International Trade and Firm Performance

by

Ruoying Wang

B.Sc., Tsinghua University, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

International Trade and Firm Performance

submitted by Ruoying Wang in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Economics

Examining Committee:

Matilde Bombardini
Supervisor

Hiro Kasahara
Supervisory Committee Member

Viktoria Hnatkovska
Supervisory Committee Member

Brian Copeland
University Examiner

John Ries
University Examiner

Additional Supervisory Committee Members:

Paul Beaudry
Supervisory Committee Member

Supervisory Committee Member

Abstract

This dissertation is a collection of three essays that study the effect of opening to trade, especially opening to the import market, on firm performance.

The first essay (Chapter 2) explores the link between innovation and import competition in China, a country that during the period we study (2000-2007) saw both a rapid increase in patenting and a lowering of import barriers due to accession to the WTO. Combining manufacturing firm survey data with customs and patent data, we find that import competition encouraged innovation, but only for the most productive firms. These top firms saw an increase in patenting rate of 3.6% for every percentage point drop in import tariffs. The result is quantitatively similar whether we use a sector-wide tariff on output or a weighted tariff at the firm level as a measure of import competition. Consistent with the main finding, top firms also feature increased R&D expenditures and an increase in domestic sales following import liberalization.

To analyze the mechanism and welfare implications underlying our empirical findings about China, the second essay (Chapter 3 and 4) builds a model. Firms engage in monopolistic competition across varieties and neck-and-neck competition within each variety. An increase in the neck-and-neck competition reduces the expected profit of not innovating, thus encouraging firms to innovate more to escape the competition. We analyze the efficiency and utility implications using a simple version of the model in Chapter 4.

The third essay (Chapter 5) examines the relationship between Canadian manufacturing firms' import behavior and their performance. The focus is on two aspects of import structure, input variety and the dynamics of import relationships. Firms importing more products from a larger set of suppliers tend to be larger, more productive, and more successful in export markets. Not only the number, but also the duration of supply relationships matters. Firms maintaining a higher share of continuous supply relationships also benefit in size and productivity. These results suggest that the breadth and depth of the import network are relevant factors for the performance of Canadian manufacturers.

Lay Summary

This dissertation studies two aspects of the importing market: the effect of import competition on innovation, and the effect of buyer-supplier relationship on importer productivity. Chapter 2 studies how the intensified import competition after China's accession to the WTO affect Chinese manufacturing firm's innovation behavior. We found that import competition induced more innovation from the most productive firms. Chapter 3 builds a model that explains the incentives behind firms' innovation reactions, while Chapter 4 analyzes the efficiency and welfare implications of the model. Chapter 5 studies how the structure of buyer-supplier relationship affect firm performance. Employing detailed buyer-supplier information for Canadian importers, we find that higher import variety and deeper relationship with foreign suppliers are beneficial to firms in terms of size, productivity, and performance in the export markets. We propose an empirical method to identify causality.

Preface

This dissertation is based on two unpublished working papers.

Chapter 2, 3 and 4 are based on co-authored work with Matilde Bombardini, and Bingjing Li. I started the project and conducted initial data cleaning and exploration after getting access to the Chinese patent database. Matilde guided the big picture of the project and the model setup. Bingjing and I conducted model simulations and implemented the various empirical specifications. I performed additional robustness checks and wrote the first draft of the paper.

Chapter 5 is based on co-authored work with Matilde Bombardini, Keith Head, and Maria D. Tito. Matilde, Keith and Maria started the project, and I joined later. I contributed to running empirical specifications in Statistics Canada in Ottawa, and designing the identification strategy of our baseline regression. The views presented in the chapter represent those of the authors and do not necessarily coincide with those of Statistics Canada. The contents of this chapter have been subject to vetting and pass the Disclosure Rules and Regulations set forth by Statistics Canada.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

I would like to express my sincere gratitude for the guidance and support I received throughout my PhD studies, from the Vancouver School of Economics, and especially from my committee members, Professor Matilde Bombardini, Professor Viktoria Hnatkovska, Professor Paul Beaudry, and Professor Hiroyuki Kasahara.

I owe particular thanks to my supervisor, Matilde Bombardini, who led me into the field of International Trade, and helped me grow into an empirical economist. Matilde's unfailing encouragement has supported me through the most difficult times. I very much appreciate the last year of my PhD career, working closely with Matilde at Stanford University, which gave me so much exposure to the academic world outside UBC. I would like to thank Professor Viktoria Hnatkovska and Professor Amartya Lahiri, from whom I learnt a lot through working as an RA and eventually co-authoring the paper on China's One Child policy and structural change. I feel lucky to co-work with Maria D. Tito and Bingjing Li, who shared with me their skills and insights. I am also grateful for the helpful comments I received from Professor Keith Head, Professor Nancy Gallini, Professor Tomasz Swiecki, Professor Jesse Perla, and participants at the UBC Trade Study Group.

Last but not the least, I would like to thank my parents and my husband. I could not have the luxury of delving into the ivory tower without them standing beside me.

Chapter 1

Introduction

This dissertation is a collection of three essays that study the impact of import liberalization on firm performance. The first essay examines the effect of import liberalization on Chinese firm's patenting behavior and discovers a novel non-linear relationship between import competition and firm innovation. Based on this observation, the second essay formulates a trade model to study why import competition can generate heterogenous innovation incentives among firms and examines the model's implications on resource allocation and aggregate welfare. The third essay studies the effect of the breath and depth of import relationships on firm performance, using the Canadian manufacturing firm data. In this chapter, we summarize the methodology and main findings of each essay but postpone the discussion of literature and contribution to the respective chapters.

The first essay of this dissertation (Chapter 2) studies how import liberalization after China's accession to the World Trade Organization related to its rapid growth in innovation during the post-accession period. We use the mandated drop of import tariff by the WTO as the policy shock to foreign competition and use invention patent applications in China to measure innovation. We find that import competition encouraged innovation among the most productive firms. The magnitude of this competition channel is comparable to roughly 10% of the total increase in patenting rate among these firms during the period 2003-2007. We exploit the variation of tariff changes at the industry-year level, and control for other possible channels through which trade could affect innovation, such as access to a larger exporting market and access to better imported intermediate inputs. We conduct a falsification test by regressing the pre-policy patent activity changes during 1998-2001 on future import tariff changes during 2001-2005 and find no significant relationship between the two. This result suggests that (1) there is no differential pre-trend in the propensity to innovate related to future tariff changes between more and less productive firms; and (2) firms do not respond to expectations in tariff decrease before they see the actual increase in foreign competition. In addition to reduced form regressions, we also conduct a two-stage control function estimation where we use the import tariff as instrument for the import volume. Consistent with the previous findings, the most productive firms patent more in the face of increased import competition caused by the tariff drop.

The empirical finding that some firms can innovate more as a result of import competition cannot be easily explained by standard trade models with heterogeneous firms. Under monopolistic competition, the profit difference between improving quality or technology always decreases when there is higher import competition (the rent destruction effect of competition). Thus, in such models, import liberalization should have a negative effect on innovation, holding other conditions constant. On the other hand, there is a separate class of models that can generate a positive effect of competition on innovation. For example, Aghion et al. (2001, 2005, 2009) find that in models with neck-and-neck competition, innovation helps firms survive and maintain their market share. So, when competition intensifies, firms will find it optimal to invest more in innovation. Such incentive is called in that literature the "escape competition effect". However, these models usually eliminate any general equilibrium price effects, which is exactly the reason why trade models generate a uniformly negative effect. Therefore, it is useful to introduce the escape competition effect into an otherwise standard trade model. And this is what we do in Chapter 3.

More specifically, in our model, firms engage in monopolistic competition across varieties and neck-and-neck competition within each variety. Therefore, the escape competition effect and the rent destruction effect both exist in our model. When import competition intensifies, the positive escape competition effect dominates the negative rent destruction effect for the more productive firms. And the rent destruction effect dominates for the less productive ones.

In Chapter 4, we further study the welfare implications of the heterogeneous reaction of firms. Using a simplified framework, we find that, relative to a constrained social optimum, the decentralized equilibrium always features under-investment, because the consumer gains form innovation are not internalized by firms. The second finding is that, because of under-investment, whether there is additional welfare gains from innovation depends on whether the escape competition effect prevails over the rent destruction effect.

In Chapter 5, we use transaction-level administrative data from Statistics Canada to study how decisions regarding import relations (number of suppliers per good, and the duration of the relationship with the suppliers) affect firm performance. An immediate endogeneity issue is that the import relationship decision is itself affected by firm performance. This problem is in spirit very similar to identifying input elasticities when estimating a production function. Therefore, we borrow from that literature and formulate a control function approach to deal with endogeneity. The key condition for identification is that the import relationship is not uniquely determined by the unobserved productivity shocks. In this way, the observable productivity shock can be proxied by the input choices such as capital and material. Armed with the identification strategy, the rich data structure allows us to explore the effect of different features of buyer-supplier relationships that have not been previously explored in the literature. In particular we find that a larger variety of suppliers and a longer duration of buyer-supplier relationships affect the performance of the Canadian buyers positively.

Chapter 2

Import Competition and Innovation: Evidence from China

2.1 Introduction

The link between innovation and prosperity is not only at the heart of a large literature on endogenous growth, but also the target of much attention by policymakers. A recent paper by Akcigit et al. (2017) documents a long-run relationship between innovation and growth in the United States. This broad consensus on the value of innovation stands in contrast with the disagreement on what its main drivers are. We focus on a question that has proven particularly difficult to settle, that is the link between competition, in particular import competition, and innovation. In fact, despite recent evidence on the effect of trade on innovation in several developed countries (Aghion et al., 2017; Autor et al., 2017; Bloom et al., 2016), no clear consensus has emerged on whether import competition encourages or discourages innovation. This paper contributes to this discussion by analyzing firm level evidence from China. We ask the following questions: what is the impact of import competition on Chinese firm's capacity to innovate? And what firms are affected the most?

China serves as an interesting case to study the relationship between import competition and innovation because it experienced both a rapid growth in patenting and intensified foreign competition after its accession to the WTO in December 2001. From 2001 to 2007, invention patents filed in the State Intellectual Property Office (SIPO) of China grew at an average annual rate of 25 percent, comparing to 6 percent in the US patent office (USPTO). In 2007, the number of total patents filed in SIPO has reached 53 percent that of the USPTO. During this post-WTO period, China experienced large drop in import tariff barriers and other kinds of local protections mandated by WTO, which significantly increased the presence of foreign competition. The effectively applied tariff dropped by 6.2 percentage points, from 0.166 in 2001 to 0.104 in 2005. The Non-Tariff Barriers (NTBs) were quite low during that period[1]. Total import quadrupled from 243.5 billion USD in 2001 to 956.1 billion USD in 2007. We exploit the variation in the import tariff decreases across industries and over time to identify the shock to import competition on firms. It is worth noting that, even though China's first application to the GATT dated to 1986, the schedule of tariff changes was not known until September 2001. The negotiation took 18 meetings between 1996 and 2001 and was characterized by Vice Minister LONG Yongtu, Head of the Chinese Delegation,

[1]The NTBs used in China during this period (2000-2007) were mainly the anti-dumping duties. Among all HS6 products, on average, only 0.17 percent of the products were subject to some form of anti-dumping duties. This number was 0.74 percent for the US.

in his Statement on Sep 17, 2001: "The complexity and difficulty of this process are beyond the imagination of almost everybody." It is unlikely that anybody knew the timing and extent of tariff cuts and much less the impact of those tariff cuts on imports.

Our firm level variables come from the Chinese Annual Survey of Manufacturing Firms, which is matched to the patent and customs data. The matched firms account for over two thirds of total patents filed by Chinese enterprise assignees during 2003-2007. We estimate the elasticity of firm patent application on two-period lagged industry output tariff through a Poisson count data model, controlling for industry fixed effects and time trends. Output tariff is measured as the average import tariff faced by products that the industry (firm) produces. The idea is that, when the import tariff decreases, there are more competing foreign goods (final or intermediates) in the domestic market. Thus, a decrease in the import tariff proxies for an increase in total import of the industry. Moreover, we use the WTO accession tariff that was scheduled at the time of China's WTO entry, instead of using the applied tariff for each year. In this way, we try to minimize the concern that concurrent tariffs may be affected by industry lobbying that is correlated with productivity or innovation, our outcome variable.

We find that for firms above the 75th percentile in terms of productivity, a one percentage point decrease in output tariff could induce about 3.6 percent increase in patenting. This increase is due to both strengthening firms' core technology and enlarging the technology scope.

There are several challenges in our measurement and identification of the problem at hand. First, we measure innovation with patent applications in SIPO. Of course, patent application is not the only output of innovation. Many innovative activities such as improvement in management or business model is not patentable, and firms may prefer to keep some new formula secrecy[2]. However, survey evidence shows that all outputs of innovation are positively correlated (Hall et al., 2014; Moser, 2013), and comparing to productivity, patenting is a more direct and precise way of measuring technology progress at the firm level (Griliches, 1990; Nagaoka et al., 2010). Therefore, we use patent application count as our benchmark measure of innovation and maintain the assumption that patent count is a sufficient statistic in measuring firm innovation.

