'A waiver helps to keep generic manufacturers safe from patent litigation. But they won't even get to that stage without the cooperation of the inventors.' On top of that, not all that you need to produce these vaccine generics is patented and, thus, disclosed in the patent application. Much is protected against competitors not via patent law but by keeping it secret. 'You can't force the company that hold these secrets to pass it on to you.' "

The practice of combining trade secrets and patents in chemical innovations has a long history. Arora (1997) describes how dyestaff producers in the first half of the 20th century

[1]The quote is taken from the interview with Jayashree Watal, a professor at the Georgetown University School of Law in Washington D.C., who worked for more than three decades at the WTO secretariat ("Three Crises and One Waiver", Verfassungsblog, May 7th, 2021. For additional discussions, see also "The COVID-19 Vaccine Patent Waiver: The Wrong Tool for the Right Goal", Bill of Health, Petrie-Flom Center at Harvard Law School, May 5h, 2021.

patented codified individual chemical compounds but kept tacit knowledge on how to combine these compounds secret. The same approach was used by ammonia producers:

> "The Haber-Bosch process for ammonia, a truly significant process innovation, was protected by more than 200 patents that covered the apparatus, temperatures, and pressures, but avoided particulars about the catalysts employed or their preparation. The catalyst was critical to the successful operation of the process, and keeping it secret significantly increased the expense and time for firms trying to circumvent the Haber-Bosch patent ..."

Combining secrecy with patents seems to be inconsistent with the disclosure requirements of patents. For example, inventors should disclose their preferred method for carrying out the invention ("best mode") in order to "restrain inventors from applying for patents while at the same time concealing from the public preffered embodiments of the inventions they have in fact conceived."[2] However, given high uncertainty about the limits of this requirement firms try minimize the amount of disclosed knowledge. For instance, in *Fonar Corp vs. Gen. Elec. Co.* case (software) inventors did not disclose their source code, and in *Amgen, Inc. vs. Chungai Pharm. Co.* case (biotech) inventors did not disclose the specific cell lines used in their products. In both cases, courts supported the inventors. Jorda (2008) provides a general discussion on the limits of the "best mode" requirement from a legal point of view. For example, it applies only to the knowledge that inventors had at the time of patent filing. Given that patents are often filed at the early stage of research, preferred embodiments are often discovered later. In the case *C&F Packing Co. v. IBP, Inc.*, C&F had "developed a process for making and freezing a precooked sausage for pizza toppings" that was superior to existing technologies. C&F got two patents: one on the equipment and another on the process itself. After that they continued to improve the technology but kept it secret. C&F shared these secrets under a confidentiality agreement with a supplier who leaked them to its customer, Pizza Hut.

[2]See *Teflex, Inc. v. Ficosia N. Am. Corp.*

The court ordered Pizza Hut to pay 10.9$ million to C&F for trade secret misappropriation.

A.2 Data Appendix

A.2.1 Data Details

Patents. The main source of patent data is PatentsView. Kogan *et al.* (2017) and Autor *et al.* (2020) provide a matching of patent assignees to Compustat and CRSP firm names for publicly traded firms. The main problem with matching assignee names to other data sets is that names in patent files are not standardized: the name of the same company can be written in different ways and with typos. I use the existing matching of assignee names to Computstat firms for the period 1976–2014 in Autor *et al.* (2020) to extend it for years up to 2019. For the rest of the patents, I follow the procedure outline in Autor *et al.* (2020) for cleaning and standardizing firm names (e.g., replace "Incorporated" with "INC"). Finally, I matched around 100 thousand patents manually for the largest assignees.

Suppliers and Customers. I take the supplier-customer data from Compustat Segments data set. For publicly traded firms, it lists names of the main customers, which are mostly other firms but can also be government agencies. Regulation SFAS No. 131 requires publicly traded firms to report the identity of any customer representing more than 10% of their total sales. Using Compustat Segments, Barrot and Sauvagnat (2016) constructed a data set of suppliers and customers for the U.S. publicly traded firms for the period 1976–2013. I extend their data up to 2019 by manual inspection of Compustat Segments data.

A.2.2 Details to Section 2: Concentration of Patent Citations

Section A.2.2 provides details on the construction of the concentration measure. Section A.2.2 gives details on the Monte-Carlo exercise that equates citation rates across firms. Section A.2.2 discusses the Monte-Carlo exercise that equates citation rates and the number of patents across firms across firms. Section A.2.2 gives details on how to identify "specific" patents that are used in Section 1.5.1. Section A.2.2 provides details on the concentration of citations for scientific publications.

Construction of the Concentration Measure

The concentration measure from Figure 1.1 is constructed in the following way. First, I identify top 5% of the most cited patents within each grant year and technology class. Second, for these patents I compute the share of citations coming from the most citing firm. Finally, I aggregate these measures within and across technology classes.

The first step is to identify top 5% of the most cited patents. Denote $y_{km} = 1$ if patent $m \in \mathcal{P}$ makes a citation to patent $k \in \mathcal{P}$ and $y_{km} = 0$ otherwise, where $\mathcal{P}$ is the set of all granted patents. Each patent has an assignee (owner) or in rare cases (around 3%) multiple assignees. For the majority of patents, the assignee is a corporate firm but it can also include universities, government agencies, and individual inventors. In the second step, I will compute the distribution of citations across different organizations, so I exclude citations from individual inventors and patents with missing assignee information.[3] Each patent $k \in \mathcal{P}$ has a grant year t_k^g and a primary technology class c_k. I define the technology class at the group level in the Cooperative Patent Classification (primary class is a class listed first in the patent file). Denote the set of all groups in Cooperative Patent Classification (CPC) system by $\mathcal{CPC}$. For each patent $k \in \mathcal{P}$, I compute the number of

[3]Formally, I set $y_{km} = 0$ where patent m belongs to an individual inventor and does not have have an assignee information.

citations within a 5-year window defined by grant years

$$n_k = \sum_{m \in \mathcal{T}_k} y_{km} \text{ where } \mathcal{T}_k = \{m \in \mathcal{P} : \text{Grant Date}_m - \text{Grant Date}_k \leqslant 5 \cdot 365 \text{ Days}\} \tag{A.2.1}$$

For each grant year t and technology class c, I define the set of all granted patents receiving at least one citation

$$\Omega_{t,c} = \left\{k \in \mathcal{P} : t_k^g = t \text{ and } c_k = c \text{ and } n_k > 0\right\}$$

and take top 5% of patents in terms of the number of citations within this set. Denote it by $\Omega_{t,c}^{top}$, and define the *"Main"* sample as top patents for all years and technology classes:

$$Main = \left\{\Omega_{t,c}^{top}\right\}_{c \in \mathcal{CPC} \text{ and } t=1976...2014}$$

In the second step, for each patent in the *Main* sample I compute the share of citations coming from the most citing organization. To account for patents with multiple assignees, I define a weighted citation as $y_{km}^w = y_{km}/F_m$ where F_m is the number of assignees for patent m. I also group all universities as one assignee, all foreign government agencies as one assignee, and all U.S. government agencies as one assignee. The results are almost identical if I do not group them or exclude citations from these organizations at all because the concentration of citations is driven by corporate patents (see Figure A.6).[4] Define the number of citations to patent $k \in \mathcal{P}$ from organization i as

$$n_{k,i} = \sum_{m \in i, m \in \mathcal{T}_k} y_{km}^w$$

where $m \in i$ means that organization i is an assignee for patent $m \in \mathcal{P}$. Then, the

[4]The logic for the main specification is the following. I leave citations from non-corporate organizations in order to capture the most cited patents. I group universities, foreign governments, and the U.S. government in three groups because citations from these organizations are unlikely to reflect sharing of trade secrets. Figure A.7 shows that the concentration of citations is robust if I do not group these organizations together.

concentration measure for patent $k \in \mathcal{P}$ is

$$\mathcal{C}_k = \max_i \left\{ \frac{n_{k,i}}{n_k} \right\}$$

In the third step, I aggregate these measures. Specifically, within each grant year (t) and technology class (c) I compute a simple average across patents[5]

$$\mathcal{C}(t,c) = \frac{1}{|\Omega_{t,c}^{top}|} \sum_{k \in \Omega_{t,c}^{top}} \mathcal{C}_k \tag{A.2.2}$$

where $|\Omega_{t,c}^{top}|$ is the number of patents in $\Omega_{t,c}^{top}$. Then I aggregate across technological classes using the weighted average of $\mathcal{C}(t,c)$ where weights are defined by the number of patents in each $\Omega_{t,c}^{top} \neq \varnothing$

$$\mathcal{C}(t) = \sum_{c \in \mathcal{CPC}} \frac{|\Omega_{t,c}^{top}|}{\sum_{c \in \mathcal{CPC}} |\Omega_{t,c}^{top}|} \mathcal{C}(t,c) \tag{A.2.3}$$

The variable $\mathcal{C}(t)$ for $t = 1976 \ldots 2014$ is shown in Figure 1.1.

Monte-Carlo Simulations to Equalize Citation Rates Across Firms

The details of the Monte-Carlo simulations that equate citation rates across firms are the following. First, for each patent from the *Main* sample I take the citing patents and identify control patents that have the same application year, geographical location of inventors, and detailed technological class. Second, for each patent I randomize citation links across citing and control patents. Then for each patent I compute the concentration measure on the simulated sample. Third, I repeat this procedure 300 times to construct the distribution of concentration measures. Finally, I aggregate different moments of this distribution in a way similar to the actual concentration measure.

In the first step, I divide all granted patents into disjoint groups based on common

[5]The results are robust if instead of a simple average I use a citation-weighted average, or median instead of an average.

observational characteristics. Each patent k has an application year t_k^a, a geographical location of the majority of inventors ℓ_k, and the detailed technological class $\tilde{c}_k$. I define the geographical location at the state level if an inventor is located in the U.S., and at the country level if an inventor is located outside the U.S. For example, the location for an inventor living in Cambridge, MA, USA is (USA, MA), and for an inventor living in Berlin, Germany is *Germany*. If a patent has several inventors in different locations, I define the location for a patent based on the location of the majority of inventors. In the case of a tie, I take the location based on the alphabetical order. For a detailed technology class, I take the main subgroup level in CPC, which has more than 7000 categories. It is nested within a group level, $\widetilde{\mathcal{CPC}} \subset \mathcal{CPC}$ where $\widetilde{\mathcal{CPC}}$ is the set of all main subgroups in CPC. I define the characteristics as a triple $h = (t^a, \ell, \tilde{c})$, and each patent is assigned to a unique group of characteristics. Define the set of all characteristics as

$$\mathcal{H} = \Big\{ h = (t^a, \ell, \tilde{c}) : t^a \in \text{All Observed Application Years}, \\ \ell \in \text{All Observed Locations}, \tilde{c} \in \widetilde{\mathcal{CPC}} \Big\}$$

For each $h = (t^a, \ell, \tilde{c}) \in \mathcal{H}$, denote the group of patents with these characteristics by $\mathcal{P}(h)$ such that

$$\mathcal{P} = \bigcup_{h \in \mathcal{H}} \mathcal{P}(h) \text{ and } \mathcal{P}(h) \cap \mathcal{P}(h') = \varnothing \text{ for } h \neq h'$$

For each $k \in Main$, I take the set of citing patents

$$Citing_k = \{ m \in \mathcal{P} : y_{km} = 1 \text{ and } m \in \mathcal{T}_k \} \tag{A.2.4}$$

Denote the number of citations to patent k from patents with characteristics h by $n_k(h)$. For each $h \in \mathcal{H}_k$, I equate citation rates across firms. Formally, for each $h \in \mathcal{H}_k$ I randomize $n_k(h)$ citations across all patents that have characteristics h and satisfy the 5-year window time constraint. Denote the set of these patents by $\mathcal{P}_k(h)$:

$$\mathcal{P}_k(h) = \{ m \in \mathcal{P}(h) : m \in \mathcal{T}_k \}$$

where $\mathcal{T}_k$ is defined in (A.2.1). The Monte-Carlo exercise takes $n_k(h)$ random patents from the set $\mathcal{P}_k(h)$ and assigns citations to these patents. Under this randomization patents in $\mathcal{P}_k(h)$ from different firms make a citation to patent k with the same probability

$$p_k(h) = \frac{n_k(h)}{|\mathcal{P}_k(h)|}$$

I repeat the randomization procedure 300 times, and each time I compute the counterfactual concentration of citations for patent k. Denote the concentration measure for patent k in round s by $\mathcal{RC}_{k,s}$. I compute the median and the 95th quantile based on the distribution $\left\{\mathcal{RC}_{k,s}\right\}_{s=1}^{300}$, which are denote by $\mathcal{RC}_k(q95)$ and $\mathcal{RC}_k(q50)$. Then I aggregate these measures across patents in the same way as with the actual concentration, see equations (A.2.2) and (A.2.3).