Another concern is that the measured increase in patenting is totally driven by the change in the propensity to patent. For example, studies have shown that stronger IP protection due to entrance into the WTO would boost domestic patent filing, and the patent system may shift firms' innovation effort from unpatentable to patentable products (Qian, 2007; Arora et al., 2015). There definitely exist nation-wide trends in IP protection and propensity to patent in China. The question is whether the propensity to patent due to IP protection has differentially affected sectors that experienced different degrees of import liberalization. Because we show, similarly to Brandt et al. (2017), that tariff cuts are unrelated to observable pre-trends like productivity growth, we do not find the conditions for our regressions to pick up this type of spurious correlation. Furthermore,

[2] Hall et al. (2014) reviewed the literature on choice between formal IP and secrecy. They concluded that although the choice is made strategically and is affected by various industry and market characteristics, empirical evidence shows that secrecy and IP are usually complements. And the choices within formal IP — patent, trademark and copyrights — are used as complements as well.

if patenting propensity differentially changes by aggregate sectors, our sector-time dummies will absorb different trends in industry patenting propensity.

Our empirical findings point to an "escape competition" motif for firms in the face of intensified competition, which is stronger for the top firms. We postpone a detailed theoretical exploration to the next Chapter. For now, it is helpful to state the intuition, which we borrow from Aghion et al. (2001) and Aghion et al. (2009). Competition introduces two impacts on firm profitability. On the one hand, it increases the probability that the domestic firms be replaced by the foreign competitor. If innovation could help the firm to retain its market, then there is more incentive to innovate when there is more foreign competition. On the other hand, the entering of foreign competitors erodes away markups by firms, thus induces a rent destruction effect, which could reduce the incentive to innovate. The net effect of competition depends on which force dominates. In the case of China, we find that the escape competition effect dominates the rent destruction effect.

Our paper is related to several strands of literature. First, it is related to studies about China's gains from accessing the WTO. In addition to the usual gains from trade such as selection through exports, or access to more imported varieties, evidence has shown that accessing the WTO helped correct resource misallocation and accelerated the market reforms that is sometimes difficult to implement within political constraints (Khandelwal et al., 2013; Lu and Yu, 2015). In a recent and very related paper, Brandt et al. (2017) studies the effect of lowering import tariffs on industry average productivity and mark-up. They find that when import competition intensifies, there is a decrease in output price level and mark-up, and an increase in aggregate productivity. While confirming their findings, we go one step further to study the innovation channel of productivity increase, among other potential channels such as purchasing new machines.

The second strand of literature looks empirically at channels through which trade could impact firm's capacity to innovate. Four channels have received most attention. First, trade could encourage innovation by technology diffusion (Coe and Helpman, 1995; Eaton and Kortum, 1999; Buera and Oberfield, 2017; MacGarvie, 2006). Second, getting access to bigger markets due to export liberalization has been found to induce firms to switch to skill intensive technology (Bustos, 2011), increase R&D spending (Aw et al., 2011), and engage in more innovation (Aghion et al., 2017; Lim et al., 2017). Thirdly, import liberalization could enable access to better imported inputs, which helps to enhance knowledge diffusion (MacGarvie, 2006), complements R&D spending (Bøler et al., 2015), and induces quality upgrading (Fieler et al., 2016). While the positive effects of the aforementioned channels are quite unambiguous, the fourth channel which is the focus of this paper — import competition — has received more mixed evidence. Bloom et al. (2016)[3] and Teshima (2008) find positive effects of import competition on innovation for European countries and Mexico,

[3]Although Bloom et al. (2016) also find a positive effect of import competition on innovation, the underlying channel that they propose is quite specific to the situation of Europe, thus difficult to generalize to the case of China. In particular, they motivate their empirical findings by a model of "trapped factors": since international competition usually comes from the less developed countries, openning to trade lowers the opportunity cost for European firms to switch into more novel products. For China, on the other hand, competition usually comes from the more advanced countries so that the cost of innovation actually should increase due to competition, caritas paribus. The model we propose in Chapter 3 captures this feature through the rent-destruction effect.

respectively. Autor et al. (2017), on the other hand, documents a drop in patent production in the US manufacturing sector in response to the rising Chinese competition. Our empirical analysis provides evidence on China itself. Our analysis on the heterogeneous effects also suggests that the net effect of foreign competition depends on the productivity of the firm and whether the firm could overcome foreign competition through innovating.

Our paper is also related to the literature that explores the reasons behind the rapid increase in patenting in China after 2000. The most studied causes are increased investment in R&D (Wei et al., 2017; Hu et al., 2005; Hu and Jefferson, 2009), improvement in Intellectual Property Right (IPR) protection (Ang et al., 2014), ownership reforms, government's pro-patenting policies, and FDI. Fang et al. (2017) find that privatization of state-owned firms motivates more patenting, especially in prefectures with higher IPR protection. Xie and Zhang (2015) find that rising wages have propelled labor-intensive sectors to become more innovative, and firms in female-intensive industries have exhibited more innovations than those in male-intensive industries. Jiang et al. (2018) study the technology transfer from foreign joint venture partners to the Chinese partners. They find that industries with an increase in foreign JV presence experienced an overall increase in TFP growth. While we focus on the effect of import competition and market structure, we take into consideration these other forces through controlling for region and ownership characteristics.

The remainder of this chapter is organized as follows: Section 2.2 describes the data and summary statistics. Section 2.3 shows our empirical framework, and Section 2.4 discusses the empirical results. Section 2.5 concludes.

2.2 Data

2.2.1 International trade

Industry level

Our baseline measure of import competition uses the average import tariff China imposes on the products of each four-digit industry. The import tariff information is obtained from China's WTO accession document, which specifies the tariff targets for each year since 2001 for each six-digit HS product. Figure 2.1 plots the actually applied tariff against the bounded tariff for 2001-2005. As mentioned in the introduction, China's WTO negotiation ended in September 2001, and then it quickly entered the WTO in December 2001. So for almost all 2001, Chinese imports were not subject to any WTO restrictions. From Figure 2.1 we can see that, in 2001, only 32% of the 5,085 six-digit HS products complied with the WTO bounded tariff. This rate quickly raised to 97.5% in 2002 and remained above 97% thereafter. Figure 2.2 shows the average of the WTO accession tariff and the applied tariff during 1997-2007. During this period, the average bounded tariff dropped from 0.1372 to 0.1002, and the average applied tariff dropped from 0.1588 to 0.0982.

Figure 2.1: Actually applied tariff and bounded tariff

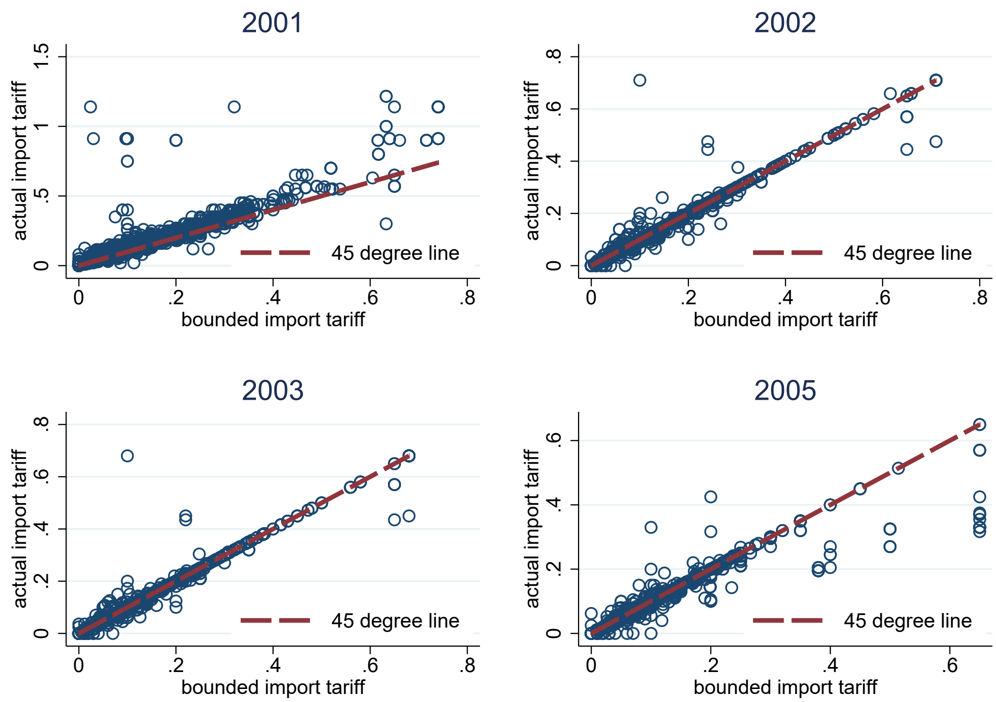
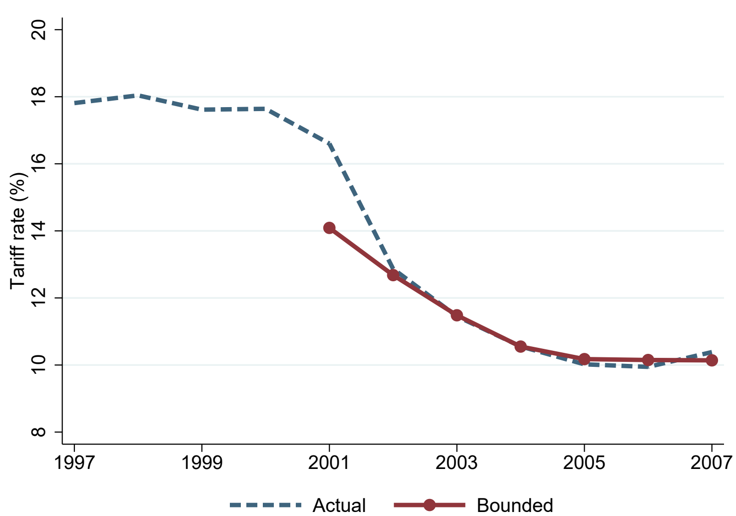


Figure 2.2: Tariff trends



We map the products into the China Industrial Classification (CIC) system at the four-digit level (424 sectors), using the concordance developed in Brandt et al. (2017)[4], and take simple

[4]Their concordance is based on the HS-CIC concordance table constructed by the National Bureau of Statistics (NBS).

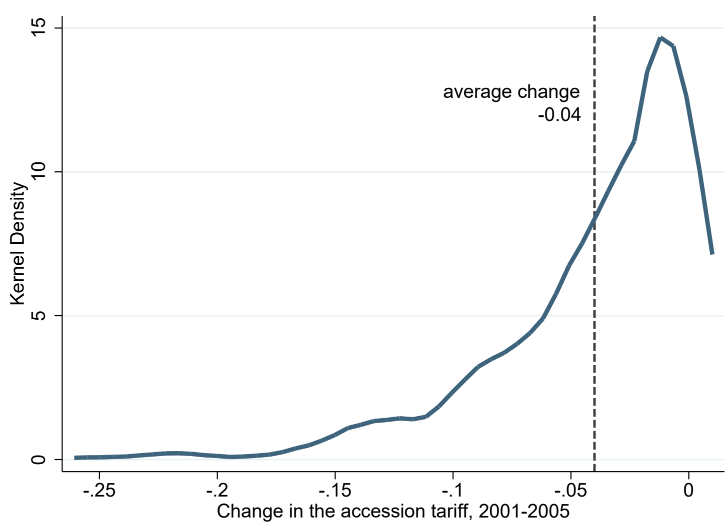
average[5] to arrive at the industry level output tariff. Then we take log plus one to arrive at our measure of log output tariff,

$$\tau_{st}^{\text{output}} = \log\left(1 + \frac{1}{H_s}\sum_{h=1}^{H_s} \text{tariff}_{ht}^{\text{import}}\right). \tag{2.1}$$

For product h that is matched to industry s, $\text{tariff}_{ht}^{\text{import}}$ is the WTO specified import tariff for China in year t. H_s denotes the total number of six-digit HS products within industry s. Figure 2.3 shows that the tariff drop between 2001-2005 varied across industries.

We use the accession tariff instead of the actually applied tariff, because the applied tariff may be subject to the same contemporaneous forces that affect firms' innovation incentives. While the pre-determined accession tariff is exempt from such endogeneity concerns, we require it to be also independent from any expectations in future innovation trends. We test such restriction by regressing the change in the accession tariff from 2001 onward on the change in patenting or firm productivity during 1998-2000. Table 2.1 shows that we cannot reject the hypothesis that the accession tariff is not correlated with pre-trend in patenting or productivity growth.

Figure 2.3: Distribution of tariff changes across industries



[5] There is a possible bias with trade volume weighted averages: Trade volume is negatively correlated with tariff levels. Taking weighted average will tend to give more weight to the most liberalized product lines and thus underestimate the change in effective protection and could cause an upward bias in the estimated effect of trade liberalization.

Table 2.1: WTO accession tariff and initial period growth rates

	(1) $\tau_{2002}^{output} - \tau_{2001}^{output}$	(2) $\tau_{2003}^{output} - \tau_{2001}^{output}$	(3) $\tau_{2004}^{output} - \tau_{2001}^{output}$	(4) $\tau_{2005}^{output} - \tau_{2001}^{output}$
Panel A: Initial TFP growth				
tfp_{2000}-tfp_{1998}	0.001	0.001	-0.001	-0.003
	(0.005)	(0.009)	(0.013)	(0.015)
R^2	0.330	0.344	0.373	0.373
Obs	424	424	424	424
Panel B: Initial patent growth				
patent_{2000}-patent_{1998}	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.333	0.349	0.376	0.375
Obs	424	424	424	424

Note: *** p<0.01, ** p<0.05, * p<0.1

Any effect that industry output tariff has on the industry could affect other industries through the input-output linkages. If the upstream industries experienced higher competitive pressure, the downstream firms would likely to get cheaper and higher quality inputs. We capture such effect through the change in input tariff. Specifically, we define input tariff as

$$\tau_{st}^{\text{input}} = \log\left(1 + \sum_{k} \nu_{sk} \text{tariff}_{kt}^{\text{output}}\right). \tag{2.2}$$

The input share of industry k good used in industry s, υ_{sk}, is obtained from China's 2002 input-output table. The sum of υ_{sk} is smaller than 1, and the other shares include non-manufacturing inputs, labor and capital inputs. The effects of these other inputs are subsumed into the error term of our specification and is assumed to be independent from our main explanation variable: change in trade induced competition in the manufacturing sector. The input tariff captures any competition effects transferred from import liberalization in input industries. This is what Fieler et al. (2016) call the magnifying effect in their quantitative analysis.

For the importers, the input tariff calculated in equation (2.2) also captures the direct effect of getting access to cheaper or better foreign inputs. Therefore, we also control for an importer dummy and the interaction between the input tariff and the dummy. A firm is defined as importer after its initial appearance in the customs importer registry.

Another channel through which trade could affect innovation is the market size effect brought by export liberalization as studied in Aghion et al. (2017) and Lim et al. (2017). We control for such effect through export demand shock from other countries, which is defined as:

$$E_{st}^{\text{demand}} = \sum_{h,c} \frac{X_{hc,2000}}{X_{s,2000}} \log M_{hct}, \tag{2.3}$$

where M_{hct} denotes country c's import from the world other than China of product h at time t. After taking log, we weigh the country-product demand shocks by the export share of China during year 2000. $X_{hc,2000}$ denotes China's export of product h to country c in industry s in year 2000.[6]

Finally, we focus on year 2001-2005 because this is the period that experienced highest tariff change after the WTO accession for most products and industries.

Firm level

In addition to the industry level measure of trade shocks, we also construct firm level trade shocks.

$$\tau_{ist}^{\text{output}} = \ln\left(1 + \sum_{h} \frac{X_{ih,t-1}}{\sum_{h'} X_{ih',t-1}} \text{tariff}_{ht}^{\text{import}}\right); \tag{2.4}$$

$$E_{ist}^{\text{demand}} = \ln\left(1 + \sum_{h,c} \frac{X_{ihc,t-1}}{\sum_{h'c'} X_{ih'c',t-1}} \log M_{hct}\right); \tag{2.5}$$

$$\tau_{ist}^{\text{input}} = \ln\left(1 + \sum_{h} \frac{M_{ih,t-1}}{\sum_{h'} M_{h'i,t-1}} \text{tariff}_{ht}^{\text{import}}\right). \tag{2.6}$$

where X_{iht-1} and M_{iht-1} denotes firm i's export and import in product h in the previous year, respectively. To reduce missing values, for firms that import (export) with gap years, we use the most recent year that it had imported (exported) to calculate the weight. The firm level export demand shock is constructed from the product-country level export demand faced by China, where c, again, denotes destination countries.

2.2.2 Patent and other firm-level variables

The firm level sample of our study comes from the Annual Survey of Manufacturing Firms conducted by the National Bureau of Statistics (NBS) of China (Hereafter referred to as the NBS data), 1998-2007. The survey covers all state-owned firms, and private firms with annual sales larger than 5 million RMB. It has become a standard data set for studying firm level behavior in China's manufacturing sector (Brandt et al., 2012; Hsieh and Song, 2015). In addition, we match the NBS data with the customs data and patent data to get firm level trade and patenting information.

We distinguish between processing and non-processing firms. A firm is defined as processing firm if during all years with available customs data (2000-2007), over 90% of its total export is through processing export. Among the firms defined as processing firms, 88% are foreign or HMT owned. On the one hand, processing trade is not subject to any import tariffs, and since most processing firms are oriented abroad, they are not likely to be subject to domestic competition either. On the other hand, fall in import tariff could affect firm's choice between ordinary and processing trade

[6]The definition of demand shock in equation (2.3) is consistent with those in Bombardini et al. (2015) and Aghion et al. (2017). An alternative measure is $E_{st}^{\text{alternative}} = \log\left(\sum_{h,c} \frac{X_{hc,2000}}{X_{s,2000}} M_{hct}\right)$. This measure gives similar results.

mode (Brandt and Morrow, 2016), thus indirectly affect firms' incentive to innovate. To avoid such complication, we drop the processing firms from our NBS sample (account for 12% in terms of patents filed during 2003-2007).

We measure innovative activity using invention patents applied by firms. There are three categories of patents in the Chinese system: invention, utility and industrial design. The invention patent is equivalent to the utility patent in the US, and is subject to the agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), which requires, for example, a search in the international patent database to determine the novelty of patents during examination. Each application under the invention category needs to go through two rounds of examination for novelty and non-obviousness, while the other two categories got granted immediately[7]. It takes on average two to four years from application to patent granting. The length of protection for invention patents is twenty years, while that for the other two categories is only ten years. For these reasons, we focus on the invention patent category as innovation outputs.

Patent data is obtained from the State Intellectual Property Office (SIPO) of China. It covers all invention patents applied during 1985-2015. We identify whether a patent belongs to an NBS firm by matching assignee names to the list of NBS firms. Since the NBS data is an unbalanced panel, there are years during which an NBS firm does not have observation in the NBS data (most likely because it is not big enough), but it has observation in the patent data. Table 2.2 shows the total patent count and the matched patent count for year 2007. Of all invention patents, 72% were applied by firms, among which 39% belonged to firms located in China. The NBS firms cover 62% of all patents applied by Chinese corporates. For comparison, Autor et al. (2017) finds that the share of US corporate patents applied by Compustat firms in the manufacturing sector is around 56% in 1999 and around 50% in 2007[8]. Our percentage is higher than theirs because their firm dataset only includes publicly traded firms, whereas ours cover a larger universe of Chinese firms. Out of the patents belonging to NBS firms, 81% fall in our sample of non-processing firms with non-missing data. 5,904 firms in our final sample filed for at least one patent in 2007.

Autor et al. (2017) emphasized the importance of controlling for different industry trends. Table 2.3 shows the evolution of patent distribution and application per firm for 1999, 2003 and 2007. The sample used is the primary sample without Huawei and ZTE. As with the US, we do see different trends among sectors. The share of patent count in chemicals and petroleum declined from 39% in 1999 to 23.8% in 2007; the share of the metal and metal products sector decreased from 12.9% in 1999 to 7.5% in 2007. On the other hand, the computers and electronics sector experienced increase in patenting share from 20.3% in 1999 to 34.6% in 2007. The machinery and equipment sector also experienced an increase, from 9.2% in 1999 to 17.2% in 2007. In terms of application per firm, all sectors experienced an increase, with the most notable increase happening for the computers and electronics sector. In 2003 and 2007, the top three patenting sectors are chemicals

[7]The granted utility or design patents can be revoked if another party sue the patent holder in the court.

[8]In Autor et al. (2017), the percentage of US corporate patents in Compustat is 72%. Out of these patents, the share of manufacturing patents is 77.2% and 70% in 1999 and 2007, respectively. That is how we arrive at 56% and 50%.

and petroleum, computers and electronics, and machinery and equipment. Together they account for over a quarter of manufacturing patent application. Therefore, when controlling for differential sectoral trends, we control for these three sectors separately.

Table 2.2: Patent sample construction

	# application	# patenting firms	patent per firm
SIPO data			
All assignee	233,271		
Firm assignee	167,670	29,212	5.74
Firm located in China	65,621	13,799	4.76
matched to NBS	40,057	7,279	5.50
and non-processing	32,348	5,904	5.48
NBS data		# firms	
All		329,836	
Non-processing		317,467	

Note: Statistics for year 2007.

Table 2.3: Patent distribution across sectors

	(1)	(2)	(3)	(4)	(5)	(6)
	Patent share			No. patents per firm		
Application year	1999	2003	2007	1999	2003	2007
Chem., Petrol., Rubber	39.0%	35.5%	23.8%	1.79	2.30	3.64
Computers, Electronics	20.3%	30.6%	34.6%	2.39	5.35	8.38
Metal, Metal Products	12.9%	6.7%	7.5%	2.20	2.72	4.42
Machinery, Equipment	9.2%	11.8%	17.2%	1.40	1.74	2.91
Food, Tobacco	6.6%	4.4%	4.3%	1.45	1.91	3.85
Clay, Stone, Glass	5.0%	3.2%	2.2%	1.20	1.84	2.15
Transportation	3.3%	4.2%	6.0%	1.56	2.43	4.35
Paper, Print	1.3%	1.6%	1.3%	1.38	1.68	3.16
Textile, Apparel, Leather	1.2%	1.3%	2.2%	1.11	1.44	3.01
Wood, Furniture	1.1%	0.4%	0.5%	1.29	1.38	2.00
Other Manufacturing	0.1%	0.3%	0.2%	1.00	1.39	2.33

Notes: The sample used is the primary sample of non-processing NBS firms, dropping Huawei and ZTE. Industries are ordered by column (1), ranking of patent share in 1999. Columns (1)-(3) show the share out of total patent count for each sector. Columns (4)-(6) shows the average number of patent application per patenting firm.

2.2.3 Tariff reforms and industry competition

Since our hypothesis is that the WTO accession tariff reduction affects innovation incentives through changes in the competition environment, we now examine the industry market structure and its relationship with tariffs and productivity.

Following Aghion et al. (2015), we measure competition through variations of the Lerner Index, which is usually defined as total profit net of financial cost, divided by total revenue or value added. We test three versions of the Lerner Index definition for robustness:

$$
\begin{aligned}
\text{Lerner I}_{st} &= \frac{\sum_{is}\text{Profit}_{ist} - \sum_{is}\text{Finance Fee}_{ist}}{\sum_{is}\text{Value Added}_{ist}} \\
\text{Lerner II}_{st} &= \frac{\sum_{is}\text{Profit}_{ist}}{\sum_{is}\text{Value Added}_{ist}} \\
\text{Lerner III}_{st} &= \frac{\sum_{is}\text{Profit}_{ist} - \sum_{is}\text{Finance Fee}_{ist}}{\sum_{is}\text{Revenue}_{ist}}.
\end{aligned}
$$

The indices are measured at the sector s - year t level, by aggregating up firm level values. We regress them on the output tariffs,

$$\log \ \text{Lerner}_{st} = \beta_0 + \beta^{\tau}\tau_{st}^{\text{output}} + \delta_s + \delta_t + \varepsilon_{st}. \tag{2.7}$$

Table 2.4 shows the results. Taking column (1) as our baseline, a one percentage point drop in tariff is related to 5.6% drop in the Lerner Index, which is 5.6% more competition.

Table 2.4: WTO accession tariff and Lerner Index

	(1) Lerner I_{st}	(2) Lerner II_{st}	(3) Lerner III_{st}
$\tau_{st}^{\text{output}}$	5.624***	2.990***	5.149***
	(1.111)	(0.761)	(1.132)
R^2	0.458	0.465	0.509
Obs	1848	2016	1848

Notes: Time period, 2001-2005. Industry fixed effects and year fixed effects controlled. *** p<0.01, ** p<0.05, * p<0.1.

2.2.4 Tariff reforms and foreign investment

In addition to self-improvement through research and development, one important channel that brings about innovation among firms is technology transfer through FDI. Holmes et al. (2015) and Jiang et al. (2018) both find evidence that there is significant technology transfer to China through the form of Joint Ventures. We agree that FDI is an important source of technology growth for China, the identification of our story would not be threatened as long as changes in FDI patterns do not exactly map changes in industry output tariffs.

In Table 2.5 we look at the correlation between industry output tariff changes and the level and share of foreign equity in the industry. Interestingly, we find that during the post-WTO period, industries that experienced a larger drop in output tariff would see less FDI. This could make sense if exporting to and investing in China are substitutable ways to selling goods in China. When it is easier to export to China, firms would choose to export rather than to establish joint ventures, possibly to avoid transferring technology to China (Holmes et al., 2015). With the results in Table 2.5, our estimates of the effect of import competition on innovation are most likely to be under-estimated if there is any confounding effects coming from the correlated changes in FDI. In Section 2.4, we also put the FDI measures as a control for our baseline regressions.

Table 2.5: FDI and output tariff

	(1)	(2)	(3)	(4)
Dep. var	ln(foreign equity)		ln(foreign+HMT equity)	
$\tau_{s,t-2}^{output}$	3.428***	0.672	3.624***	1.400*
	(0.628)	(1.031)	(0.591)	(0.830)
Year dummy	y	y	y	y
Industry dummy		y		y
R^2	0.033	0.895	0.037	0.925
Obs	2,053	2,051	2,085	2,084

	(5)	(6)	(7)	(8)
Dep. var	foreign equity share		(foreign + HMT) equity share	
$\tau_{s,t-2}^{output}$	0.380***	-0.180	0.749***	0.017
	(0.053)	(0.133)	(0.074)	(0.151)
Year dummy	y	y	y	y
Industry dummy		y		y
R^2	0.044	0.771	0.061	0.849
Obs	2,119	2,119	2,119	2,119

2.3 Estimation Framework

In the empirical analysis, we estimate the effect of tariff reduction on firm's innovation capacity measured by patent application. We assume firms apply for patent at Poisson rate λ_{ist}, so that $\text{Pat}_{ist}|\lambda_{ist} \sim \text{Poisson}(\lambda_{ist})$, and $E(\text{Pat}_{ist}|\lambda_{ist}) = \lambda_{ist}$, where Pat_{ist} denotes patent applied at time t for firm i in sector s. The Poisson arrival rate is affected by firm and industry characteristics, as well as changes in market structure.