Monte-Carlo Simulations to Equalize Citation Rates and Patenting Across Firms

The Monte-Carlo exercise that equates both citation probabilities and patenting is similar to the one from Section A.2.2 except the details on the randomization of citations. In the exercise that equates citation probabilities only, $n_k(h)$ citations are allocated randomly across patents in $\mathcal{P}_k(h)$. Therefore, firms that have more patents in $\mathcal{P}_k(h)$ are more likely to cite patent k. To equate patenting, I assume that $n_k(h)$ citations are allocated randomly to firms with the same probability. Formally, denote by $\mathcal{F}_k(h)$ the set of firms that have at least one patent in $\mathcal{P}_k(h)$. Each citation out of $n_k(h)$ is randomly allocated to firm $j \in \mathcal{F}_k(h)$ with probability

$$\frac{1}{|\mathcal{F}_k(h)|}$$

where $|\mathcal{F}_k(h)|$ is the number of firms in $\mathcal{F}_k(h)$.

Identifying "Specific" Patents

In Sections 1.5.1 and A.2.7, I identify patents where inventors were likely to have access to some private knowledge relative to others. Formally, the procedure is the following. Monte-Carlo simulations from Section A.2.2 construct the distribution of counterfactual concentration measures $\left\{\mathcal{C}_{k,s}\right\}_{s=1}^{300}$ for each patent $k \in Main$. I call citing patents $m \in Citing_k$ "specific" to patent $k \in Main$ if they are responsible for the majority of citations from a particular firm, and the concentration measure for patent $i \in Main$ exceeds the $95th$ quantile in the Monte-Carlo simulations. Formally,

$$m \in Citing_k \text{ is "specific" to } k \in Main \Leftrightarrow m \in i = \arg\max_j \left\{ \frac{n_{k,j}}{n_k} \right\} \text{ and } \mathcal{C}_k > \mathcal{RC}_k(q95)$$

where $m \in i$ means that patent m is assigned to firm i, and $\mathcal{RC}_k(q95)$ denotes the 95th quantile of the distribution $\left\{\mathcal{C}_{k,s}\right\}_{s=1}^{300}$.

Concentration for Scientific Publications

Figure 1.2 compares the concentration of citations for patents and scientific publications. To construct the concentration measure for publications, I use Microsoft Academic Graph data (Marx and Fuegi (2020)). The concentration measure is constructed in a way similar to the measure for patents. For each year between 1976 and 2014 and each scientific discipline in Web of Science classification, I track citations within a five-year window for the top 5% of the most cited publications. For each patent, the concentration is defined as a share of citations coming from the most citing organization. These organizations are predominantly universities. There are around 2 million most cited publications, and to reduce the computational burden I take a random 10% sample of publications from the set of the most cited ones. Then, I take the average of the concentration measures within a field of science, and the average across fields of science weighted by the number of publications in each field. The purple squared line in 1.2 shows the average concentration

measure for the most cited publications.

A.2.3 Movement of Inventors and Citation Patterns

To distinguish whether the concentration of citations is driven by firms or inventors, I track citations of inventors who worked for multiple companies. I compute the concentration measure similar to the one in Section 1.2.1 but within inventors-movers, and then I do the decomposition of the concentration measure similar to the one in Section 1.2.2. This exercise follows the same procedure as the Monte-Carlo exercises in Sections A.2.2 and A.2.2 except that the sample is restricted to inventors who worked in multiple companies, and citations are randomized within an inventor.

To increase the sample size, I consider all citations rather than the ones within a five year window. As a result, I consider the trend in citation patterns for cited granted patents until 2009, so that they have 10 years to accumulate citations. I also focus on the sample of patents granted to publicly listed firms in Compustat.[6] Moreover, I exclude patents assigned to multiple companies because it is impossible to distinguish which company an inventor represents.

For each patent, I compute the distribution of citations across inventors. I leave only patents that received at least 20 citations from one inventor. The results are robust to other thresholds. This is done in order to ensure greater variability in the concentration measure. For example, if an inventor cited a patent only one time, then this patent would always receive a citation from one firm only, and the within-inventor concentration measure would always be 100%. For each citing patent, I find all patents that were filed by the same inventor in the same U.S state or foreign country and the same main subgroup category in

[6]Matching of patents to Compustat firms is cleaner in a sense that I use the data from Autor *et al.* (2020) to control for potential subsidiary-parent relationship. If a patent is granted to a subsidiary of a certain firm, I match it to the parent company. Therefore, when the same inventor has patents in two firms in Compustat, these firms are more likely to represent different organizations relative to cases where the same inventor has patents in two private firms or foreign firms not listed in the U.S.

Cooperative Patent Classification system.[7] Patents should also be applied in the same time period: I find all patents applied in the period $[t_j^a, t_j^a + 2]$ where t_j^a is the application year of the citing patent. Then, I equate citation rates across firms by randomizing citations within each inventor across all these patents with similar characteristics: citing ones and control patents that are observationally similar to the citing ones. The procedure is the same as in Section A.2.2. I remove citing patents where no inventor worked in at least two companies and filed for similar patents. To equate citation rates and patenting, I randomize citations across firms that had observationally similar patents filed by the same inventor. The procedure is the same as in Section A.2.2.

The final data set is the following. Each cited patent has at least one citing inventor who filed similar patents in multiple firms. I compute the actual concentration of citations within each of these citing inventors (if there are many). Next, I compute the same concentration in Monte-Carlo simulations where citations are allocated randomly. For each cited patent, I take the average of the concentration measures across all citing inventors-movers. This gives a within-inventor actual and counterfactual average concentrations of citations for each cited patent. Then, I aggregate within and across technology classes in a way similar to Section 1.2. Figure 1.3 shows the results. The average within-inventor concentration measure is significantly higher relative to the 95th quantile of the same measure in Monte-Carlo simulations. This means that citations are driven by firms rather than inventors: inventors tend to cite different patents in various companies despite doing similar technologies.

A.2.4 Robustness to Section 2: Concentration of Citations

Figure A.6 shows that the concentration of citations is similar if we restrict the sample to corporate patents only. I also consider different thresholds for the most cited patents:

[7] If there are several inventors in the citing patent, I do this procedure for each of them.

top 10% and top 1%. Finally, I exclude the sample of citing patents that are assigned to superstar firms. Specifically, in each year and group level in Cooperative Patent Classification I find top 1% of firms in terms of the number of patents, and exclude their patents from the sample of citing patents. Figure A.7 shows that the results are robust if one uses the citation-weighted average or the median instead of the average to aggregate concentration measures within technological classes. I also plot on this graph the concentration measure when the non-corporate sector is disaggregated into different organizations rather than types of organization. Figure A.8 shows that the results are the same when I exclude self-citations of firms to itself, so the concentration is driven by citations between firms rather than self-citations. Kuhn *et al.* (2020) argue that the quality of citations as a measure of knowledge flows has declined over time due to a small number of patents responsible for a large share of backward citations. Figure A.8 shows that the results on the concentration are robust when I exclude top 1% of patents in terms of the number of backward citations. Finally, I exclude citations between patents sharing a common law firm to ensure that the concentration is not driven by lawyers citing themselves. I also group citations from patents from the same within-country family (continuations, continuations-in-part, divisionals) as a single citation. This ensures that the rise in concentration is not driven by increasing patent families. Finally, for the period after 2001 I separate citations made by patent examiners and non-examiners. Figure A.10 shows the concentration based on citations from patent examiners is around two times lower than the concentration based on citations from non-examiners.[8]

I also check whether citations are not driven by lawyers. Specifically, I track citations of lawyers who worked in multiple firms similar to the movement of inventors in Section A.2.3. The only difference is that there is no data for the location of lawyers, so I consider patents filed by the same lawyer in at least two companies, that have a similar application

[8]The USPTO started to separate examiner and non-examiner citations only around 2001.

period and are classified to the same main subgroup in Cooperative Patent Classification system. Figure A.11 shows that the actual concentration of citations across firms is around 95% within lawyers who represented similar companies. It is significantly higher relative to the 95th quantile of the concentration measure where citation rates are equated across companies within a lawyer.

A.2.5 Details to Section 4.1: Citations in a Production Network

Section 1.4.1 estimates the probability of a citation between two patents as a function of the network relationship between firms that own them. The details are the following. I impose the following restrictions on the set of all citations between patents granted between 1976 and 2019. First, since the supplier-customer data is available for the U.S. publicly traded firms only, I leave only cited and citing patents that are owned by a publicly traded U.S. firm. Second, to control for truncation that older patents have more time to accumulate citations I consider only citations within a 10 year window. Third, to focus on citations between firms I exclude self-citations. Finally, I remove cited patents granted after 2014 to allow patents to accumulate some citations.[9] Formally, denote by $y_{km} = 1$ if patent m makes a citation to patent k and $y_{km} = 0$ otherwise. Denote the grant year of patent k by t_k^g, the assignee of a patent k by f_k. I focus on the following set of cited-citing pairs

$$\text{Cited-Citing Set} = \Big\{(k,m) \in \mathcal{P}^2 : y_{km} = 1, \text{Grant Date}_m - \text{Grant Date}_k \leqslant 10 \cdot 365, \\ t_k^g \leqslant 2014, f_k \neq f_m\Big\}$$

[9] In all regression specifications, I use year fixed effects, which also control for differences in citations for patents granted in different years. I take the patents until 2014 in order to keep the sample of patents used for the computation of the concentration measure. I consider citations within a 10 year window (rather than 5) in order to increase the sample size. The results are robust when I restrict citations to be within a 5-year window, or take granted patents until 2009.

Next, following Jaffe *et al.* (1993), Thompson and Fox-Kean (2005), Singh and Marx (2013) I construct the control set of patents that have similar characteristics to citing ones. For each citing patent, I define the control set as all patents with the same application year, from the same main subgroup CPC class, and coming from a different firm. Formally, denote by t_k^a the application year of patent k, and by $\tilde{c}_k$ the main subgroup CPC class of patent k. Then the cited-control set for a pair $(k, m) \in$ Cited-Citing Set is

$$\text{Control Set}_{k,m} = \{r \in \mathcal{P} : t_r^a = t_m^a, \tilde{c}_r = \tilde{c}_m, f_r \neq f_m, f_r \neq f_k\}$$

I randomly choose one patent from $\text{Control Set}_{k,m}$, denote it by $r(k, m)$. Then the cited-control set is

$$\text{Cited-Control Set} = \Big\{(k, r) \in \mathcal{P}^2 : \text{ there is } m \in \mathcal{P} \text{ such that } (k, m) \in \text{Cited-Citing Set}, \text{ and } r = r(k, m)\Big\}$$

The estimation of (1.4.1) is done on the following sample of patent pairs

$$\text{Sample} = \text{Cited-Citing Set} \cup \text{Cited-Control Set}$$

For estimation of (1.4.1) I use the data on supplier-customer relationship. The details of the construction of this data are given in appendix A.2.1. Denote the first year when firm A reports firm B as a customer by t_{first}, and the last year by t_{last}. In rare cases, there might be gaps in years when A reports B as a customer, but this is likely to be caused by the 10%-threshold for reporting explained in appendix A.2.1. Therefore, I assume that firm A is supplier to firm B in all years between t_{first} and t_{last}.

Controls include the spatial proximity between inventors. I define the distance (in nautical miles) between two patents by taking the minimum out of all pairwise distances between patents' inventors. I include a dummy whether two patents come from the same country or not. The country for a patent is defined by the country of the majority of inventors as in section A.2.2. The controls also include a dummy whether cited and citing

patents are in the same main subgroup of CPC, and dummies for a lag between a citing (control) application year and cited grant year.

A.2.6 Details to Section 4.2: Import Competition with China

Section 1.4.2 estimates how import competition with China affected citations patterns. For this exercise, I follow the methodology in Autor *et al.* (2020) (Appendix B.3) for the analysis at the technology class level . Specifically, I do the following.

First, I take the set of top 5% of the most cited patents ($Main$ sample defined in appendix A.2.2). For the specification with corporate patents, I leave only patents assigned to corporate firms (both public and private). As in the previous sections, denote by f_k the assignee of patent k and by t_k^a the application year of patent k. I group patents into five 7-year periods based on the grant year. Formally, I define the following sets

$$S_{1977} = \{k \in Main : f_k \in \text{Corporate}, t_k^a \in [1977, 1983]\}$$

$$S_{1984} = \{k \in Main : f_k \in \text{Corporate}, t_k^a \in [1984, 1990]\}$$

$$S_{1991} = \{k \in Main : f_k \in \text{Corporate}, t_k^a \in [1991, 1997]\}$$

$$S_{1998} = \{k \in Main : f_k \in \text{Corporate}, t_k^a \in [1998, 2004]\}$$

$$S_{2005} = \{k \in Main : f_k \in \text{Corporate}, t_k^a \in [2005, 2011]\}$$

where the sample $Main$ is defined in Section A.2.2.