We run the following baseline specification:

$$\begin{aligned} \text{Pat}_{ist} \quad = \exp(\quad & \beta_1 \tau_{s,t-2}^{\text{output}} \times \text{Top}_{is,t-2} + \beta_2 \tau_{s,t-2}^{\text{output}} + \beta_3 \text{Top}_{is,t-2} \\ & + \text{EXP CONTROL}_{is,t-2} + \text{IMP CONTROL}_{is,t-2} \\ & + \delta_s + \delta_{St}), \end{aligned} \tag{2.8}$$

where $\tau_{s,t-2}^{\text{output}}$ is the industry import tariff as defined in equation (2.1), $\text{EXP CONTROL}_{is,t-2}$ and

IMP CONTROL$_{is,t-2}$ control for effects brought by shocks from export demand and the imported inputs,

$$\text{EXP CONTROL}_{is,t-2} = \alpha_1 \text{D}^{\text{exporter}}_{is,t-2} + \alpha_2 E^{\text{demand}}_{s,t-2} + \alpha_3 E^{\text{demand}}_{s,t-2} \times \text{D}^{\text{exporter}}_{is,t-2} \quad (2.9)$$

$$\text{IMP CONTROL}_{is,t-2} = \gamma_1 \text{D}^{\text{importer}}_{is,t-2} + \gamma_2 \tau^{\text{input}}_{s,t-2} + \gamma_3 \tau^{\text{input}}_{s,t-2} \times \text{D}^{\text{importer}}_{is,t-2}. \quad (2.10)$$

$E^{\text{demand}}_{s,t-2}$ measures the market size effect brought by export tariff changes, as defined by equation (2.3). $\tau^{\text{input}}_{is,t-2}$ is the two period lagged input tariff measure which is defined by equation (2.2). We use two-period lagged tariff shocks to take into account that it takes a while for innovative ideas to be turned into patents. Further, we control for industry fixed effect δ_s at the four-digit level, and sector-year fixed effect δ_{St} to take into account the different sectoral trends. S is a categorical variable on four sectors: chemicals and petroleum, computers and electronics, machinery and equipment sector, and others .

Following Bustos (2011), we divide firms into four groups for any industry-year cell, according to two period lagged TFP quartiles. The TFP estimation procedure follows Ackerberg et al. (2015), De Loecker and Warzynski (2012) and Brandt et al. (2017), which is discussed in detail in Appendix A.1. Top$_{is,t-2}$ is then defined to be a dummy which equals to 1 if the firm is above the 75th percentile in terms of productivity among firms in industry s two periods before. Coefficient $-\beta_1$ measures the differential percentage change in patenting rate for the top firms after the industry output tariff decreases by 0.01.

For firms with product composition information, we could also construct competition, export, and imported input shocks at the firm level, which are defined by equations (2.4)-(2.6). We run the following firm level specification:

$$\begin{aligned} \lambda_{ist} = \exp(& \beta_1 \tau^{\text{output}}_{is,t-2} \times \text{Top}_{is,t-2} + \beta_2 \tau^{\text{output}}_{is,t-2} + \beta_3 \text{Top}_{is,t-2} \\ & + \alpha E^{\text{demand}}_{is,t-2} + \gamma \tau^{\text{import}}_{is,t-2} + \delta_s + \delta_{St} + \delta_i). \end{aligned} \quad (2.11)$$

Since import tariff declined most during 2001-2005, we restrain our time period of analysis to 2001-2005 for the tariff shocks. And since we assume it takes one to two years for patents to come out, our patent variables cover years 2003-2007.

Table 2.6 shows the summary statistics of the firm level variables, for the top and non-top firms in 2003 and 2007. Table 2.7 shows the mean and correlation among the trade shocks. The output tariff is negatively correlated with the export demand, and positively correlated with the input tariff.

Table 2.6: Summary statistics, firm level

	2003				2007			
	Top 25%		Non-top		Top 25%		Non-top	
Firm performance and inputs								
Patent per firm	0.06	(1.72)	0.03	(0.48)	0.19	(6.44)	0.07	(1.43)
Patent dummy	0.02	(0.13)	0.01	(0.10)	0.03	(0.17)	0.02	(0.14)
log TFP	1.10	(0.50)	0.87	(0.47)	1.28	(0.54)	1.06	(0.46)
log R&D	0.90	(2.19)	0.73	(1.89)	0.82	(2.28)	0.66	(1.98)
log Capital	8.55	(1.80)	8.79	(1.51)	8.59	(1.76)	8.66	(1.48)
log Employment	4.86	(1.24)	4.92	(1.19)	4.75	(1.17)	4.72	(1.06)
Firm level trade shocks								
firm $\tau_{is,t-2}^{output}$	0.14	(0.07)	0.15	(0.07)	0.10	(0.06)	0.10	(0.06)
Exporter dummy	0.42	(0.49)	0.41	(0.49)	0.41	(0.49)	0.41	(0.49)
firm $E_{is,t-2}^{demand}$	9.94	(2.16)	9.73	(2.24)	10.21	(2.19)	10.12	(2.23)
Importer dummy	0.24	(0.43)	0.23	(0.42)	0.25	(0.43)	0.24	(0.43)
firm $\tau_{is,t-2}^{input}$	0.10	(0.05)	0.11	(0.05)	0.07	(0.04)	0.08	(0.04)
Ownership								
State dummy	0.13	(0.34)	0.20	(0.40)	0.06	(0.23)	0.06	(0.24)
Foreign dummy	0.10	(0.30)	0.07	(0.26)	0.12	(0.33)	0.09	(0.29)

Note: Standard deviations are shown in parenthesis.

Table 2.7: Summary statistics, industry level

	2001		2005		2001-2005		Correlation in 2005		
τ_{st}^{output}	0.134	(0.067)	0.096	(0.049)	0.109	(0.057)	1.000		
E_{st}^{demand}	12.273	(1.642)	12.247	(1.892)	12.237	(1.764)	-0.290*	1.000	
τ_{st}^{input}	0.053	(0.020)	0.038	(0.012)	0.043	(0.016)	0.257*	0.209*	1.000

Note: Standard deviations are shown in parenthesis. * p<0.01.

2.4 Empirical Results

2.4.1 Baseline Estimates

Table 2.8 shows the regression results for specification (2.8). From left to right, we gradually add in export demand and input tariff controls. All columns control for the four-digit industry fixed effects and sector-year effects. For firms below the 75th percentile of TFP, the effect of import competition is almost zero, with big standard errors. Relative to them, the top firms are highly responsive to import tariff drops. Taking column (4) as our baseline result, after a one percentage point drop in import tariff, the top firms increase their patent application effort by 3.6 percentage points more, relative to the less productive firms. During the period of 2003-2007, the annual growth rate of the average patenting rate among the top firms is 37.5 percentage points. Thus a one percentage point drop in import tariff roughly contributes to 10 percent of the growth in top firm innovation.

Being an exporter increases average patent application per firm by 0.9. Increase in export demand in general discourages non-exporters to innovate, while it tends to encourage exporters. This result is consistent with what was found for French firms in Aghion et al. (2017).

For the effect of accessing imported inputs, importers on average file for one more patent than non-importers. While change in input tariff has no effect for non-importers, the encouragement effect of innovation for importers is quite big. The average annual decrease in input tariffs is 0.00375, which would predict an increase in patenting rate of 2.7% for the importers.

Table 2.8: Output tariff and patenting, industry measure

	(1)	(2)	(3)	(4)
Dep. var: Patent application counts				
Output competition				
$\tau^{\text{output}}_{s,t-2} \times \text{Top}_{is,t-2}$	-3.525**	-3.675**	-3.418**	-3.577**
	(1.412)	(1.428)	(1.474)	(1.466)
$\tau^{\text{output}}_{s,t-2}$	-0.506	-0.161	0.966	1.275
	(1.643)	(1.655)	(1.896)	(1.883)
$\text{Top}_{is,t-2}$	1.210***	1.234***	1.177***	1.209***
	(0.142)	(0.143)	(0.147)	(0.146)
Export control				
$D^{\text{exporter}}_{ist-2}$		1.278***		0.909***
		(0.333)		(0.329)
$E^{\text{demand}}_{s,t-2}$		0.063***		0.066***
		(0.021)		(0.021)
$E^{\text{demand}}_{s,t-2} \times D^{\text{exporter}}_{ist-2}$		0.026		0.023
		(0.025)		(0.025)
Import control				
$D^{\text{importer}}_{ist-2}$			1.799***	0.984***
			(0.188)	(0.168)
$\tau^{\text{input}}_{is,t-2}$			-0.336	-4.936
			(11.268)	(10.972)
$\tau^{\text{input}}_{is,t-2} \times D^{\text{importer}}_{ist-2}$			-11.314***	-7.385**
			(3.848)	(3.639)
obs	800,292	800,292	800,292	800,292

Notes: The Top dummy equals to 1 if the firm is above 75th percentile in industry s at time $t-2$. All columns control for four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. The four sectors are: chemicals and petroleum, computers and electronics, machinery and equipment sector, and others. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1.

In Table A.3 in the appendix, we show the effect of output tariff changes on the four productivity quartiles separately, by interacting the output tariff with the lagged TFP quartile dummies, instead of only the top dummy. Consistent with what we found in the baseline specification in Table 2.8, firms in the top quartile innovate more when there is a larger drop in output tariff.

As discussed in Section 2.2.4, foreign investment is another potential channel that affect firms' innovation capacity and could be confounded with the import competition channel. We add the industry level FDI as a further control to our baseline specification in Table A.4 in the appendix. The magnitude and significance of our estimates stay stable and robust.

While in Table 2.8, and throughout the main text of this chapter, we are showing the reduced form relationship between firm patenting and import tariff changes, in Appendix A.2 we test the underlying mechanism that a decrease in import tariffs first causes increase in import competition, leading to innovation reaction among firms. More specifically, in Table A.5, we regress patent counts on industry import volume changes that were induced by tariff changes. Instead of an instrumental variable procedure, we use the control function approach that is widely used in the literature when dealing with Poisson count data regressions (Aghion et al., 2009; Wooldridge, 2010; Blundell and Powell, 2003). The results show that a drop in industry output tariff τ^{output} indeed causes increase in imports of the competing goods in that industry, which causes an increase in patent application among top firms.

Table 2.9 shows the results for firm level specification (2.11). The coefficients on the heterogeneous effect of output competition remain stable across columns, and the magnitude is close to the industry specification in Table 2.8. The export demand elasticity increases relative to the industry specification. While the imported input effects, on the other hand, becomes not significant in the firm specification. In Table A.8 in the appendix, we show results for the OLS specification. The coefficients are comparable.

Table 2.9: Output tariff and patenting, firm measure

	(1)	(2)	(3)	(4)
Dep. var: Patent application counts				
Output competition				
firm $\tau^{\text{output}}_{is,t-2} \times \text{Top}_{is,t-2}$	-4.791**	-4.874**	-5.285**	-5.295**
	(2.423)	(2.335)	(2.659)	(2.571)
firm $\tau^{\text{output}}_{is,t-2}$	-1.368	-0.586	-0.617	-0.086
	(1.284)	(1.322)	(1.453)	(1.455)
$\text{Top}_{is,t-2}$	1.372***	1.260***	1.291***	1.266***
	(0.232)	(0.226)	(0.262)	(0.252)
Export control				
firm $E^{\text{demand}}_{is,t-2}$		0.102***		0.082***
		(0.021)		(0.024)
Import control				
firm $\tau^{\text{input}}_{is,t-2}$			1.208	1.381
			(2.290)	(2.322)
obs	138,640	132,945	72,500	70,246

Notes: Poisson specification. All columns control for four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. See Table A.8 for an OLS specification. The Top dummy equals to 1 if the firm is above 75th percentile in industry s at time $t-2$. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1.

In all the specifications so far, we use patent application counts as the measure of innovation outcome. The granting rate for the patent applications is around 60% during the sample period. To better control for the quality of patent application, in Table A.6, we run the same specification as in column (4) of Table 2.8 using patent grants (column 1) and citation weighted patent counts (column 2) as our outcome variables. The effect of a one percentage point drop in output tariff remains the same as in our benchmark. One may also be concerned that firms file for multiple patents under the same technology to better protect itself in case of law suits. Therefore, giving each application the same weight would possibly overstate the effort to innovate. Furthermore, if firms become more strategic due to competition, our estimate would be upward biased. We check for the specification with patent dummy, instead of patent counts, as our dependent variable. The coefficient magnitude is not readily comparable, but the direction and significance of the effect remains. In column (4) and (5) of Table A.6, we use alternative specifications that have been used in the literature, other than Poisson, and still the direction and significance of the estimated effect remains.

In Table 2.10 column (1)-(2), we show the long term regression by running the industry specification on years 2003 and 2007 only. Both columns controlled for ownership, region, year and industry dummies. Column (2) also includes the export and import controls. The long term effect of a one percentage point decrease in output tariff encourages top firms to increase patent rate by

4.6 percentage points.

Table 2.10: Long term effects and falsification test

Dep. var	I. 2003-2007 $\text{Pat}_{it,(t=2003,2007)}$ (1)	(2)	(3)	(4)	II. 1998-2001 (pre-exposure) $\text{Pat}_{it,(t=1998,2001)}$ (5)	(6)
Past output competition						
$\tau^{\text{output}}_{s,t-2} \times \text{Top}_{is,t-2}$	-4.587***	-4.685***	-3.005*	-2.982*		
	(1.534)	(1.576)	(1.546)	(1.622)		
$\tau^{\text{output}}_{s,t-2}$	-0.029	0.646	-0.333	0.733		
	(2.523)	(2.623)	(2.323)	(2.709)		
$\text{Top}_{is,t-2}$	1.335***	1.338***	1.225***	1.197***		
	(0.152)	(0.156)	(0.143)	(0.145)		
Pre-exposure trends						
$\text{Pat}_{i,t-6}$			0.128***	0.117***		
			(0.015)	(0.013)		
Future output competition						
$\tau^{\text{output}}_{s,t+4} \times \text{Top}_{is,t+4}$					-1.768	-1.815
					(2.189)	(2.193)
$\tau^{\text{output}}_{s,t+4}$					0.099	0.920
					(2.724)	(3.400)
$\text{Top}_{is,t+4}$					0.897***	0.859***
					(0.193)	(0.190)
obs	337,029	337,029	141,190	141,190	119,146	119,146

Notes: The specifications are Poisson with two years stacked. All columns control for ownership and region dummies, four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. The even columns include the export and import controls in addition. Standard errors are clustered at the industry-year-top level. *** p<0.01, ** p<0.05, * p<0.1

The time period we look at is one where the patent rate in China picked up rapidly. One concern is that the differential patenting behavior between top and other firms and across industries was caused by an unobservable factor that also determined the tariff measures. Therefore, in columns (3)-(4), we add to columns (1)-(2) firms' patent applications in the pre-exposure years as additional control. Indeed, patent application is a rather persistent feature for firms. The coefficient in front of past patent is highly significant and positive. Since patenting was a rather high-tech activity, we should expect that firms that patent before the WTO accession would continue patenting. The point estimate of the interaction term becomes smaller, with slightly higher standard errors, making the estimate less precise. Second, we run a falsification test in columns (5)-(6) by regressing the pre-exposure patent application in 1998 and 2001, on the future tariff rates in 2001 and 2005. The interaction term is weakly negative, and not significant at the 10 percent significance level. Therefore, we do not find evidence that the pre-exposure patenting behavior is related to the WTO accession tariffs.

2.4.2 Technology deepening v.s. technology scope

Next, we further investigate the dimensions of innovation that are induced by import competition. Specifically, we decompose the total patent count into patents filed in the core technology of a firm, versus the total number of technology classes the firm file patents into. Technology class is defined according to the six-digit International Patent Classification (IPC)[9]. A technology class is defined as the core technology if a firm has accumulated the most patent applications in that class up to the previous year. The technology scope is the sum of the number of classes a firm files patent in a specific year.

Table 2.11 shows the estimation result for the industry specification applying on core technology and technology scope. Column (1) is repeating column (4) in Table 2.8 as benchmark. The results suggest that the top firms react to increase in import competition by both increasing innovation in the core technology as well as broadening its technology space. The point estimate for the effect on patent scope (column 2) is smaller than the overall effect (column 1), whereas the point estimate for the core patent (column 3) is larger than the overall effect. The result remains very similar when we only look at firms that have applied for patents before (columns 4-6). Column (7) shows the effect on the ratio of scope to core. The second row show that on average, firms react more by increasing their patent in core technology, which is consistent with previous columns. There isn't a differential effect for the top firms in terms of the relative magnitude of core and scope innovation.

[9]There are 4944 six digit IPC in 2007. For example, in 2007, Huawei filed patents in 144 technology classes. According to our definition, its core patent class was H04L12, "Data switching networks". Other technology classes that it filed patent in are H04L29, "Arrangements, apparatus, circuits or systems", and H04L1, "Arrangements for detecting or preventing errors in the information received", etc.

Table 2.11: Technology core vs. scope

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample		All			Patented before		$\frac{\text{scope}}{\text{core}}$
Dep. var	application	scope	core	application	scope	core	
Output competition							
$\tau^{\text{output}}_{s,t-2} \times \text{Top}_{is,t-2}$	-3.577**	-2.397**	-3.968**	-3.659***	-2.323**	-4.833**	-0.472
	(1.466)	(0.974)	(1.832)	(1.591)	(1.092)	(2.063)	(0.821)
$\tau^{\text{output}}_{s,t-2}$	1.275	1.188	1.334	1.247	2.281	0.705	3.654**
	(1.883)	(1.337)	(1.564)	(2.333)	(1.650)	(1.917)	(1.454)
$\text{Top}_{is,t-2}$	1.209***	0.933***	1.091***	1.020***	0.768***	0.953***	0.205***
	(0.146)	(0.090)	(0.192)	(0.152)	(0.088)	(0.217)	(0.071)
Export control							
$D^{\text{exporter}}_{ist-2}$	0.909***	1.582***	0.401	-0.022	0.517*	-1.163**	0.287
	(0.329)	(0.252)	(0.409)	(0.455)	(0.308)	(0.502)	(0.196)
$E^{\text{demand}}_{s,t-2}$	0.066***	0.079***	0.082***	0.059*	0.052**	0.056*	0.018
	(0.021)	(0.019)	(0.022)	(0.030)	(0.022)	(0.033)	(0.018)
$E^{\text{demand}}_{s,t-2} \times D^{\text{exporter}}_{ist-2}$	0.023	-0.040**	0.028	0.068**	0.012	0.124***	-0.003
	(0.025)	(0.020)	(0.033)	(0.034)	(0.023)	(0.039)	(0.016)
Import control							
$D^{\text{importer}}_{ist-2}$	0.984***	0.968***	0.631***	0.469**	0.174	-0.043	0.172
	(0.168)	(0.163)	(0.202)	(0.205)	(0.176)	(0.195)	(0.172)
$\tau^{\text{input}}_{is,t-2}$	-4.936	-9.983	-6.497	-2.990	-13.008	-4.033	-5.160
	(10.972)	(6.907)	(7.036)	(12.473)	(8.584)	(9.981)	(8.317)
$\tau^{\text{input}}_{is,t-2} \times D^{\text{importer}}_{ist-2}$	-7.385**	-4.063	1.744	-6.094	-0.314	7.254*	-4.132
	(3.639)	(3.685)	(3.905)	(5.047)	(4.247)	(4.330)	(4.427)
obs	800,292	800,005	800,005	23,931	23,917	23,917	12,721

Notes: All columns control for ownership, region, four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1

2.4.3 Effect on firm scale and productivity

In this section, we look at the effect of import liberalization on other firm outcome variables.

First, we are interested in whether surviving firms get bigger. Table 2.12 shows the effect of trade shocks on domestic sales in columns (1)-(3) and domestic market share in columns (4)-(6). From columns (4) -(6), the domestic market share decreases for all firms following import competition, which is a mechanical result to be expected. From column (3), there is a weak increase in the domestic output for firms surviving the competition, 0.44 percent increase, after a one percentage point drop in output tariff.

Table 2.12: Effects on domestic output

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var		log domestic output			Domestic market share	
Output competition						
$\tau^{\text{output}}_{s,t-2} \times \text{Top}_{is,t-2}$		-0.491**	-0.435*		-0.031	-0.033
		(0.243)	(0.238)		(0.033)	(0.036)
$\tau^{\text{output}}_{s,t-2}$	-0.177	-0.038	-0.427	0.064**	0.072**	0.036
	(0.397)	(0.398)	(0.432)	(0.027)	(0.032)	(0.023)
$\text{Top}_{is,t-2}$		0.657***	0.650***		0.010	0.010
		(0.033)	(0.032)		(0.009)	(0.009)
Export control						
$D^{\text{exporter}}_{ist-2}$			0.396***			0.102
			(0.118)			(0.120)
$E^{\text{demand}}_{s,t-2}$			0.008*			0.004
			(0.005)			(0.003)
$E^{\text{demand}}_{s,t-2} \times D^{\text{exporter}}_{ist-2}$			-0.037***			-0.009
			(0.010)			(0.011)
Import control						
$D^{\text{importer}}_{ist-2}$			0.744***			0.025
			(0.082)			(0.023)
$\tau^{\text{input}}_{is,t-2}$			4.853***			0.694
			(1.633)			(0.496)
$\tau^{\text{input}}_{is,t-2} \times D^{\text{importer}}_{ist-2}$			-7.499***			-1.303
			(1.745)			(0.982)
obs	741,978	741,978	741,978	808,123	808,123	808,123

Notes: All columns control for four-digit CIC industry fixed effects, aggregate sector by year fixed effects. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1

In Table 2.13 we look at the effect on productivity, R&D, capital and employment inputs. The effect on productivity is quite pronounced. After a one percentage point drop in output tariff, the top firms see increase in productivity by 0.17 percent. This is consistent with the estimate of 0.19 percent in Brandt et al. (2017). In column (2), we estimate the effect of import competition on the R&D input. Consistent with the result for the patent application, the top firms react more to import competition and put more effort into research and development in the face of more liberalized import market. Column (3) and (4) shows that the elasticity of capital and labor on output tariff is 1.12 and 0.36, respectively.

Table 2.13: Effect on TFP, R&D, capital and labor

Dep var	(1) TFP *OLS*	(2) ln (R&D) *OLS*	(3) ln (capital) *OLS*	(4) ln (labor) *OLS*
Output competition				
$\tau^{\text{output}}_{s,t-2} \times \text{Top}_{is,t-2}$	-0.167***	-1.333***	-1.122***	-0.365**
	(0.036)	(0.226)	(0.264)	(0.152)
$\tau^{output}_{s,t-2}$	-0.447*	0.276	0.522	-0.462*
	(0.236)	(0.374)	(0.364)	(0.279)
$\text{Top}_{is,t-2}$	0.221***	0.302***	-0.003	0.051**
	(0.005)	(0.027)	(0.036)	(0.021)
Export control				
$D^{\text{exporter}}_{ist-2}$	0.043***	0.140	0.339***	0.242***
	(0.015)	(0.158)	(0.075)	(0.046)
$E^{\text{demand}}_{s,t-2}$	0.002	-0.000	-0.006	-0.005*
	(0.002)	(0.006)	(0.004)	(0.003)
$E^{\text{demand}}_{s,t-2} \times D^{\text{exporter}}_{ist-2}$	-0.003***	0.022*	0.009	0.019***
	(0.001)	(0.013)	(0.006)	(0.004)
Import control				
$\text{D}^{\text{importer}}_{ist-2}$	-0.007*	0.639***	0.957***	0.449***
	(0.004)	(0.101)	(0.048)	(0.024)
$\tau^{input}_{is,t-2}$	-0.822	-0.121	8.180***	4.130***
	(0.808)	(2.096)	(1.270)	(1.481)
$\tau^{input}_{is,t-2} \times \text{D}^{\text{importer}}_{ist-2}$	0.180**	-4.153	-4.012***	-1.105**
	(0.080)	(2.537)	(0.960)	(0.473)
R^2	0.567	0.137	0.228	0.196
obs	802598	563144	798414	802598

Notes: The Top dummy equals to 1 if the firm is above 75th percentile in industry s at time $t-2$. All columns control for four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. The four sectors are: chemicals and petroleum, computers and electronics, machinery and equipment sector, and others. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1

2.5 Conclusion

The China Miracle has been a manufacturing success. But after over forty years of rapid growth with cheap labor, imitation, and institutional reforms, China's manufacturing sector has arrived at a crossroad where further growth depends much on indigenous innovation. In this chapter, we study the impact of change in competition environment brought about by foreign imports on Chinese firm's innovation capacity, measured by patent application. Using a newly combined data set that covers the universe of medium to large manufacturing firms, and more than 60% of corporate innovators, we find that the increase in import competition following China's accession to the WTO during 2001-2005 induced more productive firms to innovate more.

Our finding adds to the debate on the effect of international competition on innovation. For a developing country like China, opening to international competition served as a stimulating mechanism for the top firms to invest in research to improve products and processes. In the mean time, a less productive firm may find it not as attractive to innovate. Whether the aggregate effect is positive or negative depends on the extent of technology spillover and other effects we do not consider in this work. We believe this is a fruitful future research path to pursue.

Chapter 3

Import Competition and Innovation: A Theory

3.1 Introduction

In Chapter 2, we showed empirical evidence that the effect of import competition on firm's incentives to innovate is heterogeneous in the case of China. Across the 4-digit industries, a 1 percentage point drop in import tariff induces a 3.6 percent increase in patenting rate for the top 25% most productive firms. We also showed that the effect of competition on innovation of initially less productive firms is not significantly different from zero.

In this Chapter, we propose a model of international trade with endogenous innovation to illustrate the mechanisms behind the heterogeneous response of R&D and innovation to trade liberalization. More specifically, we build on the trade model with firm heterogeneity as proposed in Melitz and Ottaviano (2008) and introduce two modifications. The first one is the ability to innovate, subject to a convex cost, and the second one is neck-and-neck competition between domestic and foreign firms within each variety of differentiated goods. We show that import competition brings about a negative "rent-destruction" effect and a positive "escape-competition" effect. The net effect of import competition depends on the dominating force. We study the welfare implications in Chapter 4.