Second, for each set S_t and each technology class I compute the aggregated concentration measure. I take the simple average across concentration measures. Autor *et al.* (2020) provides a mapping from USPC technological classes and SIC industries. Moreover, there exists a matching from USPC to NBER technology categories that will be used as controls. Therefore, for technology classes I use USPC system. Denote the aggregate concentration measure for technology class j and set S_t by $\mathcal{C}_{j,t}$.

Third, given the constructed $\mathcal{C}_{j,t}$ the analysis proceeds as described in Section 1.4.2. Data construction with non-corporate patents is the same except in the first step I leave only non-corporate patents from the $Main$ sample.

A.2.7 Details to Section 5.1: Estimation of the Value of Private Knowledge

Section 1.5.1 estimates a regression of the firm's growth in sales and profits on the composition of its patent portfolio. For firm f and year t, this portfolio includes the total number of granted patents, N_{ft}, and the number of patents making at least one citation to the top 5% of the most cited patents, N^g_{ft}. Specifically,

$$N^g_{ft} = |\{m \in \mathcal{P} : \text{ there is } k \in Main \text{ such that } m \in Citing_k\}|$$

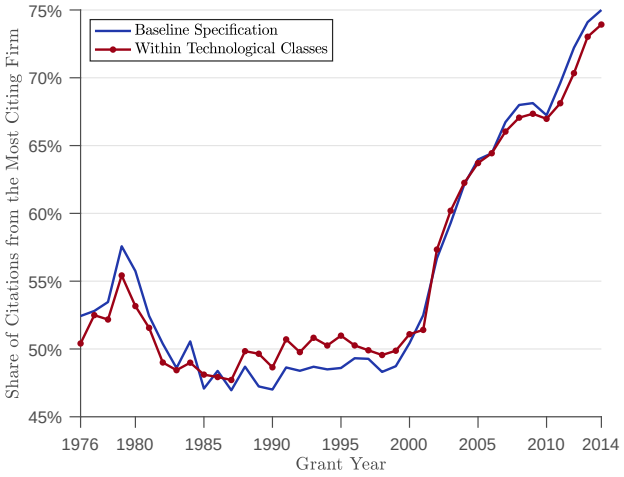
where $\mathcal{P}$ is the set of all granted patents, the sample $Main$ is defined in Section A.2.2, and the sample $Citing_k$ is defined in (A.2.4). Finally, the firm's portfolio of patents includes the number of "specific" patents N^c_{ft} that are responsible for the majority of citations to a highly-concentrated patent in the $Main$ sample. The precise definition is given in section A.2.2.

The details of specification (1.5.1) are the following. The outcome variable $y_{f,t}$ represents sales (Compustat variable "sale") or profits (Compustat variables: "sale" − "cogs"). Both variables are deflated using by CPI (= 100 for 1982 year). Variable X_{ft} includes firm employment (Compustat variable "emp"), firm capital stock (Compustat variable "ppegt" deflated by the NIPA price of equipment), and a lag of profit or sales. All variables are winsorized at the 1% level using annual breakpoints. The variable $\zeta_{i(f),t}$ is an industry-year fixed effect where industry is defined using historical 4-digit SIC classification. For years with missing industry information I take the industry for the nearest year. Standard errors are clustered by firm and year.

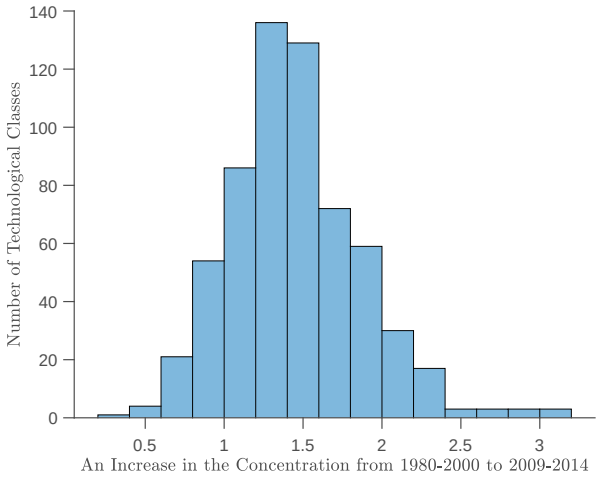
A.2.8 Joint Innovations Between Suppliers and Customers

In this section, I consider the dynamics of joint innovations between suppliers and customers. I focus on patents with multiple assignee where at least of these assignees are from publicly traded firms in Compustat. Around 3% of patents granted in USPTO since 1976 have at least two assignees. Multiple assignees on a patent means that there are several organizations sharing property rights on a patent. To study the joint research between suppliers and customers, I compute two statistics. First, I compute the share of patents granted to firms in a vertical relationship relative to total number of patents granted to two at least two publicly traded firms in the U.S. Panel a of Figure A.4 shows that the share of joint patents between suppliers and customers among all joint patents has significantly increased over time. However, the concern with this rise might be related to changes in the reporting requirements in 10-K files. In other words, the increase in the share of joint patents among suppliers and customers might be driven by greater disclosure of supplier-customer relationships over time rather than greater activity in joint innovations. To address this concern, in Panel b I plot the share of supplier-customer pairs who had joint patents in a particular period relative to the total number of supplier-customer pairs in this period. This panel shows that the share of firms who reported a vertical relationship and had joint patents has increased over time.

A.3 Figures and Tables



(a) Baseline vs. Within Technological Classes



(b) Distribution of the Increase Across Classes

Figure A.1: Concentration of Citations Within Technological Classes

Figure A.1a shows the aggregate concentration of citations that is driven by changes within technological classes. In the baseline specification, I aggregate concentration measures across classes by taking an average weighted by the number of patents in a class. The dotted red line shows the concentration in which the average across classes is unweighted. In Figure A.1b, for each technology class (a group category in CPC) I compute the ratio of the average concentration between 2009 and 2014 to the average concentration between 1976 and 2000. Figure A.1b shows the distribution of the increase in the concentration measure across classes.

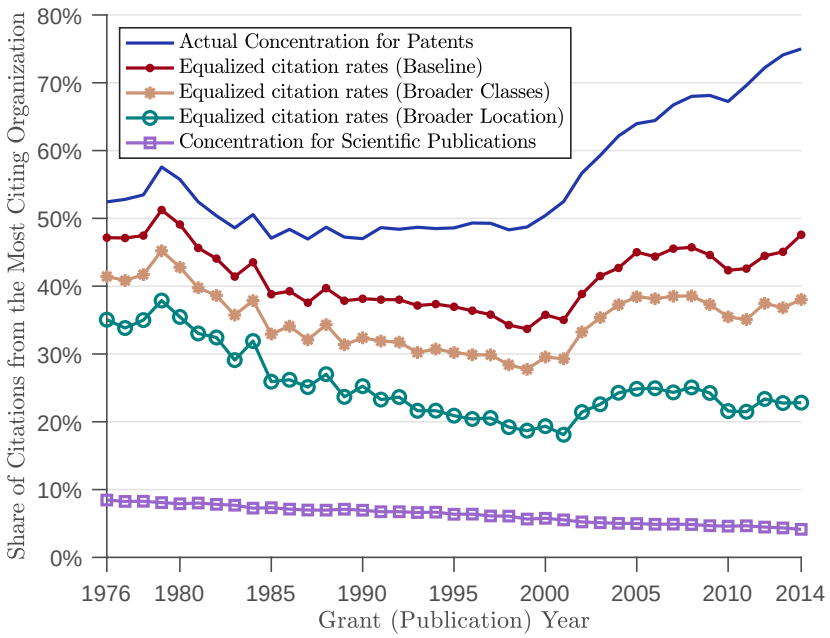


Figure A.2: Concentration of Citations: Patents, Monte-Carlo, and Publications

This figure shows the average concentration of citations for the most cited patents and scientific publications between 1976 and 2014. For each cited patent and publication, the concentration is defined as the share of citations coming from the most citing organization. For patents the most citing organization is usually a corporate firm, and for scientific publications it is usually a university. The most citing patents are defined in the following way. For each patent, I track its citations within a 5 year from a grant date, and for each grant year and the group level in Cooperative Patent Classification, I take top 5% of the most cited patents. The most citing scientific publications are defined in a similar way. For each year and the scientific field in Web of Science classification, I take top 5% of the most cited publications where citations are taken within a 5 year from the publication year. Since the number of publications is much larger relative to the number of patents, to reduce the computational requirements I take a random sample of 10% publications from the most cited ones. The solid blue line shows that the actual concentration of patent citations. The dotted (red) line, the starred (brown) line, and the circled (green) line show the medians of the concentration measure where citation rates are equalized across firms. They are constructed using Monte-Carlo simulations where citations are allocated randomly across all patents observationally similar to the citing patents (see Section 1.2.2 and Appendix A.2.2). The dotted line shows the median counterfactual concentration from the baseline specification in Section 1.2.2, the starred line shows the median concentration when patents with similar characteristics are grouped based a location of the majority of inventors, an application year, and on a broader technological class level (group CPC). The circled line shows the median concentration in which patents with similar characteristics are grouped based only on an application year, and on a narrow technological class level (main subgroup CPC). The squared purple line shows the concentration of citations for scientific publications.

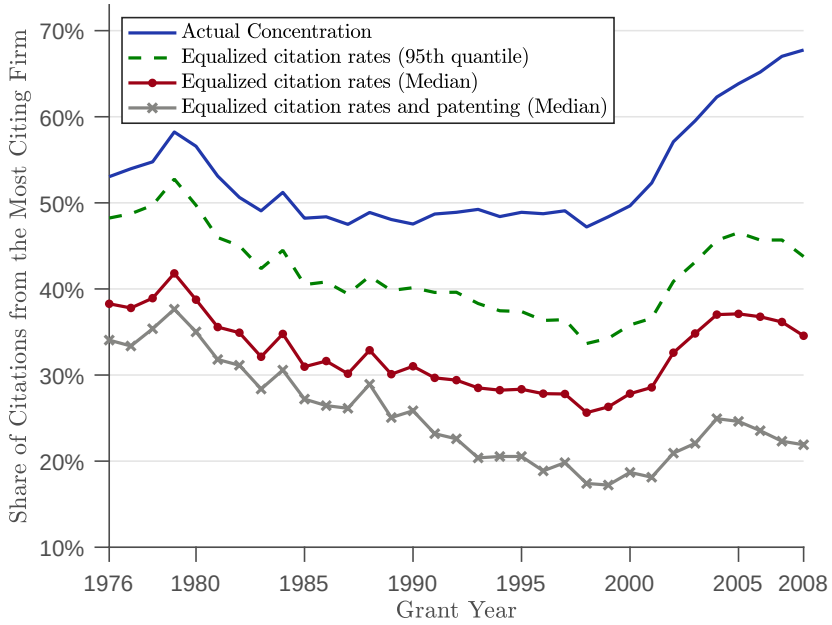
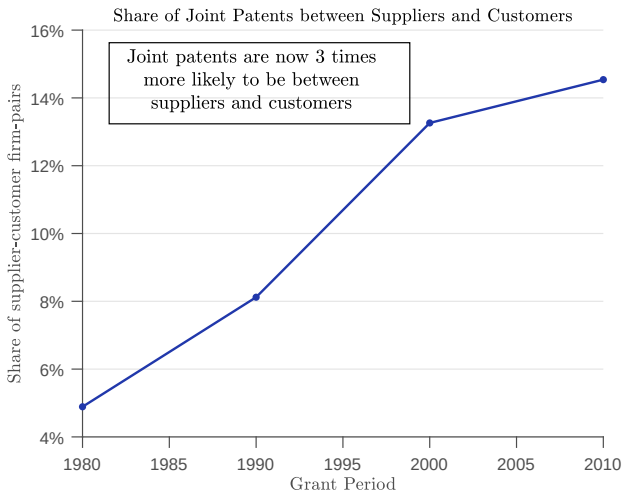
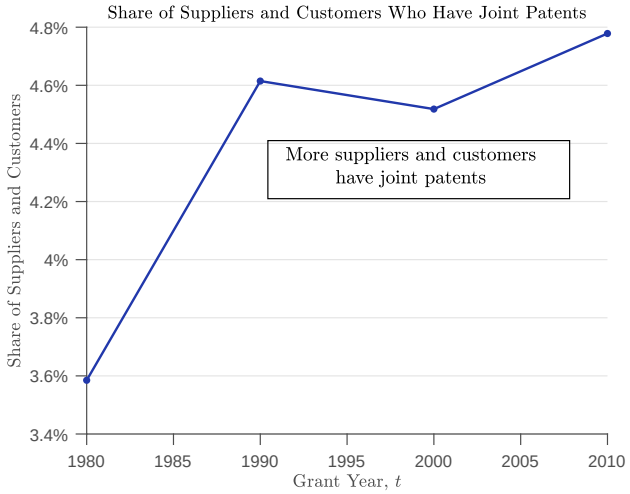


Figure A.3: Decomposition of the Concentration of Citations Based on the Textual Similarity Between Patents

This figure shows the decomposition of the concentration of citations described in Section 1.2.2 and Figure 1.2 where instead of technological classes I group patents based on their textual similarity. Specifically, for each citing patent I find patents with the same application year, and then I choose at most 200 patents closest in terms of the textual similarity based on the data from Arts *et al.* (2018). Finally, among the selected patents I choose the ones where the majority of inventors are from the same location as for the citing patent. Then, I do Monte-Carlo exercises described in Section 1.2.2.



(a) Share of Joint Patents Among Suppliers and Customers.



(b) Share of Suppliers and Customers who have Joint Patents.

Figure A.4: Joint Innovations Between Suppliers and Customers

Panel a shows the share of joint patents between suppliers and customers among all joint patents. Panel b shows the share of supplier-customer pairs who had joint patents in a particular period relative to the total number of supplier-customer pairs in this period.

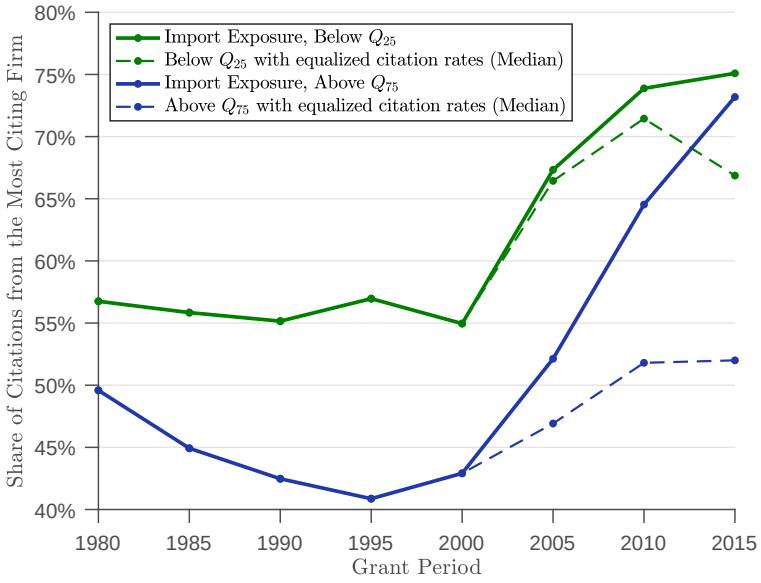


Figure A.5: Trade with China. Actual Concentration vs. Counterfactual One with Equalized Citation Rates

This figure shows the concentration measure from Section 1.2 (Figure 1.1) for different technological classes divided into quartiles based on their exposure to import competition from China, ΔIP_{i2} in (1.4.2). The solid lines show the concentration measures over time for the technologies most (the blue line) and least (the green line) exposed to competition from China. The dashed lines show the counterfactual concentration measures that are evolved due to changes in patenting across firms holding firms' citation rates fixed at the level of 2000. Formally, for patents granted after 2000 I compute the median concentration measures in the in the Monte-Carlo simulations where citations are allocated randomly across patents observationally similar to citing ones (see A.2.2). I add the difference between the actual and the counterfactual concentration measures in 2000.

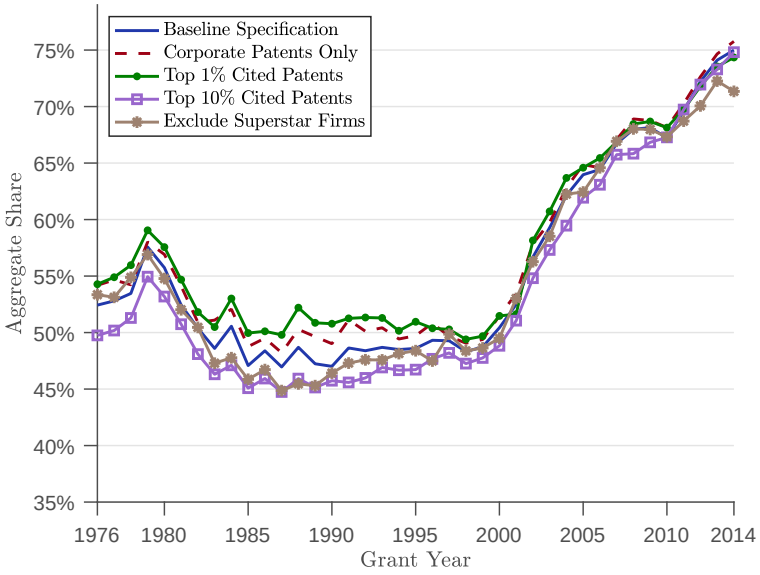


Figure A.6: Corporate patents, top 1%, top 10%, exclude superstar firms in patenting

This figure shows the concentration of citations in the sample of corporate patents only. I also consider different thresholds for the most cited patents: top 10% and top 1%. Finally, I exclude the sample of citing patents superstar firms. Specifically, in each year and group level in Cooperative Patent Classification I find top 1% of firms in terms of the number of patents, and exclude their patents from the sample of citing patents.

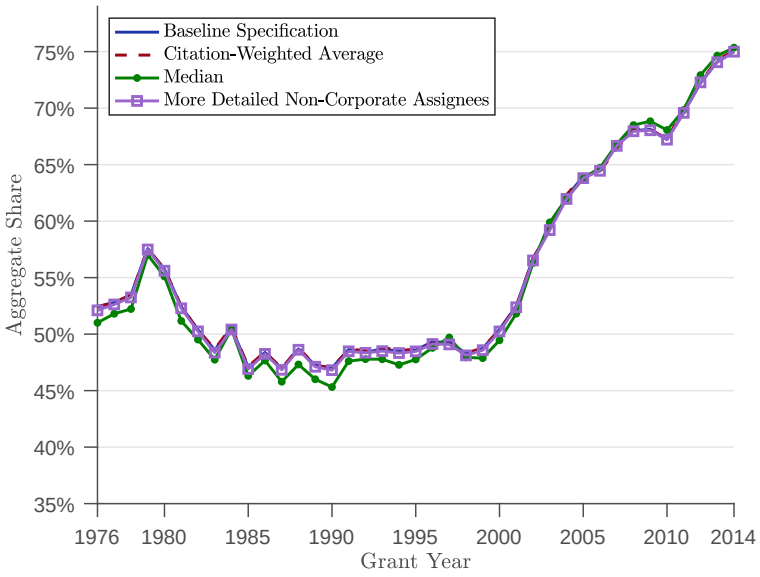


Figure A.7: Alternative aggregation: citation-weighted average and median within classes

This figure shows the concentration of citations if one uses the citation-weighted average or the median instead of the average to aggregate concentration measures within technology classes. I also plot on this graph the concentration measure when the non-corporate sector is disaggregated into different organizations rather than types of organization.

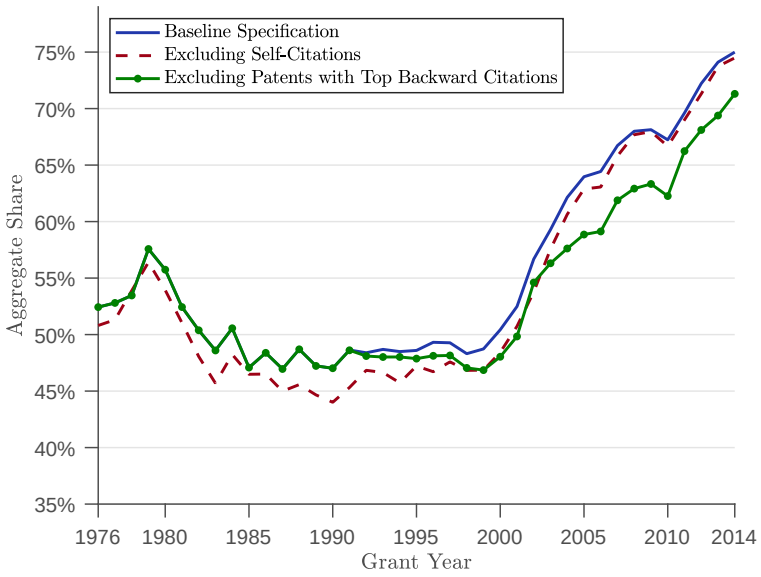


Figure A.8: Exclude Self-Citations and top patents in terms of backward citations

This figure shows the concentration in the sample without self-citations of firms to itself. It also shows the concentration in the sample without top 1% of patents in terms of the number of backward citations.

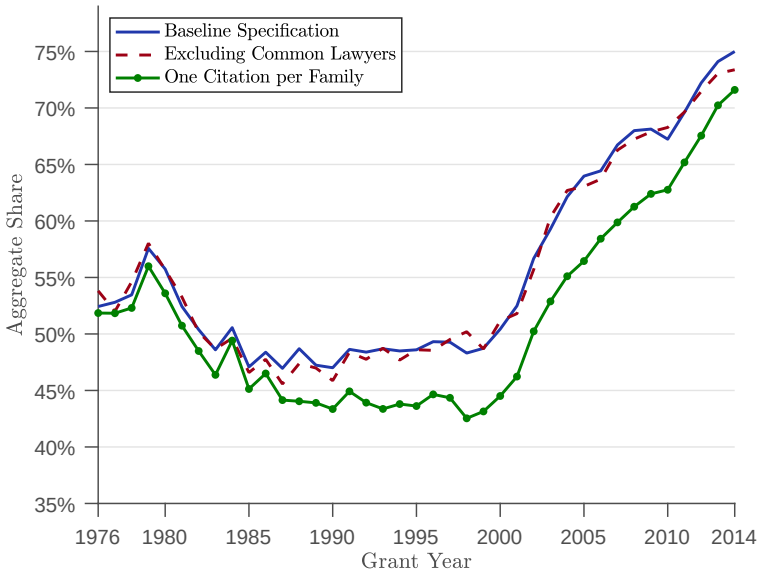


Figure A.9: Exclude Citations from common lawyers and group patents from one family

This figure shows the concentration of citations in the sample without citations between patents sharing a common law firm. I also group citations from patents from the same within-country family (continuations, continuations-in-part, divisionals) as a single citation.

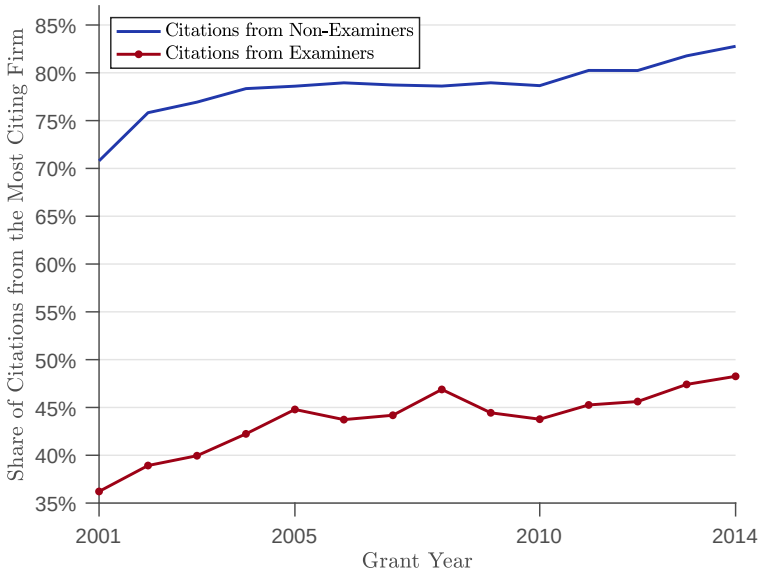


Figure A.10: Concentration of Citations: Examiners vs. Non-examiners

This figure shows the concentration of citations across firms in which I separate citations from examiners and non-examiners. USPTO started to distinguish citations from examiners in 2001.

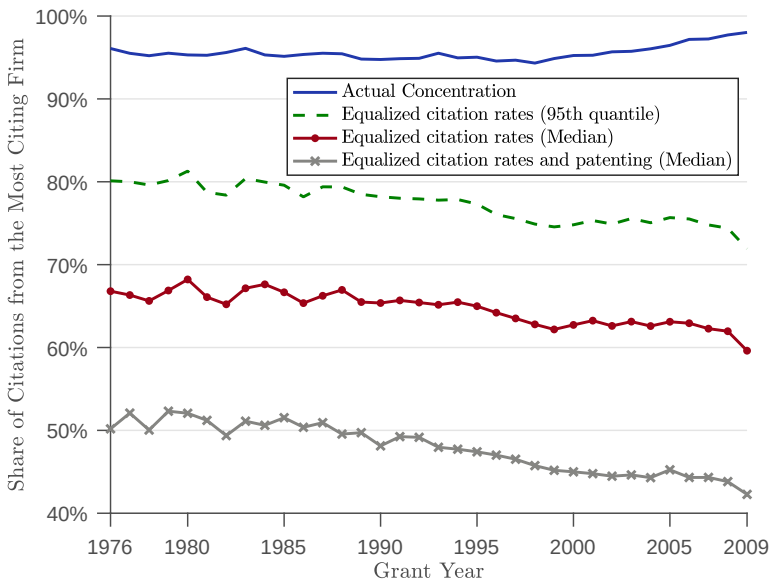


Figure A.11: Concentration of Citations for Lawyers Who Represented Several Firms

This figure shows the concentration of citations across firms within lawyers who represented multiple companies. I compare patents with similar characteristics: the same main subgroup level in Cooperative Patent Classification, the same geographical location (states in the U.S., and countries for foreign inventors), and application time (within 2 years from the citing application year). The solid line shows the actual aggregate within-lawyer concentration of citations across firms measured by the share of citations coming from one firm only. The dashed line shows the 95th quantile of the same measure in Monte-Carlo simulations where citations are allocated randomly within an inventor.