The escape-competition effect induced by the neck-and-neck competition is a common feature to several papers by Aghion et al. (2001, 2005, 2009)[10]. However, these models usually have zero cross-price elasticity, which reduces the negative effect of rent destruction introduced by more competition. Such rent-destruction effect is important in any trade models with monopolistic competition. In addition, in these papers, the change in the competition environment is usually governed by exogenous parameters that are difficult to map to a variable empirically. Our model contributes to studying the escape-competition and the rent-destruction effects in a unified model, where the competition environment can be easily summarized by an empirical variable — the import tariffs.

Our model also contributes to introducing within sector oligopoly competition into a typical trade model. Looking at narrow product niches, we usually see several big players strategically competing, instead of hundreds of ignorable small players. Therefore, by introducing the neck-and-neck competition, we not only could introduce the missing escape-competition effect, but also could

[10] See also the IO literature surveyed in Gilbert (2006).

push the model closer to reality.

Our theory is also related to the literature on trade and innovation in a heterogeneous firm framework, although it is worth noting that the vast majority of these papers feature increased market size as the incentive for further innovation. A key contribution that features this mechanism is the one by Atkeson and Burstein (2010). Their main result is that as variable trade costs decline, any increase in average productivity due to additional process innovation is compensated (in special cases exactly) by reduced entry and therefore lower product variety. The mechanism in that paper is due to the fact that a shock to the export market initially increases expected profits. But to satisfy the free entry condition, expected profits have to decline. This happens because as firms become more productive by investing more, the average firm is more productive and expands, demanding more labor. This puts pressure on wages, thus further decreasing average profits for the other firms that enter the market. Perla et al. (2015) also generate dynamic gains from trade in a model where firms invest to learn from existing firms in the market. Due to Melitz-type selection, trade opening improves the pool of firms that other firms can learn from. In their setting welfare rises because growth increases due to costly imitation, but, similarly to Atkeson and Burstein (2010), their welfare gains are reduced by decreased entry. Other papers in this literature are Rivera-Batiz and Romer (1991) and Hsieh et al. (2018), Grossman and Helpman (1991).

Four contemporary papers that are related to the mechanism described in this paper are Fieler and Harrison (2018), Aghion et al. (2017), Akcigit et al. (2017) and Lim et al. (2017). Fieler and Harrison (2018) and Lim et al. (2017) share some common features that are different from our model. They work with a constant elasticity of substitution utility function which features different nests. As firms innovate they can escape to another nest, where they face lower competition. The difference between these two papers is the source of increased competition. In Fieler and Harrison (2018), the rise in competition comes form foreign firms entering the domestic (Chinese in their case) market, whereas in Lim et al. (2017) the increase in competition is a consequence of the rising domestic entry due to export opportunities. The model is closest to Aghion et al. (2017) and Akcigit et al. (2017). Aghion et al. (2017) also builds on the framework of Melitz and Ottaviano (2008), but focuses on the effect of exporting on innovation for French firms. They find a negative effect of competition coming from the price index effect which would discourage innovation among firms away from the technology frontier. In Akcigit et al. (2017), importing happens in sectors where the home firm is lagging behind, and therefore, only firms at the middle-lower part of the productivity distribution react to import competition. In contrast, in our model, foreign competition can be present along the whole productivity distribution, and we let the data tell us which firms are affected the most.

The remainder of this chapter is organized as follows: Section 3.2 lays out a baseline model with firm heterogeneity. Section 3.3 analyses the escape-competition and the rent-reduction channels through the lens of the model. Section 3.4 shows simulation results of the model. And Section 3.5 concludes.

3.2 Model setup

In this section, we present a model of import competition in which firms from two countries, Home (H) and Foreign (F), compete in the domestic (Home) market. There is a finite number of varieties and firms compete in a Bertrand fashion when they produce the same variety. Moreover, after entering the market, firms have a chance to further invest in cost-reducing innovation. The model asks for a fixed number of varieties of goods. Fixing the number of varieties enables "neck-and-neck" competition (Aghion et al., 2005, 2009; Akcigit et al., 2017). If the number of potential varieties is unlimited, then two firms will never enter the market in the same variety. For a given productivity draw it is weakly more profitable to enter a variety not previously produced. As will be shown in section 3.3, the presence of the "neck-and-neck" state will drive the positive reaction of some firm's innovation to the increased foreign competition.

3.2.1 Consumer preferences

Consumers enjoy utility from consumption of a homogeneous good, denoted by q_0^c, and a mass N_e of potential varieties of the differentiated good, each denoted by q_i^c. The utility is a quadratic aggregator of the goods, as in Ottaviano et al. (2002) and Melitz and Ottaviano (2008):

$$U = q_0^c + \alpha \sum_{i=1}^{N} q_i^c di - \frac{1}{2}\gamma \sum_{i=1}^{N} (q_i^c)^2 di - \frac{1}{2}\eta \left[\sum_{i=1}^{N} q_i^c di\right]^2 . \tag{3.1}$$

where $N \leq N_e$ is the number of differentiated goods being produced. Parameter α measures the relative importance of the differentiated good over the numeraire. The parameter η also governs the cross-price elasticity of demand. The parameter γ governs the own-elasticity of demand among the differentiated varieties.

Consumers choose quantities q_0^c, q_i^c to maximize utility subject to the budget constraint $\int_{i\in\Omega} p_i q_i^c di = E^c$, where E^c is the total income of an individual, given by the wage w and a share of profits when positive.

We make the common simplifying assumptions that the homogeneous good is the numeraire, freely traded, produced and consumed in positive quantities in each country and that its production is one-to-one with labor, implying that the wage w is equal to 1. The inverse demand for each differentiated good i is given by the following equation:

$$p_i = \alpha - \gamma q_i^c - \eta Q^c,$$

and the total demand for each variety is

$$q_i = L q_i^c = \frac{\alpha L}{\eta N + \gamma} - \frac{L}{\gamma} p_i + \frac{\eta N L}{\gamma (\eta N + \gamma)} \bar{p} \quad \forall i = 1, ..., N \tag{3.2}$$

Variable N is the number of active varieties, i.e. those for which the following inequality is satisfied:

$$p_i \leq \alpha \equiv p_i^{\max}$$

where $p_i^{\max}$ represents the choke price at which demand for a variety is driven to 0.

Each consumer earns a equal share of the profits from the firms in the economy. The indirect utility can thus be written in terms of firm profit and the mean and variance of prices.

$$U = 1 + \Pi + \frac{1}{2} \frac{1}{\eta + \frac{\gamma}{N}} (\alpha - \overline{p})^2 + \frac{1}{2} \frac{N}{\gamma} \sigma_p^2 \tag{3.3}$$

where $\sigma_p^2 \equiv \frac{1}{N} \int_i (p_i - \bar{p})^2 \, di$ is the variance of price. Π denotes the total profits earned by the firm. Utility increases when firms are more profitable, when average price drops, or when there is a higher dispersion in price.

3.2.2 Production and Market Structure

Within each variety, there are two potential producers, one in Home and one in Foreign and they engage in Bertrand competition. Each of them draws their initial productivity from a country-specific cost distribution $G_n(c)$, where $n \in \{H, F\}$, and c denotes the production cost. We assume Bertrand competition within each variety, so that only the firm with the lowest cost draw enters the market and becomes the incumbent for that variety. The incumbent can then further choose its innovation effort and decrease production costs by a fixed step size δ with some probability proportional to its innovation effort.

Foreign firms can sell in the domestic market subject to an iceberg transport cost $\tau > 1$. Since we are interested in the effect of unilaterally decreasing the transport cost τ, we make the simplifying assumption that Home exports are in terms of the homogeneous good only. This is inconsequential in this partial equilibrium setting and thus allows us to more clearly isolate the effect of increased import competition from other potential innovation incentives coming from the export market.

Similar to Akcigit et al. (2017), we assume that firms incur a small cost $\varepsilon \to 0$ to set the price. Therefore, only the firm with the highest productivity for each variety will set the price and produce. In this way, we sacrifice the more realistic limit pricing setting (Bernard et al., 2003) for analytical simplicity. In reality, there is no barriers to setting a lower price; therefore, the best firm could not charge a price that is higher than the cost of the second-best firm. To make the model easier to illustrate and simulate, we keep the simplifying assumption that the best firm could still charge the monopoly price. While it doesn't hurt our analysis in this chapter, this is not an innocuous assumption for welfare analysis, as will become clear in Chapter 4. We keep the simplification for now and would return to limit pricing in Chapter 4.

In this setup the presence of foreign firms generates an increase in the probability that the domestic producer will exit the market. This is the key force that drives additional innovation. If innovation can make the domestic producer more productive than foreign firms, then more foreign competition will induce domestic firms to innovate more.[11]

Labor is the only factor of production. Given marginal cost c_i for the firm operating in variety i, the firm's static problem is to maximize profits $\pi(\cdot)$:

$$\max_{p_i} \pi(c_i) = (p_i - c_i)\, q_i$$

subject to the demand function equation (3.2).

Although there is a potential number of N_e varieties, some may not be produced in equilibrium if all firms draw a cost of production that is too high. We denote by c_D the highest cost a firm can draw that will still allow it to make non-negative profits. Then, the cutoff cost, price, quantity, profit and revenue expression for both the Home production of the differentiated good, as well as the imported Foreign goods are:

$$c_D = \frac{\alpha\gamma}{\eta N + \gamma} + \frac{\eta N}{\eta N + \gamma}\overline{p}; \tag{3.4}$$

$$p(c) = \frac{1}{2}(c_D + c); \qquad p^*(c^*) = \frac{1}{2}(c_D + c^*\tau) \tag{3.5}$$

$$q(c) = \frac{L}{2\gamma}(c_D - c); \qquad q^*(c^*) = \frac{L}{2\gamma}(c_D - c^*\tau) \tag{3.6}$$

$$\pi(c) = \frac{L}{4\gamma}(c_D - c)^2; \qquad \pi^*(c^*) = \frac{L}{4\gamma}(c_D - c^*\tau)^2 \tag{3.7}$$

$$r(c) = \frac{L}{4\gamma}\left((c_D)^2 - c^2\right); \qquad r^*(c^*) = \frac{L}{4\gamma}\left((c_D)^2 - (c^*\tau)^2\right) \tag{3.8}$$

where the foreign variables are indicated by a star.

3.2.3 Innovation decision

After entering the market, domestic incumbent firms can invest in R&D to lower their costs to δc where $0 < \delta < 1$.[12] A common way to model this problem is to have the firm choosing the probability I of a successful innovation by paying a cost that is increasing and convex in this probability. This cost takes the form of $\frac{I(c)^2}{2\phi}$. The innovation decision will have a key component given by the probability of survival. A domestic firm with cost c will survive with probability $1 - G_F\left(\frac{c}{\tau}\right)$ if it fails to innovate and with probability $1 - G_F\left(\frac{\delta c}{\tau}\right)$ if it succeeds, where $G_F(.)$ is the CDF of foreign cost c^*.

Therefore the optimization problem for the domestic firm starting with production cost c consists

[11] There is also an additional effect of foreign firms operating through the average price, but as we will discuss later, this is not sufficient to generate a competition-induced increase in innovation. In Chapter 4, introducing limit pricing will mean foreign firms have the additional effect of lowering the price the domestic firm can charge.

[12] We abstract from the innovation response of domestic firms.

in maximizing the function $V(c)$:

$$\max_{I(c)} V(c) = I(c)\left(1 - G_F\left(\frac{\delta c}{\tau}\right)\right)\pi(\delta c) + (1 - I(c))\left(1 - G_F\left(\frac{c}{\tau}\right)\right)\pi(c) - \frac{1}{2\phi} I(c)^2 . \quad (3.9)$$

The solution to (3.9) gives the following innovation policy function:

$$I(c) = \begin{cases} \phi\left(\tilde{\pi}(\delta c) - \tilde{\pi}(c)\right) & \text{if } c \leq c_D \\ \phi\tilde{\pi}(\delta c) & \text{if } c_D < c \leq \frac{c_D}{\delta} \end{cases} \quad (3.10)$$

where

$$\tilde{\pi}(c) = \frac{L}{4\gamma}\left(1 - G_F\left(\frac{c}{\tau}\right)\right)(c_D - c)^2 .$$

and the function $\tilde{\pi}(\delta c)$ represents expected profits when the firm succeeds in reducing its cost to δc.

3.2.4 Firm entry and exit

We assume that in each of the N_e differentiated good varieties, there is only one domestic firm that takes a production cost draw from the domestic talent (entrepreneur skill) distribution $G_H(c)$. They can choose not to produce if their cost realization is too low, thus exiting the market. This setup corresponds to the short-run equilibrium in Melitz and Ottaviano (2008), in which there is no free entry and incumbent firms can earn positive profits.

An alternative setup is to allow for a pool of entrants who can take the productivity draw by incurring a fixed entry cost. This could generate a domestic incumbent distribution that is related to the number of firms taking the productivity draw. Although potentially consequential, we leave the analysis of this free-entry condition to future research.