Table A.1: Distribution of Citations across firms for IBM/Novellus's patent 6126798.

Citing Firm	% of Citations	Relationship to IBM	Source
Novellus Systems	64%	Supplier	Compustat Segments, Lim (2009)
Lam Research	16%	Supplier	IBM's website
Applied Materials	11%	Supplier	Lim (2009)
Semitool Inc	3%	Supplier	Compustat Segments, Lim (2009)
Toshiba, Ebara	2%	Share a Supplier	Compustat Segments
Entegris Inc	0.05%	Joint Collaboration	Entegris's website
IBM	0.05%	Self-Citations	—
Other	3%	No Info	—

Table A.2: Distribution of Citations across firms for Motorola's patent 5391517.

Citing Firm	% of Citations	Relationship to Motorola	Source
Semitool Inc	13%	Supplier	Compustat Segments
ASM International	11%	Semitool's Industry	Compustat
IBM	9%	Share a Supplier	Compustat Segments
Applied Materials	7%	Supplier	Compustat Segments
U.S. Navy	7%	No Info	—
Taiwan Semiconductor Manufacturing	5%	Share a Supplier	Compustat Segments
Advanced Micro Devices	4%	Share a Supplier	Compustat Segments
Intel Corp	4%	Share a Supplier	Compustat Segments
Motorola	4%	Self-Citations	—
Other Suppliers	3%	Suppliers	Compustat Segments
Other Firms Sharing a Supplier	17%	Share a Supplier	Compustat Segments
Other Firms	7%	No Info	—

Table A.3: Placebo Tests: Trade with China and Increase in the Concentration of Citations

	Non-Corporate Patents					Lag Outcomes			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ Tech Class Exposure	0.83	0.79	0.74	1.67	1.68	−0.10	−0.10	−0.71	−0.82
to Chinese Imports	(0.71)	(0.70)	(0.64)	(1.28)	(1.30)	(0.71)	(0.72)	(1.25)	(1.25)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Δ Citations		Yes	Yes	Yes	Yes		Yes	Yes	Yes
2 Lags of outcomes			Yes	Yes	Yes				
11 sectors, 6 Tech				Yes	Yes			Yes	Yes
Software Patents					Yes				Yes

Notes: This table shows the results for the falsification tests in specification (1.4.3) and Table 1.4. Changes in US import exposure are instrumented with Chinese exports to non-U.S. high-income markets (Autor *et al.* (2020)). In columns (1)–(4), I regress the change in the concentration of citations for non-corporate patents on the changes in import competition from China. In columns (5)–(7), I regress the change in the concentration measure pre-period (pre 1991) on future changes in import exposure. Standard errors are clustered at the technology class level. All specifications are weighted by the number of Compustat-matched U.S.-inventor patents in a technology class. All coefficients are insignificant at the 10% level.

A.4 Theory Appendix

A.4.1 Proposition 1: $\mathcal{I}$'s profits under different contracts

If $\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing with a confidentiality agreement, then under knowledge sharing $\mathcal{I}$'s profits in the final stage will be

$$\begin{aligned}\pi_{\mathcal{I}}(ex=1,ca=1) &= \frac{1-\alpha}{2}(1-\gamma)\lambda + \alpha(1-\gamma)\left(\lambda-\lambda_0\right) + \\ &+ \frac{1}{2}\Big(\frac{1-\alpha}{2}\lambda + \alpha\left(\lambda-(1-\gamma)\lambda_0\right) - \frac{1-\alpha}{2}(1-\gamma)\lambda - \alpha(1-\gamma)\left(\lambda-\lambda_0\right)\Big) \\ &= \frac{1-\alpha}{2}(1-\gamma)\lambda + \alpha(1-\gamma)\left(\lambda-\lambda_0\right) + \frac{1+\alpha}{4}\gamma\lambda\end{aligned} \tag{A.4.1}$$

If $\mathcal{I}$ and $\mathcal{S}$ sign a confidentiality agreement only, then under knowledge sharing $\mathcal{I}$'s

profits in the final stage will be

$$\begin{aligned}\pi_{\mathcal{I}}(ex=0,ca=1) &= \frac{1-\alpha}{2}(1-\gamma)\lambda + \alpha\cdot\max\left\{0,(1-\gamma)\lambda-\lambda_0\right\}+\\ &\quad+\frac{1}{2}\cdot\max\left\{0,\left(\frac{1-\alpha}{2}\gamma\lambda+\alpha(\lambda-\lambda_0)-\alpha\cdot|(1-\gamma)\lambda-\lambda_0|\right)\right\}\\ &= \begin{cases}\frac{1-\alpha}{2}(1-\gamma)\lambda+\alpha(\lambda-\lambda_0)+\frac{1-3\alpha}{4}\gamma\lambda & \text{if } \alpha\cdot\left(3\gamma-4(1-\frac{\lambda_0}{\lambda})\right)\leqslant\gamma\\ \frac{1-\alpha}{2}(1-\gamma)\lambda & \text{if } \alpha\cdot\left(3\gamma-4(1-\frac{\lambda_0}{\lambda})\right)>\gamma\end{cases}\end{aligned} \tag{A.4.2}$$

If $\mathcal{I}$ and $\mathcal{S}$ do not sign any contract in the first stage, then under knowledge sharing $\mathcal{I}$'s profits in the final stage will be

$$\begin{aligned}\pi_{\mathcal{I}}(ex=0,ca=0) &= \frac{1-\alpha}{2}(1-\gamma)\lambda+\\ &\quad+\frac{1}{2}\cdot\max\left\{0,\left(\frac{1-\alpha}{2}\lambda+\alpha\cdot 0-\frac{1-\alpha}{2}(1-\gamma)\lambda-\alpha\left(\lambda-(1-\gamma)\lambda\right)\right)\right\}\\ &= \frac{1-\alpha}{2}(1-\gamma)\lambda+\max\left\{0,\frac{1-3\alpha}{4}\gamma\lambda\right\}\end{aligned} \tag{A.4.3}$$

In all cases, the first line is $\mathcal{I}$'s outside option, the second line is a half of the gain in the total surplus of $\mathcal{I}$ and $\mathcal{S}$, and the third line is a simplified expression.

If $\mathcal{I}$ does not share its knowledge with $\mathcal{S}$, then its profits in the final stage will be

$$\pi_{\mathcal{I}}\left(\text{No Sharing}\right)=\frac{1-\alpha}{2}(1-\gamma)\lambda+\alpha(1-\gamma)\left(\lambda-\lambda_0\right) \tag{A.4.4}$$

For $\alpha>\frac{1}{3}$, $\mathcal{I}$ and $\mathcal{E}$ prefer to sign exclusive dealing, and $\mathcal{I}$'s profits under knowledge sharing with exclusive dealing are higher under knowledge sharing with $\mathcal{E}$ than without sharing. Under $\alpha\leqslant\frac{1}{3}$, $\mathcal{I}$'s profits under a confidentiality agreement exceed its profits without knowledge sharing. For the case without a contract (Null), $\mathcal{I}$ shares knowledge with $\mathcal{S}$ under condition (1.3.6).

A.4.2 Two Suppliers and a Common Customer

Proposition 10. *The equilibrium contract in the model from Section 1.3.2 is the following*

$$Contract = \begin{cases} \textit{Exclusive with CA} & \textit{if } \alpha > \frac{\gamma}{2(\lambda_1-\lambda_0)+\gamma(2\lambda_0-1)} \\ \textit{CA only} & \textit{if } \frac{\gamma}{2(1-\lambda_0)+\gamma(2\lambda_0-1)} < \alpha \leqslant \frac{\gamma}{2(\lambda_1-\lambda_0)+\gamma(2\lambda_0-1)} \\ \textit{No Contract} & \textit{if } \alpha \leqslant \frac{\gamma}{2(1-\lambda_0)+\gamma(2\lambda_0-1)} \end{cases} \tag{A.4.5}$$

Firm $\mathcal{I}$ always shares knowledge with $\mathcal{B}$, and $\mathcal{B}$ shares $\mathcal{I}$'s secrets with $\mathcal{E}$ only when $\mathcal{I}$ and $\mathcal{B}$ do not sign any contract in the first stage.

Proof. Joint surplus of $\mathcal{I}$ and $\mathcal{B}$ under different contracts is the following

$$\text{Exclusive, CA:} \qquad TS(ex=1,ca=1) = \frac{1-\alpha}{2} + \alpha \tag{A.4.6}$$

$$\text{CA Only:} \qquad TS(ex=0,ca=1) = \frac{1-\alpha}{2} + \alpha + \frac{1-\alpha}{2}(1-\gamma)\lambda_1 \tag{A.4.7}$$

$$\text{No Contract:} \qquad TS(ex=0,ca=0) = \frac{1-\alpha}{2} + \alpha + \frac{1-\alpha}{2} = 1 \tag{A.4.8}$$

$\mathcal{I}$ and $\mathcal{B}$ sign a contract that maximizes their joint surplus subject to the constraint that $\mathcal{I}$ has incentives to share its secrets with $\mathcal{B}$. If $\mathcal{I}$ shares its secrets with $\mathcal{B}$, then its profit under different contracts will be the following

$$\begin{aligned} \pi_{\mathcal{I}}(ex=1,ca=1) &= \frac{1}{2}\cdot\left(\frac{1-\alpha}{2}+\alpha\right) = \frac{1+\alpha}{4} \\ \pi_{\mathcal{I}}(ex=0,ca=1) &= \frac{1}{2}\cdot\left(\frac{1-\alpha}{2}+\alpha-\alpha\lambda_1\right) = \frac{1+\alpha}{4} - \frac{\alpha}{2}\lambda_1 \\ \pi_{\mathcal{I}}(ex=0,ca=0) &= \frac{1}{2}\cdot\left(\frac{1-\alpha}{2}+\alpha-\alpha\right) = \frac{1+\alpha}{4} - \frac{\alpha}{2} \end{aligned}$$

If $\mathcal{I}$ does not share its secrets with $\mathcal{B}$, then its profit will be

$$\pi_{\mathcal{I}}\left(\text{No Sharing}\right) = \frac{1}{2}\left(\frac{1-\alpha}{2}+\alpha\cdot(1-\lambda_0)\right)\cdot(1-\gamma) = \frac{1+\alpha}{4}(1-\gamma) - \frac{\alpha}{2}\lambda_0(1-\gamma)$$

Under exclusive dealing with a confidentiality agreement, $\mathcal{I}$ always has incentives to share

its secrets with $\mathcal{B}$. Under a confidentiality agreement only, it shares its secrets if

$$\frac{1+\alpha}{4}\gamma \geqslant \frac{\alpha}{2}\left(\lambda_1 - (1-\gamma)\lambda_0\right) \Leftrightarrow \alpha \leqslant \frac{\gamma}{2(\lambda_1-\lambda_0)+\gamma(2\lambda_0-1)} \tag{A.4.9}$$

Without any contract in the first stage, $\mathcal{I}$ shares its secrets with $\mathcal{B}$ if

$$\frac{1+\alpha}{4}\gamma \geqslant \frac{\alpha}{2}\left(1 - (1-\gamma)\lambda_0\right) \Leftrightarrow \alpha \leqslant \frac{\gamma}{2(1-\lambda_0)+\gamma(2\lambda_0-1)} \tag{A.4.10}$$

The joint surplus of $\mathcal{I}$ and $\mathcal{B}$ is always higher with less contractual restrictions. Therefore, they sign a contract so that $\mathcal{I}$ has incentives to share its secrets with $\mathcal{B}$. The combination of (A.4.9) and (A.4.10) gives the contract in (A.4.5). □

A.4.3 Welfare Implications of Restrictions on the Vertical Contracts

Proposition 11. *Consider the model with a common supplier who can sell inputs to two downstream firms from Section 1.3.1. With restrictions on the use of confidentiality agreements, the equilibrium contract between $\mathcal{I}$ and $\mathcal{S}$ is*

$$\begin{cases} \text{Exclusive Dealing} & \text{if } \alpha > \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)} \\ \text{Null} & \text{if } \alpha \leqslant \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)} \end{cases} \tag{A.4.11}$$

Under exclusive dealing only $\mathcal{S}$ learns $\mathcal{I}$'s secrets, and without any contract between $\mathcal{I}$ and $\mathcal{S}$ both $\mathcal{S}$ and $\mathcal{E}$ learn $\mathcal{I}$'s secrets.