3.2.5 Endogenous technology distribution and the price index

We now derive the endogenous distribution of firm costs, taking into account the innovation decisions. This also allows us to derive the price index. Let $F(c)$ denote the ex-post cumulative distribution after the realization of incumbent's innovation investments. The relation between $G_H(c)$ and $F(c)$ is given by the following equation:

$$F(c) = G_H(c) + \int_c^{\frac{c}{\delta}} I(c')\, dG_H(c') \quad (3.11)$$

The PDF of the post-innovation cost distribution is given by the following equation:

$$dF(c) = (1 - I(c))\, dG_H(c) + I\left(\frac{c}{\delta}\right) dG_H\left(\frac{c}{\delta}\right) \tag{3.12}$$

The number of varieties consumed by the Home consumers in equilibrium is given by:

$$N = N_e \left[F(c_D) + (1 - F(c_D))\, G_F\left(\frac{c_D}{\tau}\right) \right]. \tag{3.13}$$

which depends on the exogenous number of potential varieties N_e in both countries. The first part shows the number of varieties that can be produced by the home firm (although it could be a foreign firm that is producing it by crowding out the home firm). The second part shows the number of varieties that is produced by foreign firms only because the domestic producer's cost is too high.

The joint distribution of the minimum of the domestic and foreign cost can be then calculated as follows:

$$\begin{aligned} H(c) &\equiv \Pr\left(\min\left(c^H, c^F\right) < c\right) \\ &= 1 - (1 - F(c))\left(1 - G_F\left(\frac{c}{\tau}\right)\right) \end{aligned} \tag{3.14}$$

Having obtained the final cost distribution, we can now write down the average price for the domestic economy.

The price index is simply given by:

$$\overline{p} = \frac{\int_0^{c_D} (c + c_D)\, dH(c)}{H(c_D)} \tag{3.15}$$

3.2.6 Market clearing

Finally, we need to make sure that the resource constraint is not violated and that after exporting, the remaining homogeneous good for domestic consumption is non-negative. More specifically, we need to impose that:

$$q_0 = L - R - R^* \geq 0 \tag{3.16}$$

where

$$\begin{aligned} R &= \int_0^{c_D^H} \left(1 - G_F\left(\frac{c}{\tau}\right)\right)\left((c_D)^2 - c^2\right) dF(c) \\ R^* &= \int_0^{c_D^F} (1 - F(c^*))\left((c_D)^2 - (c^*)^2\right) dG_F(c^*) \end{aligned}$$

Since domestic wage is 1, R is equal to the labor costs employed in entry, producing the differentiated goods, and innovating. R^* is the total expenditure on the foreign differentiated good, so to maintain trade balance, Home needs to export to Foreign the corresponding value in terms of homogeneous good. Therefore R^* is equal to the labor used in producing the exported homogeneous good.

3.2.7 Solving the model: Algorithm

There are two sets of endogenous variables in the model: the domestic cutoff cost, c_D, and the innovation policy function $I(c)$. We can solve the model iteratively following the steps below:

1. Guess a value of the cutoff cost, c_D^0.

2. Solve for the innovation function using equation (3.10). Then we can obtain the ex-post domestic price distribution $F(c)$. Domestic variety N, and the average cost $\bar{p}$ can be solved using equation (3.13) and (3.15). We then update the cutoff cost to c_D^1 according to equation (3.4).

3. We repeat step 2 until c_D converges.

3.3 The escape-competition and rent-destruction effects

In this section, we analyze the two opposing effects that determine how innovation reacts to competition, namely, the escape-competition effect and the rent-destruction effect. Then, we show simulation results under different production cost distributions.

We can rewrite the innovation function equation (3.10) as:

$$I(c) = \left[G_F\left(\frac{c}{\tau}\right) - G_F\left(\frac{\delta c}{\tau}\right)\right]\pi_1 + \left[1 - G_F\left(\frac{c}{\tau}\right)\right](\pi_1 - \pi_0) \tag{3.17}$$

where $\pi_1 \equiv \frac{L}{4\gamma}(c_D - \delta c)^2$ denotes the profit if the firm succeeds in innovating, while $\pi_0 \equiv \frac{L}{4\gamma}(c_D - c)^2$ denotes the profit if the firm does not succeed. In order to illustrate the effects of foreign competition on domestic innovation we differentiate the innovation function with respect to the iceberg cost τ,

$$\begin{aligned} \frac{dI(c)}{d\tau} = & \frac{\partial\left[G_F\left(\frac{c}{\tau}\right) - G_F\left(\frac{\delta c}{\tau}\right)\right]}{\partial\tau}\pi_1 \\ & + \frac{\partial\left[1 - G_F\left(\frac{c}{\tau}\right)\right]}{\partial\tau}(\pi_1 - \pi_0) \\ & + \left[G_F\left(\frac{c}{\tau}\right) - G_F\left(\frac{\delta c}{\tau}\right)\right]\frac{\partial\pi_1}{\partial\tau} + \left[1 - G_F\left(\frac{c}{\tau}\right)\right]\frac{\partial(\pi_1 - \pi_0)}{\partial\tau}. \end{aligned} \tag{3.18}$$

The first line of the equation above shows the "escape-competition" effect: when the foreign firm's production cost, after adjusting for the iceberg cost, lies in the region $[\delta c, c]$, the domestic firm can only survive if it succeeded in innovating. Higher competition would thus increase innovation in this case.

The second and third line show the rent-reduction effect. When foreigners enter in the region $\left(\frac{c}{\tau}, +\infty\right)$, the firm can survive without innovation. So, the gains from innovation comes from the usual profit gain. Since the probability of foreign firms entering in this region decreases as trade costs drop, competition reduces the expected innovation gains, and thus serves as a source of "rent-reduction". The other source of rent-reduction effect comes from the change in π_1 and π_0. In general, when foreign competition intensifies, the cutoff cost c_D decreases, making profits shrink. This would in turn induce firms to innovate less. Whether the net effect of competition is positive or negative depends on whether the escape-competition or the rent-reduction effect dominates.

It is worth discussing here whether our assumption of the monopoly pricing made in Section 3.2.2 affect the escape competition effect more or less relative to the limit pricing market structure assumption that we will make in Chapter 4. When there is limit pricing, no matter the domestic firm innovates or not, there will be four possible cases for the final price of a good. In the fist case, the foreign cost is significantly lower than the domestic one. Then the foreign firm could charge a desired mark-up. In the second case, the foreign cost is lower than domestic but not low enough so that the foreigner has to do limit pricing and charge the domestic firm's cost. In the third case, the foreign firm has slightly higher cost than the domestic firm and the domestic firm has to do limit pricing. And in the fourth case, the foreign firm is so lagged behind that the domestic firm could still charge the monopoly price. Since neither the domestic nor the foreign firms could always charge the monopoly price, the magnitude of both the escape competition effect and the rent reduction effect should decrease. Whether one decreases more than the other depends on the distribution of foreign firms. More specifically, it depends on how the four cases compose for each domestic firm, and how that compare to equation (3.18).

3.4 Simulation

This model has no analytical solution, so we now use a numerical simulation to show: i) how innovation reacts to decreases in the import transport cost τ and how that varies across the distribution of production costs, and ii) how welfare changes. We assume the production cost is distributed *Weibull* for both foreign and domestic firms. Table 3.1 shows the parameter values used in this simulation.

Table 3.1: Parameter values for full model simulation

Parameter	Description	Value
Demand		
α	Demand shifter	1.5
γ	Elasticity of substitution	6
η	Cross-price elasticity	$\frac{1}{3}$
Technology		
λ^H, λ^F	Scale parameter for the *Weibull* distribution	0.3
k^H, k^F	Shape parameter	3
δ	Cost reduction if innovated successfully	0.8
ϕ	Innovation cost coefficient	1
L	Domestic labor supply	100
N_e	Number of the differentiated varieties	20
Iceberg cost		
τ_0	Initial iceberg cost	1.5
τ_1	Iceberg cost after liberalization	1.1

We assume that the iceberg cost of importing goods drops from 1.5 to 1.1. Figure 3.1 shows the production cost distribution for domestic and foreign producers before and after trade liberalization. We can see that the yellow line (representing foreign cost after liberalization) lies above the red line (representing foreign cost before liberalization) when the effective cost is below about 0.6. The escape-competition effect, corresponding to the first line of equation (3.18), is positive for these firms.

Figure 3.2 shows the innovation level as function of the production cost draws of the domestic producers. The blue line shows before liberalization, and the orange line shows the schedule after liberalization. The model can generate a behavior consistent with what we find in China during the WTO accession. Firms with lower production costs increase their innovation efforts after import competition intensifies while other, less productive firms do not increase their investment.

Figure 3.1: Production cost distribution for the simulation

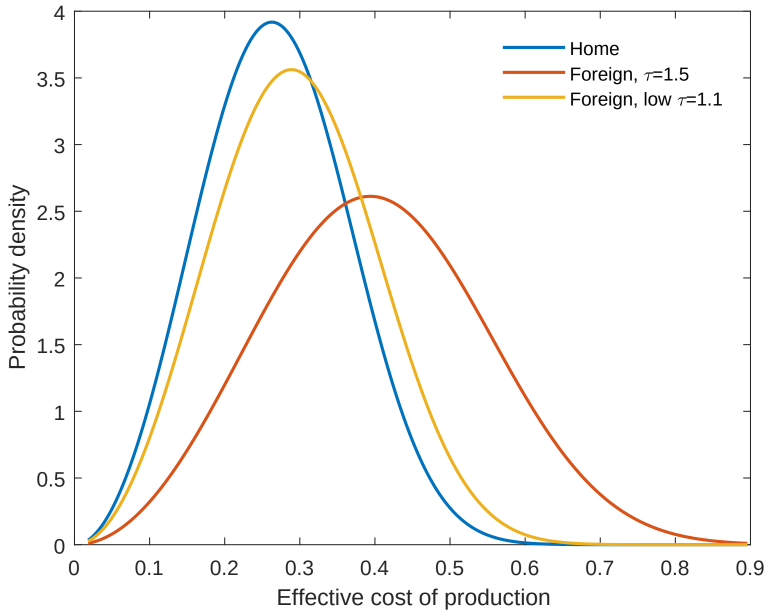
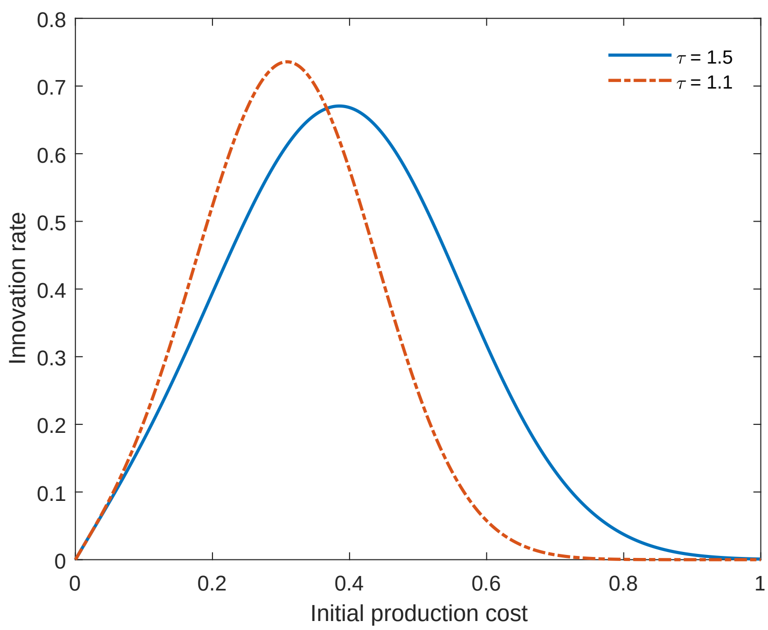


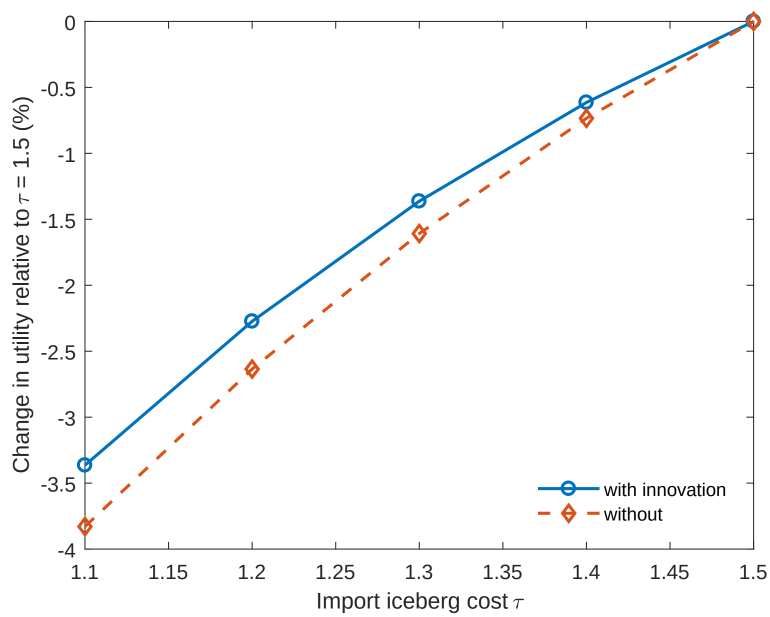
Figure 3.2: Innovation efforts for different firms



Finally, we are interested in how the heterogeneous reaction in innovation affects aggregate welfare, which is then compared to the case without innovation as a benchmark. In addition to the two cases where $\tau_0 = 1.5$ and $\tau_1 = 1.1$, we also add some intermediate points to illustrate the transition. Figure 3.3 shows the simulated results. The solid blue line shows, for the case with innovation, the percentage change in utility relative to the initial state where $\tau_0 = 1.5$. The dashed orange line shows how utility changes when innovation is allowed. At first it may appear unintuitive that the two curves are upward sloping, meaning that as trade is liberalized, utility declines. In fact, this phenomenon is not uncommon in other trade models. For example, both Ossa (2011) and Melitz and Ottaviano (2008) find that unilateral trade liberalization is welfare reducing in models where free entry delivers the Metzler Paradox, whereby import tariff decreases cause the domestic price index to increase. In Section 4.4 in the next chapter, we give more intuition about why this happens by analyzing the different sources of welfare change in a simpler version of the model.