With restrictions on the use of exclusive dealing, if $\alpha \cdot \left(4\frac{\lambda_0}{\lambda}-1\right) > 1$, then $\mathcal{I}$ will not share its secrets with $\mathcal{S}$ regardless of whether they have a confidentiality agreement or not. If $\alpha \cdot \left(4\frac{\lambda_0}{\lambda}-1\right) \leqslant 1$, the equilibrium contract is

$$\begin{cases} \text{Confidentiality Agreement} & \text{if } \alpha > \min\left\{\frac{\gamma}{\gamma+2}, \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)}\right\} \\ \text{Null} & \text{if } \alpha \leqslant \min\left\{\frac{\gamma}{\gamma+2}, \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)}\right\} \end{cases} \tag{A.4.12}$$

Under a confidentiality agreement only $\mathcal{S}$ learns $\mathcal{I}$'s secrets, and without any contract between $\mathcal{I}$ and $\mathcal{S}$ both $\mathcal{S}$ and $\mathcal{E}$ learn $\mathcal{I}$'s secrets.

With restrictions on the use of confidentiality agreements and exclusive dealing, both $\mathcal{S}$ and $\mathcal{E}$ learn $\mathcal{I}$'s secrets if

$$\alpha \leqslant \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)} \tag{A.4.13}$$

and nobody learns $\mathcal{I}$'s secrets otherwise.

The socially optimal regulation of confidentiality agreements and exclusive dealing is

$$\begin{cases} \text{Ban on CA} & \text{if } \alpha \leqslant \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)} \\ \text{Ban on Exclusive Dealing} & \text{if } \alpha \leqslant \frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)} \text{ and } \alpha \cdot \left(4\frac{\lambda_0}{\lambda}-1\right) \leqslant 1 \\ \text{No Restrictions} & \text{if } \alpha \cdot \left(4\frac{\lambda_0}{\lambda}-1\right) > 1 \end{cases} \tag{A.4.14}$$

Proof. With restrictions on confidentiality agreements, $\mathcal{I}$ and $\mathcal{S}$ have two options: sign exclusive dealing or do not sign any contracts. Under exclusive dealing, $\mathcal{S}$ cannot gain from knowledge sharing with $\mathcal{E}$ because it does not sell an input to it. I assume that in this case it does not share $\mathcal{I}$'s secrets with $\mathcal{E}$.

Joint surplus of $\mathcal{I}$ and $\mathcal{S}$ is higher without a contract relative to exclusive dealing if

$$\frac{1-\alpha}{2}\lambda + \frac{1-\alpha}{2}\gamma\lambda \geqslant \frac{1-\alpha}{2}\lambda + \alpha\left(\lambda-(1-\gamma)\lambda_0\right) \Leftrightarrow \alpha \leqslant \frac{\gamma}{2\left(1-\frac{\lambda_0}{\lambda}\right)+\gamma\left(2\frac{\lambda_0}{\lambda}+1\right)}$$

Without a contract, $\mathcal{I}$ has incentives to share its secrets with $\mathcal{S}$ if condition (A.4.13) is satisfied. The derivation is the same as in Proposition 1. Since

$$\frac{\gamma}{4(1-\frac{\lambda_0}{\lambda})+\gamma(4\frac{\lambda_0}{\lambda}-1)} \leqslant \frac{\gamma}{2\left(1-\frac{\lambda_0}{\lambda}\right)+\gamma\left(2\frac{\lambda_0}{\lambda}+1\right)} \text{ for any } \gamma \in [0,1]$$

$\mathcal{I}$ and $\mathcal{S}$ sign exclusive dealing only when (A.4.13) is not satisfied. This leads to the equilibrium contract given in (A.4.11).

With restrictions on exclusive dealing, $\mathcal{I}$ and $\mathcal{S}$ have two options: sign a confidentiality

agreement or do not sign any contracts. For $\alpha \leqslant \frac{1}{3}$, the equilibrium contract is the same as in Proposition 1 because these two options dominate exclusive dealing. For $\alpha > \frac{1}{3}$, $\mathcal{I}$ does not share its secrets with $\mathcal{S}$ without any contract. Under a confidentiality agreement it shares them only it its profit under sharing (A.4.2) is greater than the profit without knowledge sharing (A.4.4). This condition can be simplified to

$$\alpha \cdot \left(4\frac{\lambda_0}{\lambda} - 1 \right) \leqslant 1 \tag{A.4.15}$$

If (A.4.15) is satisfied, then the equilibrium contract is described in (A.4.12). Otherwise, $\mathcal{I}$ does not share its secrets with $\mathcal{S}$.

Finally, with restrictions on both exclusive dealing and confidentiality agreements, $\mathcal{I}$ compares its profit under knowledge sharing (A.4.3) relative to the profit without sharing (A.4.4). This condition is given in (A.4.13).

The analysis above leads to the optimal regulation in (A.4.14).

□

Appendix B

Appendix to Chapter 2

Table B.1: Keywords for the Textual Analysis in Chapter 2

Type	Keywords	Type	Keywords
Technological Information	technology	Business Information	customer
Technological Information	equipment	Business Information	database
Technological Information	patenting	Business Information	client
Technological Information	patent	Business Information	pricing
Technological Information	invention	Business Information	financial
Technological Information	prototype	Business Information	insurance
Technological Information	prototyping	Business Information	finance
Technological Information	machine	Business Information	salesperson
Technological Information	hardware		
Software Technology	software	Bitech/Pharma	chemical
Software Technology	computer program	Bitech/Pharma	clinical
Software Technology	computing	Bitech/Pharma	pharmaceutical
Software Technology	algorithm	Bitech/Pharma	fiber
Software Technology	programming	Bitech/Pharma	biotechnology
Software Technology	coding	Bitech/Pharma	biotech
Software Technology	computer code	Bitech/Pharma	bioscience
Software Technology	artificial intelligence	Bitech/Pharma	biology
Software Technology	cloud computing	Bitech/Pharma	antibody
Software Technology	machine learning		
Software Technology	virtual reality		
Software Technology	natural language processing		
Software Technology	deep learning		
Breach of Contract	solicit		
Breach of Contract	solicitation		
Breach of Contract	non disclosure		
Breach of Contract	NDA		
Breach of Contract	non solicitation		
Breach of Contract	non compete		
Breach of Contract	poach		
Breach of Contract	poaching		
Breach of Contract	non poaching		

Table B.2: Trade Secret Litigation Across US States

State	Share of Cases	Employment Keywords	Employment + Contract Keywords	Technology Keywords	Business Keywords	Software Keywords	Biotech Keywords
California	13%	84%	43%	51%	76%	67%	21%
Texas	9%	77%	42%	41%	70%	60%	21%
Florida	7%	85%	54%	33%	76%	59%	16%
New York	7%	79%	51%	38%	77%	65%	19%
Pennsylvania	6%	88%	58%	42%	82%	63%	33%
Illinois	5%	82%	58%	42%	82%	71%	19%
Ohio	5%	82%	56%	34%	78%	52%	17%
New Jersey	4%	87%	57%	43%	81%	68%	31%
Georgia	3%	81%	48%	30%	72%	65%	18%
Virginia	3%	79%	52%	39%	77%	58%	15%
Michigan	3%	81%	54%	39%	75%	58%	16%
Colorado	2%	79%	53%	39%	79%	62%	13%
Massachusetts	2%	78%	53%	48%	72%	63%	26%
Minnesota	2%	82%	59%	42%	79%	66%	20%
Washington	2%	73%	41%	48%	72%	68%	15%

Table B.3: Trade Secret Litigation Across Industries

Industry (SIC3)	Share of Cases	Employment Keywords	Employment + Contract Keywords	Technology Keywords	Business Keywords	Software Keywords	Biotech Keywords
Computer Programming Data Processing (737)	12%	63%	41%	57%	74%	91%	11%
Surgical and Medical Instruments (384)	5%	71%	47%	65%	71%	55%	50%
Drugs (283)	5%	67%	36%	56%	58%	53%	88%
Computer And Office Equipment (357)	4%	58%	39%	71%	67%	72%	8%
Electronic Components And Accessories (367)	3%	52%	33%	73%	51%	64%	16%
Fire, Marine, And Casualty Insurance (633)	3%	88%	79%	9%	97%	89%	11%
Communications Equipment (366)	2%	72%	40%	74%	62%	74%	15%
Motor Vehicles And Motor Vehicle Equipment (371)	2%	68%	24%	56%	46%	48%	9%
Services Allied With The Exchange Of Securities (628)	2%	97%	92%	7%	100%	100%	0%
Wholesale: Professional And Commercial Equipment (504)	2%	87%	75%	71%	94%	65%	3%

Appendix C

Appendix to Chapter 3

C.1 Theory Appendix

Proof of Proposition 1

Total sales from plants based in k in country i can be written as

$$S_{hki}(\varphi) = \widetilde{\kappa}\varphi^{\sigma-1} \cdot \mu_{hki} \cdot (\Psi_{hi})^{\frac{\sigma-1}{\theta^a}} \cdot E_i P_i^{\sigma-1}$$

where μ_{hki} and Ψ_{hi} are given in equations (3.4.12) and (3.4.13). Both μ_{hki} and Ψ_{hi} are increasing in $T_k^a (\tau_{ki}^a)^{-\theta^a} (w_k)^{-(1-\alpha)\theta^a}$ and decreasing in τ_{jk}^s. This proves part (i) of the proposition.

For part (ii), we rewrite $S_{hk'i}(\varphi)$ as

$$\widetilde{\kappa}\varphi^{\sigma-1} \cdot T_{k'}^a (\tau_{k'i}^a)^{-\theta^a} (w_{k'})^{-(1-\alpha)\theta^a} (\Theta_{hk'}(\varphi))^{\alpha\theta^a/\theta^s} \cdot (\Psi_{hi})^{\frac{\sigma-1}{\theta^a}-1} \cdot E_i P_i^{\sigma-1}$$

where the term $(\Psi_{hi})^{\frac{\sigma-1}{\theta^a}-1}$ is increasing (decreasing) in $T_k^a (\tau_{ki}^a)^{-\theta^a} (w_k)^{-(1-\alpha)\theta^a}$ and decreasing (increasing) in τ_{jk}^s if $\sigma - 1 > \theta^a$ ($\sigma - 1 < \theta^a$).

Proof of Proposition 2

Input purchases from source country j of the plant based in k, $M_{hkj}(\varphi)$, can be written as

$$M_{hkj}(\varphi) = \left(1 - \frac{1}{\sigma_w}\right) \alpha \cdot \chi_{hjk} \sum_i S_{hki}(\varphi) \tag{C.1.1}$$

where term χ_{hjk} is given in (3.4.9), and it does not depend on $T_k^a \left(\tau_{ki}^a\right)^{-\theta^a} \left(w_k\right)^{-(1-\alpha)\theta^a}$. Proposition 2 follows from Proposition 1.

Proof of Proposition 3

Part (i) follows from (C.1.1) and Proposition 1: both χ_{hjk} and $\sum_i S_{hki}(\varphi)$ are decreasing in τ_{jk}^s. For part (ii), input purchases from source country $j' \neq j$ of the plant based in k, $M_{hkj'}(\varphi)$, can be written as

$$M_{hkj'}(\varphi) = \tilde{\tilde{\kappa}} \cdot \left(\Theta_{hk}\left(\varphi\right)\right)^{\frac{\alpha\theta^a}{\theta^s}-1} \sum_{i \in J} \left(\tau_{ki}^a\right)^{-\theta^a} \left(\Psi_{hi}\left(\varphi\right)\right)^{\frac{(\sigma-1)}{\theta^a}-1} E_i P_i^{\sigma-1}$$

where $\tilde{\tilde{\kappa}}$ includes variables which do not depend on τ_{jk}^s. Both $\left(\Theta_{hk}\left(\varphi\right)\right)^{\frac{\alpha\theta^a}{\theta^s}-1}$ and $\left(\Psi_{hi}\left(\varphi\right)\right)^{\frac{(\sigma-1)}{\theta^a}-1}$ are decreasing (increasing) in τ_{jk}^s if $\sigma - 1 \geqslant \theta^a > \theta^s/\alpha$ $(\sigma - 1 < \theta^a < \theta^s/\alpha)$.

Proof of Lemma 1

Denote by

$$\lambda_{ki} = \xi_{ki}^a \left(\sum_j \mathcal{I}_j^s \cdot \xi_{jk}^s\right)^{\frac{\alpha\theta^a}{\theta^s}}$$

For part (i), it is sufficient to prove that for all $i \in J$

$$\left(\sum_{m \neq k,k'} \mathcal{I}_m^a \lambda_{mi} + \mathcal{I}_k^a \lambda_{ki} + \lambda_{k'i}\right)^{\frac{\sigma-1}{\theta^a}} - \left(\sum_{m \neq k,k'} \mathcal{I}_m^a \lambda_{mi} + \mathcal{I}_k^a \lambda_{ki} + 0\right)^{\frac{\sigma-1}{\theta^a}} \tag{C.1.2}$$

is increasing (decreasing) in $\mathcal{I}_k^a$ if $\sigma - 1 > \theta^a$ ($\sigma - 1 < \theta^a$). Notice that

$$\sum_{m \neq k,k'} \mathcal{I}_m^a \lambda_{mi} + \mathcal{I}_k^a \lambda_{ki} + \lambda_{k'i} > \sum_{m \neq k,k'} \mathcal{I}_m^a \lambda_{mi} + \mathcal{I}_k^a \lambda_{ki} + 0$$

so the derivative of (C.1.2) with respect to $\mathcal{I}_k^a$ is positive if $\sigma - 1 > \theta^a$ and negative if $\sigma - 1 < \theta^a$.

For part (ii), it is sufficient to prove that for all $i \in J$

$$\begin{aligned} &\left(\sum_{k \in J} \mathcal{I}_k^a \cdot \xi_{ki}^a \left(\sum_{m \neq j,j'} \xi_{mk}^s + \mathcal{I}_j^s \xi_{jk}^s + \xi_{j'k}^s \right)^{\frac{\alpha\theta^a}{\theta^s}} \right)^{\frac{\sigma-1}{\theta^a}} - \\ &- \left(\sum_{k \in J} \mathcal{I}_k^a \cdot \xi_{ki}^a \left(\sum_{m \neq j,j'} \xi_{mk}^s + \mathcal{I}_j^s \xi_{jk}^s + 0 \right)^{\frac{\alpha\theta^a}{\theta^s}} \right)^{\frac{\sigma-1}{\theta^a}} \end{aligned} \tag{C.1.3}$$

is increasing in $\mathcal{I}_j^s$ if $\sigma - 1 \geqslant \theta^a$ and $\alpha\theta^a/\theta^s > 1$ and decreasing in $\mathcal{I}_j^s$ if $\sigma - 1 < \theta^a$ and $\alpha\theta^a/\theta^s < 1$. Notice that

$$\sum_{k \in J} \mathcal{I}_k^a \cdot \xi_{ki}^a \left(\sum_{m \neq j,j'} \xi_{mk}^s + \mathcal{I}_j^s \xi_{jk}^s + \xi_{j'k}^s \right)^{\frac{\alpha\theta^a}{\theta^s}} > \sum_{k \in J} \mathcal{I}_k^a \cdot \xi_{ki}^a \left(\sum_{m \neq j,j'} \xi_{mk}^s + \mathcal{I}_j^s \xi_{jk}^s + 0 \right)^{\frac{\alpha\theta^a}{\theta^s}}$$

$$\sum_{m \neq j,j'} \xi_{mk}^s + \mathcal{I}_j^s \xi_{jk}^s + \xi_{j'k}^s > \sum_{m \neq j,j'} \xi_{mk}^s + \mathcal{I}_j^s \xi_{jk}^s + 0$$

so the derivative of (C.1.3) with respect to $\mathcal{I}_j^s$ is positive if $\sigma - 1 \geqslant \theta^a$ and $\alpha\theta^a/\theta^s > 1$ and negative if $\sigma - 1 < \theta^a$ and $\alpha\theta^a/\theta^s < 1$.

For part (iii), it is sufficient to prove that for all $i \in J$

$$\left(\sum_{k \in J} \mathcal{I}_k^a \xi_{jk}^a \left(\sum_{j' \neq j} \mathcal{I}_{j'}^s \xi_{j'k}^s + \xi_{jk}^s \right)^{\frac{\alpha\theta^a}{\theta^s}} \right)^{\frac{\sigma-1}{\theta^a}} - \left(\sum_{k \in J} \mathcal{I}_k^a \xi_{jk}^a \left(\sum_{j' \neq j} \mathcal{I}_{j'}^s \xi_{j'k}^s + 0 \right)^{\frac{\alpha\theta^a}{\theta^s}} \right)^{\frac{\sigma-1}{\theta^a}} \tag{C.1.4}$$

is increasing in $\mathcal{I}_k^a$ if $\sigma - 1 \geqslant \theta^a$. Notice that

$$\sum_{k\in J} \mathcal{I}_k^a \xi_{jk}^a \left(\sum_{j'\neq j} \mathcal{I}_{j'}^s \xi_{j'k}^s + \xi_{jk}^s \right)^{\frac{\alpha\theta^a}{\theta^s}} > \sum_{k\in J} \mathcal{I}_k^a \xi_{jk}^a \left(\sum_{j'\neq j} \mathcal{I}_{j'}^s \xi_{j'k}^s + 0 \right)^{\frac{\alpha\theta^a}{\theta^s}}$$

$$\sum_{j'\neq j} \mathcal{I}_{j'}^s \xi_{j'k}^s + \xi_{jk}^s > \sum_{j'\neq j} \mathcal{I}_{j'}^s \xi_{j'k}^s + 0$$

so the derivative of (C.1.4) is positive if $\sigma - 1 \geqslant \theta^a$.

Proof of Proposition 4

We introduce the following notation

$$\begin{aligned}
\Lambda(\mathcal{I}) &= \kappa \cdot \sum_{i\in J} E_i P_i^{\sigma-1} \left(\xi_{hi}^a \left(\xi_{hh}^s + \mathcal{I}^g \cdot \sum_{j\neq h} \mathcal{I}_j^s \cdot \xi_{jk}^s \right)^{\frac{\alpha\theta^a}{\theta^s}} + \right. \\
&\left. + \mathcal{I}^g \cdot \sum_{k\neq h} \mathcal{I}_k^a \cdot \xi_{ki}^a \left(\xi_{hk}^s + \sum_{j\neq h} \mathcal{I}_j^s \cdot \xi_{jk}^s \right)^{\frac{\alpha\theta^a}{\theta^s}} \right)^{\frac{(\sigma-1)}{\theta^a}} \equiv \\
&\equiv \kappa \cdot \sum_{i\in J} \left(\Psi_{hi}(\varphi) \right)^{(\sigma-1)/\theta^a} E_i P_i^{\sigma-1} \\
F(\mathcal{I}) &= \sum_{j\in J} \mathcal{I}_j^s \cdot w_h f_{hj}^s + \sum_{k\in J} \mathcal{I}_k^a \cdot w_h f_{hk}^a + \mathcal{I}^g \cdot w_h f_h^g
\end{aligned}$$

where $\mathcal{I} = (\mathcal{I}^a, \mathcal{I}^s, \mathcal{I}^g)$.

The problem at the firm level is

$$\varphi^{\sigma-1} \cdot \Lambda(\mathcal{I}) - F(\mathcal{I}) \to \max_{\mathcal{I}} \tag{C.1.5}$$

Denote by $\mathcal{I}(\varphi)$ the optimal sourcing and assembly vector of locations. Consider two firms with $\varphi_H > \varphi_L$, then it should be

$$\varphi_H^{\sigma-1} \cdot \Lambda(\mathcal{I}(\varphi_H)) - F(\mathcal{I}(\varphi_H)) \geqslant \varphi_H^{\sigma-1} \cdot \Lambda(\mathcal{I}(\varphi_L)) - F(\mathcal{I}(\varphi_L))$$

$$\varphi_L^{\sigma-1} \cdot \Lambda(\mathcal{I}(\varphi_L)) - F(\mathcal{I}(\varphi_L)) \geqslant \varphi_L^{\sigma-1} \cdot \Lambda(\mathcal{I}(\varphi_H)) - F(\mathcal{I}(\varphi_H))$$

From these inequalities it follows that

$$\left[\varphi_H^{\sigma-1} - \varphi_L^{\sigma-1}\right] \cdot \left[\Lambda\left(\varphi_H\right) - \Lambda\left(\varphi_L\right)\right] \geqslant 0$$

which implies that $\Lambda\left(\varphi_H\right) \geqslant \Lambda\left(\varphi_L\right)$.

Proof of Proposition 5

If $\varphi \to \infty$, then the firm will activate all locations, and if $\varphi \to 0$, then the firm will not become global. From Proposition 6 it follows that $\sum_{i \in J} \left(\Psi_{hi}\left(\varphi\right)\right)^{(\sigma-1)/\theta^a} E_i P_i^{\sigma-1}$ is increasing in φ, so there exists a threshold φ_h^* such that firms headquartered in h with $\varphi > \varphi_h^*$ find it optimal to become global firms.

Proof of Proposition 6

Lemma 1 shows that under $\sigma - 1 \geqslant \theta^a \geqslant \theta^s/\alpha$, the profit function in (C.1.5) features increasing differences in different elements of vector $\mathcal{I} = (\mathcal{I}^s, \mathcal{I}^a, \mathcal{I}^g)$. Furthermore, it features increasing differences in $\left(\mathcal{I}_{ij}^r, \varphi\right)$ where $i, j = 1, 2, \ldots J$, $r \in \{a, s, g\}$. Invoking Topkis's monotonicity theorem, we can then conclude that for $\varphi_H \geqslant \varphi_L$ we have $\mathcal{I}(\varphi_H) \geqslant \mathcal{I}(\varphi_L)$. Therefore, $\mathcal{J}_h\left(\varphi_L\right) \subseteq \mathcal{J}_h\left(\varphi_H\right)$ and $\mathcal{K}_h\left(\varphi_L\right) \subseteq \mathcal{K}_h\left(\varphi_H\right)$.

Proof of Proposition 7

Consider the firms with headquarters in country h. Denote by $\tilde{\Lambda}\left(\mathcal{J}, \mathcal{K}\right) \equiv \Lambda\left(\mathcal{I}\right)$ where $\mathcal{I}_j^s = 1$ if $j \in \mathcal{J}$, $\mathcal{I}_j^s = 0$ if $j \notin \mathcal{J}$, and $\mathcal{I}_k^a = 1$ if $k \in \mathcal{K}$, $\mathcal{I}_k^a = 0$ if $k \notin \mathcal{K}$. Consider $j \notin \mathcal{J}$, $j \notin \mathcal{J}'$ and $\mathcal{J} \subseteq \mathcal{J}'$, $\mathcal{K} \subseteq \mathcal{K}'$. By definition in the proposition, $V_{h,j}^s(\varphi, \mathcal{J}, \mathcal{K}) = 1$ if

$$\varphi^{\sigma-1}\left[\tilde{\Lambda}\left(\mathcal{J} \cup j, \mathcal{K}\right) - \tilde{\Lambda}\left(\mathcal{J}, \mathcal{K}\right)\right] > f_{hj}^s$$

Under $\sigma - 1 \geqslant \theta^a \geqslant \theta^s/\alpha$, $\Lambda(\cdot)$ satisfies the increasing differences condition, so

$$\varphi^{\sigma-1}\left[\tilde{\Lambda}\left(\mathcal{J}' \cup j, \mathcal{K}'\right) - \tilde{\Lambda}\left(\mathcal{J}', \mathcal{K}'\right)\right] \geqslant \varphi^{\sigma-1}\left[\tilde{\Lambda}\left(\mathcal{J} \cup j, \mathcal{K}\right) - \tilde{\Lambda}\left(\mathcal{J}, \mathcal{K}\right)\right] > f_{hj}^s$$

Therefore, $V_{h,j}^s\left(\varphi, \mathcal{J}', \mathcal{K}'\right) = 1$. The proof for $V_{h,k}^a(\varphi, \mathcal{J}, \mathcal{K})$ is analogous.

C.2 Data Appendix

C.2.1 Matching the Census and BEA data

We build on the matching method developed by Brad Jensen and Fariha Kamal to merge the BEA and Census data. The BEA data contain several employer identification numbers (EINs) per firm, as well as name and address information. We merge these data to the Census Bureau's Business Register (BR) data, which includes EIN, name, and address information by establishment.

The matching method proceeds as follows. First we perform three merges of the BEA data to the BR separately on EIN, name and address. Not all three match successfully; we almost never find a match using the address merge. If all three methods match to a unique record in the BR, then we have found a match and we stop. However, if we find many possible matches in the BR then we follow a series of rules to choose the best match. To implement these rules we also use information on state, two-digit NAICS and employment which we have in both the BEA and BR data. We also prioritize BR records that are multi-unit and in the County Business Pattern (CBP) data. The rules proceed as follows:

1. the record that matches on EIN, name, state, and NAICS and is contained in CBP;
2. the record that matches on EIN, state, and NAICS and is contained in CBP;
3. the record that matches on the max number of EIN, name, state, and NAICS and is contained in CBP;

4. the record that matches on the max number of EIN, name, state, and NAICS, has closest ratio of BR employment to BEA employment, is contained in CBP and is multi-unit;

5. the record that matches on the max number of EIN, name, state, and NAICS, has closest ratio of BR employment to BEA employment, and is contained in CBP;

6. the match that is contained in the CBP, is multi-unit and has the closest employment ratio;

7. the match that is multi-unit;

8. the pair where the match was by EIN;

9. random.

For a subset of the largest MNEs, we use a clerical match provided by Fariha Kamal. In the event of conflicts with the original algorithm, we use the clerical matches which were done by hand. Finally, we use links between BEA firmids and Census firmids from the Business R&D and Innovation Survey.

C.2.2 Distinguishing US and Foreign-Owned Firms

The BEA data collected via survey BE-11 identifies the foreign affiliate activity by country and industry of firms operating in the United States (this is the outward data). The BEA data collected via survey BE-15 identifies firms operating in the United States that are owned by foreign parents, and provides information on the headquarter country of those parents (this is the inward data). In some cases, firms may exist in both of the BEA datasets. Consider a hypothetical example of foreign-owned firm that bases its North American headquarters in the United States. If those headquarters legally own affiliates in Canada or Mexico, they will report both outward and inward FDI activity. It is important

to note that the BEA outward data therefore includes some activity by foreign-owned firms.

In principle, the BE10 file (the parent file), contains a variable BE15_id, which can be used to determine a firm's ultimate ownership country. In practice however, this approach does not seem to work using the Census firm definitions because it appears to overstate the extent of foreign ownership in the US relative to public data posted by the BEA. This overstatement is likely due to the fact that, for a small set of very large firms, the Census firm definition tends to encompass a larger number of EINs than the BEA definition. As a result, if we designate these larger Census firms as foreign-owned whenever they contain some portion of activity that the BEA classifies as foreign-owned, they are much larger than the BEA assignment and firm definition imply. Examination of the data suggests that in some cases, these larger Census firms are foreign-owned, while in other cases they seem to be US-owned.

To address these issues and distinguish between US and foreign MNEs as systematically as possible, we supplement the BEA data using the Census Bureau's Company Organization Survey (COS), which asks firms whether they are majority-owned by a foreign firm and whether they own foreign affiliates. Before relying on the COS data, we analyze the accuracy of these previously unused variables by comparing the related party trade status and shares of firms that the COS identifies as foreign-owned or owning foreign affiliates. This analysis is available as technical documentation inside our project and provides reassuring evidence that the COS data do indeed contain relevant information for identifying MNEs.

For the subset of firms that appear in both the outward and inward BEA data and which the BEA classifies as majority foreign-owned, we use the COS and BR data to distinguish whether they are most likely US MNEs or foreign-owned firms when using the broader Census firm definition. First, we use the COS data and identify firms as "Foreign-owned" whenever those firms report that they are majority-owned by a foreign firm in the COS. (Note that in this case, the BEA and Census COS data agree so this

seems conservative.) Second, for firms that are missing the COS data, we aggregate the BEA data to the BEA-EIN level and calculate the share of the firm's employment at establishments that belong to EINs that the BEA flags as foreign-owned. We then identify firms as "foreign-owned" if their share of US "foreign-owned" employment is greater than 49 percent according to the Census firm definition. Finally, we classify the remaining firms as "US MNEs."

To summarize:

1. All firms that appear only in the BEA inward data are classified as "foreign-owned" firms,

2. All firms that appear only in the BEA outward data are classified as "US MNEs",

3. All firms that appear in both the BEA outward and inward data, and for which the firm reports the United States as the ultimate owner country to BEA are classified as "US MNEs",

4. For firms that appear in both the BEA inward and outward data, and for which the firm reports majority-ownership with the ultimate owner country not as the United States to BEA:

    - Classify as "foreign-owned" firms if they report being majority foreign-owned in the COS data,
    - Classify as "Foreign-owned" if are missing from the COS data but have greater than 49 percent of their US employment per the Census firm definition in establishments with EINs that are present in the BEA inward data,
    - Classify remaining firms as "US MNEs"

This approach results in approximately 7,600 foreign-owned MNEs and 2,800 US MNEs. These firms' share of employment, sales, and trade are reported in Table 3.1.

C.2.3 BEA Country Classifications

When matching the Census data to the BEA data, we find several countries that are aggregated in the BEA data (e.g., the French Islands, Kiribati, etc.). We aggregate the import data to match the level of aggregation in the BEA data. Generally gravity are only available for the main country in those cases. If there are multiple countries with gravity data, we use the data for the one with the largest population (e.g., in the case of Australia, Cocos Island, Norfolk Islands, Heard and McDonald Islands, etc., we use the gravity data on Australia).

C.2.4 Sample Description

The sample of firms in the paper is all firms with one or more manufacturing establishments in 2007, and/or with foreign affiliate manufacturing activity. We define foreign manufacturing activity as observations in the BEA outward data for which there is positive employment or sales activity in manufacturing. Table 3.3 shows the aggregate amount of all majority-owned affiliate activity, as well as the subset in manufacturing. The difference in these two amounts is likely related to retail, distribution, or other sectors.

We drop firms that have zero employment or sales in the United States. We do not drop the Census of Manufactures administrative records. Although these observations tend to have imputed information for values like inputs, they are surprisingly important for matching the LBD/EC data to the Customs Transactions database. Since our goal is to capture those foreign activities as completely as possible, we choose to retain these records.

We use the LFTTD data which is matched from the LBD to the trade transactions data by the Center for Economic Studies. Import data match rates are generally quite high, with the exception of nine countries like Djibouti, Tonga, etc. Since the focus of

the paper is on manufacturing, we drop mineral imports and exports (HS2=27) from our analyses.

As in AFT, we define a firm's total manufacturing inputs as: inputs mat merch and prod worker wages (assuming ww is prod worker wages).

C.2.5 Additional Tables

In Table C.1 we present summary statistics on imports and exports for non-manufacturing firms.

C.2.6 Gravity Regressions for Firm Exports

We also estimate a variant of equation (3.3.1) where the dependent variable is an indicator for whether or not the firm exports from country c. Table C.2 presents these results.

We also estimate a variant of equation (3.3.2) where the dependent variable is the log of firm exports to country c. Table C.3 presents these results.

C.2.7 Aggregate gravity regressions

In this subsection, we aggregate the data used in Section 3.3 to estimate the firm-level gravity regressions to show that the standard relationships that have become well-known in the trade literature are also present in our data.

Table C.1: Trade statistics for non-manufacturing firms in 2007, by import and MNE status

Panel A: Imports	$\frac{Imports}{Sales}$	$\frac{Importers}{Firms}$	$\frac{RPImporters}{Importers}$	$\frac{RPImports}{Imports}$
Panel A: Imports				
Importers	0.06	1.00	0.14	0.23
Foreign-Owned	0.09	0.52	0.68	0.70
US MNEs				
w/o Manuf Aff	0.02	0.73	0.63	0.07
w/Manuf Aff	0.07	0.67	0.90	0.52
Non-Manuf Total	0.03	0.05	0.15	0.32
Panel B: Exports	$\frac{Exports}{Sales}$	$\frac{Exporters}{Firms}$	$\frac{RPExporters}{Exporters}$	$\frac{RPExports}{Exports}$
Domestic	0.01	0.02	0.06	0.10
Importers	0.02	0.34	0.14	0.11
Foreign-Owned	0.02	0.44	0.46	0.26
US MNEs				
w/o Manuf Aff	0.01	0.73	0.56	0.18
w/Manuf Aff	0.02	1.00	0.53	0.22
Non-Manuf Total	0.01	0.04	0.10	0.14

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward, and BEA outward datasets. Table presents the ratios of firm imports to sales and exports to sales, the share of each firm type that imports and exports, the share of related-party importers and exporters, and the share of related-party imports and exports. "Foreign-owned" are firms that are majority-owned by a foreign firm. "US MNEs" are non-foreign-owned firms with majority-owned foreign affiliates. "Domestic" firms are non-multinationals that do not import. "Importers" are non-multinationals that import.

Table C.2: Extensive margin export regressions

Dependent variable is $exporter_{frc} = 1$ if firm exports to country c

	(1)	(2)	(3)	(4)	(5)	(6)
Common Language$_c$	0.021**	0.021**	0.021**	0.021**		
	(0.009)	(0.009)	(0.009)	(0.009)		
$log(distance_c)$	-0.023	-0.003	-0.003	-0.003		
	(0.020)	(0.012)	(0.011)	(0.011)		
$log(GDP_c)$	0.023***	0.022***	0.022***	0.022***		
	(0.002)	(0.002)	(0.002)	(0.002)		
Contiguous$_c$		0.229***	0.226***	0.226***		
		(0.013)	(0.013)	(0.013)		
Affilliate$_{fc}$			0.474***	0.512***	0.423***	0.463***
			(0.036)	(0.038)	(0.032)	(0.035)
Foreign-Owned$_{fc}$			0.590***	0.593***	0.518***	0.521***
			(0.040)	(0.040)	(0.043)	(0.043)
Affiliate in Region$_{fr}$				0.081***		0.087***
				(0.020)		(0.020)
Foreign in Region$_{fr}$				0.031**		0.035**
				(0.014)		(0.014)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Region FEs	Yes	Yes	Yes	Yes	No	No
Country FEs	No	No	No	No	Yes	Yes
Adj. R2	0.222	0.227	0.233	0.234	0.266	0.267
Observations (000s)	7,230	7,230	7,230	7,230	7,230	7,230

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward, and BEA outward datasets. Dependent variable is an indicator for whether firm f imports from country c in region r. Sample is all firms with manufacturing establishments in the United States in 2007 that export multiple countries. Observations in 1000s and rounded per Census disclosure rules. The mean and standard deviation of the dependent variable are 0.048 and 0.215, respectively. There are 188 countries in this sample. Standard errors two-way clustered by firm and by country. *,**, *** denote p<0.10, p<0.05, and p<0.01, respectively.

Table C.3: Intensive margin export regressions

Dependent variable is $log(exports_{frc})$

	(1)	(2)	(3)	(4)	(5)	(6)
Common Language$_c$	0.265***	0.267***	0.254***	0.254***		
	(0.080)	(0.078)	(0.073)	(0.073)		
$log(distance_c)$	-0.408***	-0.207*	-0.205*	-0.203*		
	(0.110)	(0.113)	(0.108)	(0.109)		
$log(GDP_c)$	0.407***	0.391***	0.363***	0.364***		
	(0.019)	(0.020)	(0.019)	(0.019)		
Contiguous$_c$		0.402***	0.394***	0.400***		
		(0.118)	(0.112)	(0.113)		
Affilliate$_{fc}$			1.973***	2.049***	1.906***	1.993***
			(0.107)	(0.103)	(0.108)	(0.102)
Foreign-Owned$_{fc}$			1.301***	1.280***	1.306***	1.286***
			(0.147)	(0.162)	(0.140)	(0.155)
Affiliate in Region$_{fr}$				0.142*		0.163**
				(0.079)		(0.078)
Foreign in Region$_{fr}$				-0.115		-0.112
				(0.121)		(0.122)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Region FEs	Yes	Yes	Yes	Yes	No	No
Country FEs	No	No	No	No	Yes	Yes
Adj. R2	0.396	0.397	0.412	0.412	0.42	0.42
Observations (000s)	350	350	350	350	350	350

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward, and BEA outward datasets. Dependent variable is the log of imports by firm f, from country c, in region r. Sample is all firms with manufacturing establishments in the United States in 2007 that export to multiple countries. Observations rounded per Census disclosure rules. Standard errors two-way clustered by firm and by country. The mean and standard deviation of the dependent variable are 4.098 and 2.153, respectively. There are 188 countries in this sample. *,**, *** denote p<0.10, p<0.05, and p<0.01, respectively.

Table C.4: Aggregate gravity regressions

	$log(imports_c)$		$log(exports_c)$	
	(1)	(2)	(3)	(4)
Common Language$_c$	1.129***	1.127***	0.936***	0.935***
	(0.274)	(0.275)	(0.165)	(0.166)
$log(distance_c)$	-0.991***	-0.984***	-1.107***	-1.106***
	(0.296)	(0.310)	(0.177)	(0.185)
$log(GDP_c)$	1.346***	1.345***	1.126***	1.126***
	(0.055)	(0.056)	(0.033)	(0.033)
Contiguous$_c$		0.11		0.016
		(1.310)		(0.800)
Adj. R2	0.78	0.78	0.87	0.87
Countries	182	182	188	188

Source: 2007 Economic censuses, LBD, LFTTD, BEA inward, BEA outward, and CEPII gravity datasets. The sample is all firms with manufacturing establishments in the United States in 2007 that import from multiple countries. The samples here are identical those used in the firm-level gravity regressions in Section 3.3. *,**, *** denote p<0.10, p<0.05, and p<0.01, respectively.