Regardless of whether the baseline trade model delivers negative gains from trade, the relevant result for our context is that utility declines more slowly in the presence of endogenous investment compared to the case of no investment, hence pointing to a potential new source of gains from trade. The reason for these additional gains from trade is the topic of next chapter, but essentially relies on the presence of under-investment in the decentralized equilibrium of the model.

Figure 3.3: Percentage change in utility



3.5 Conclusion

In this chapter, we built a model to help us analyze the heterogeneous effects of import competition on innovation among firms. We emphasize two opposing channels through which trade liberalization affects innovation: the escape-competition channel, and the rent-destruction channel, in the spirit of Aghion et al. (2001) and Aghion et al. (2009). We find that the positive escape-competition effect can dominate the negative rent-destruction effect for the most productive firms. Thus, the model rationalizes the empirical findings, shown in Chapter 2, that Chinese high productivity firms increased innovation as they faced tariff cuts in the period 2001-2005.

Chapter 4

Welfare analysis

4.1 Introduction

In this chapter, we analyze a simplified setup in order to more clearly highlight the mechanism behind the welfare consequences of competition-induced innovation. More specifically, we eliminate one of the two sources of competition present in the general model, i.e. the one coming from average price variations. Therefore, the only source of competition in this simplified setup is the foreign firm producing an identical product and competing directly with the corresponding home firm.

The goal of this chapter is to compare welfare changes in the presence of competition-induced innovation to the standard gains from trade present in a benchmark model without endogenous innovation. The chapter has three main findings. First, we find that, relative to a constrained social optimum, the decentralized equilibrium always features under-investment. This is because, as will become clear later in the chapter, the domestic social planner is not affected by the business-stealing aimed at foreign firms. The second finding is that, because of under-investment, whether there are additional welfare gains from innovation depends on whether innovation increases after competition increases. As we saw in the previous chapter, this happens if the escape-competition effect prevails over the rent-destruction effect. The third finding is that the simplifying assumption we made in the previous chapter, which eliminates limit pricing, has important welfare implications. Under such assumption, the benchmark model without endogenous innovation entails negative gains from trade. On the contrary, under limit pricing, trade is always welfare enhancing in the benchmark case of no endogenous innovation. This observation explains why in the model by Akcigit et al. (2017) welfare increases upon the imposition of higher import tariffs.

4.2 A simplified setup

In the last chapter, we assumed that there is a distribution of domestic and foreign productivities, which, together with the iceberg cost, determine the probability of foreign entry. In this chapter, we show that the welfare analysis is simplified if, instead of considering the entire foreign and domestic productivity distribution, we only consider two variables, e and θ, which we now describe. First, the reduction of transport cost can be viewed as a simple increase in the probability of foreign firm entry, denoted by e, as further explained below. Second, there are two relevant cases for welfare, one in which there is no escape-competition effect and one in which such effect is strong enough to overcome the rent-destruction effect. Whether we are in one or the other case depends on the productivity of the foreign firm relative to the domestic firm. We assume therefore a simple distribution of foreign productivity described as follows. Domestic firm production cost is c. Foreign firm production cost is:

$$c^* = \begin{cases} c_1^* = c & \text{Prob } \theta \\ c_0^* \ll c & \text{Prob } 1-\theta \end{cases}$$

where $\theta \in [0,1]$ is the probability that the foreign firm is close enough in productivity to the domestic firm that the Home firm can survive if it innovates. Moreover $1-\theta$ is the probability that the Foreign firm is too productive relative to the domestic firm and therefore it will take over the domestic market whether the domestic firm innovates or not. We present the welfare analysis in terms of the two variables e and θ under the further four simplifying assumptions.

Assumption 1 There is no competition among differentiated varieties $\eta = 0$.

Assumption 2 The innovation step $\frac{1}{\delta}$ is large enough so that if the foreign firm is of type c_1^*, the domestic firm can charge the monopoly price after innovating. $\frac{\alpha+\delta c}{2} < c$.

Assumption 3 The foreign firm of type c_0^* can always charge the monopoly price: $\frac{\alpha+c_0^*}{2} < \delta c$.

Assumption 4 When foreign and domestic firms are equally productive, the foreign firm produces.

The first assumption implies that the cutoff cost c_D is α. Thus, this assumption helps eliminate the endogenous market price change induced by the change in c_D, which simplifies the profit, innovation, and thus utility expressions. Without the endogenous price change, the rent destruction effect could become smaller, as the third line of equation (3.17) disappears. Since this effect is a market effect and thus same across firms, it does not affect the relative behavior among firms, nor should it affect the qualitative comparison between the decentralized and constrained social optimal utilities that we would discuss later.

The second and third assumptions simplifies pricing, so that even if we allow for limit pricing, the only situation where limit pricing could happen is when the foreign firm enters at c_1^*. Assumption 4 breaks the tie.

It is easy to draw a relationship between the more general setup in the previous chapter and this simplified case. The domestic firm's optimizing problem is the following:

$$\max_{I} \qquad e\theta I\pi_1 + \left[I\pi_1 + (1-I)\,\pi_0\right](1-e) - \frac{1}{2\phi}I^2 \tag{4.1}$$

where π_1 again denotes the profit after innovating, and π_0 denotes the profit without innovation. The innovation schedule, denoted by the superscript d for decentralized optimum, can be solved as follows:

$$\frac{I^d}{\phi} = (1-e)\,(\pi_1 - \pi_0) + e\theta\pi_1 \tag{4.2}$$

In comparison with equation (3.17) in Chapter 3, $(1-e)$ corresponds to $\left[1 - G_F\left(\frac{c}{\tau}\right)\right]$, i.e. the probability that the foreign firm does not enter the domestic market. This term governs the rent-destruction effect. The term $e\theta$ corresponds to $\left[G_F\left(\frac{c}{\tau}\right) - G_F\left(\frac{\delta c}{\tau}\right)\right]$, i.e. the probability that the foreign firm enters and that it is close enough to the domestic firm that it can be kept at bay if the domestic firm innovates. This term therefore governs the escape-competition effect.

We take a step back to look at the implication of our pricing rule on firms' innovation incentive, as we did in Section 3.3. Due to our assumption on the production costs and innovation step size, the profit under innovation for domestic firms, π_1, is the same for the limit pricing and monopoly pricing cases. The only difference is π_0, the profit of domestic firms when there is no foreign entry. This term would be higher under monopoly pricing assumption. Therefore, the rent reduction effect is expected to be higher, and the overall innovation incentive would be smaller when the market structure is assumed as in the previous chapter.

4.3 Constrained optimum of the social planner

The first goal of this chapter is to investigate whether the decentralized equilibrium features excessive or sub-optimal innovation. We therefore derive the socially optimal innovation schedule given firms' price and quantity choices. In this sense, this is a constrained optimum, because the social planner is still deciding how much to invest given the private choice of firms in terms of quantity and prices. The social planner problem is the following:

$$\max_{I} \quad u = q_0^c + \alpha \mathbb{E}q - \frac{1}{2}\gamma \mathbb{E}q^2 \tag{4.3}$$

$$s.t. \quad 1 = q_0^c + \mathbb{E}l_d \tag{4.4}$$

where

$$\begin{aligned}
\mathbb{E}q &= e\left(\theta\left(Iq_m(\delta c) + (1-I)\,q_l(c)\right) + (1-\theta)\,q_m(c_0^*)\right) + (1-e)\left(Iq_m(\delta c) + (1-I)\,q_m(c)\right) \\
\mathbb{E}q^2 &= e\left(\theta\left(Iq_m^2(\delta c) + (1-I)\,q_l^2(c)\right) + (1-\theta)\,q_m^2(c_0^*)\right) + (1-e)\left(Iq_m^2(\delta c) + (1-I)\,q_m^2(c)\right) \\
\mathbb{E}l_d &= e\left(\theta\left(I\delta c + (1-I)\,c\right) + (1-\theta)\,p_m(c_0^*)\right) + (1-e)\left(I\delta c + (1-I)\,c\right) + \frac{1}{2\phi}I^2
\end{aligned}$$

$\mathbb{E}l_d$ denotes expected labor demand for producing the differentiated good and the exported homogeneous good; $q_m(c) = \frac{\alpha - c}{2}$ denotes quantity under monopoly pricing; and $q_l(c) = \alpha - c$ denotes the quantity under limit pricing. If we substitute the homogeneous good quantity using the resource constraint (4.4), utility (4.3) can be written as

$$u^{innov} = 1 - \frac{1}{2\phi}I^2 + IW_1 + (1-I)\,W_0 \tag{4.5}$$

where

$$\begin{aligned}
W_1 &= e\left[\theta\left(\pi_1 + \mathrm{CS}_m(\delta c)\right) + (1-\theta)\,\mathrm{CS}_m(c_0^*)\right] + (1-e)\left[\pi_1 + \mathrm{CS}_m(\delta c)\right] \\
W_0 &= e\left[\theta \mathrm{CS}_l(c) + (1-\theta)\,\mathrm{CS}_m(c_0^*)\right] + (1-e)\left[\pi_0 + \mathrm{CS}_m(c)\right]
\end{aligned}$$

$\mathrm{CS}_a(c) = \frac{1}{2}(\alpha - p_a(c))\, q_a(c)$ denotes the consumer surplus. When the customers are charged a markup, $a = m$. When the customers are charged at cost (the firm has to do limit pricing) $a = l$.

Denote the constrained optimal solution by the superscript co, the innovation function is as follows:

$$\begin{aligned}\frac{I^{co}}{\phi} &= e\theta\pi_1 + (1-e)(\pi_1 - \pi_0) \\ &\quad + e\left[\theta\left(\mathrm{CS}_m(\delta c) - \mathrm{CS}_l(c)\right)\right] + (1-e)\left(\mathrm{CS}_m(\delta c) - \mathrm{CS}_m(c)\right)\end{aligned} \tag{4.6}$$

The first line is exactly equal to the decentralized innovation in equation (4.2) and represents the additional profit derived from innovation. If there is entry, the additional profit is π_1. If there is no entry the additional profit is $(\pi_1 - \pi_0)$. The second line is the consumer surplus created by innovation, which is not internalized by the firm when it is making a private innovation investment decision. Under Assumption 2, $\frac{\alpha+\delta c}{2} < c$, it is easy to show that $\mathrm{CS}_m(\delta c) > \mathrm{CS}_l(c)$ and therefore the consumer surplus part, the second line of equation (4.6), is always positive.

Lemma 1. *The decentralized equilibrium always features under-investment: $I^d < I^{oc}$.*

4.4 Openness and welfare

In this section we investigate how utility is affected by an increase in openness, represented here by an increase in the probability of foreign entry,e. The goal of this section is to show that openness will provide additional utility gains when the innovation response to trade is positive, i.e. when the escape-competition effect dominates.

First, it is helpful to have a benchmark utility where there is no change in innovation as a result of increased openness. It is still important to have some initial investments in order to start from the same average productivity level in the Home country. Therefore, setting innovation to the initial decentralized optimal level I_0^d, we can rewrite the utility function as follows:

$$u^{noinnov} = 1 - \frac{1}{2\phi}\left(I_0^d\right)^2 + I_0^d W_1 + \left(1 - I_0^d\right) W_0. \tag{4.7}$$

We can then take the total differential of utility with respect to e at the initial decentralized optimum for both the case of endogenous innovation and the case of no innovation:

$$\frac{du^{noinnov}}{de} = \frac{\partial u\left(I_0^d\right)}{\partial e} \tag{4.8}$$

$$\frac{du^{innov}}{de} = \frac{\partial u\left(I_0^d\right)}{\partial e} + \frac{\partial u}{\partial I^d}\frac{\partial I^d}{\partial e} \tag{4.9}$$

We are interested in the sign of the two derivatives $\frac{du^{noinnov}}{de}$ and $\frac{du^{innov}}{de}$, and in whether the gains from trade are larger under endogenous innovation, i.e. $\frac{du^{noinnov}}{de} \gtrless \frac{du^{innov}}{de}$.

Lemma 2. *When openness increases, the utility without innovation is increasing. That is, for* $\forall\theta$, $\frac{du^{noinnov}}{de} > 0$.

Appendix B provides the proof of the lemma. The intuition is that, foreign entry either enables the consumers to consume a better product (when $c^* = c_0^*$) or consume the same quality product at a lower price (when $c^* = c$). There will be losses to the domestic firms. But since there is less distortion in the aggregate economy, utility will increase.

The market structure that requires limit pricing is crucial here. When we make the assumption that allows for monopoly pricing, as we did in Chapter 3, and in Akcigit et al. (2017), welfare can decrease as e increases, because the consumer gains are dominated by the producer losses. Take the extreme case where all foreigners come in at the same production cost as the domestic firm ($\theta = 1$). As foreign entry increases, consumers face the same prices since domestic firms are replaced by foreign firms, but both are charging monopoly prices. The only relevant change is a transfer of domestic profits to the foreigners (Assumption 4), which reduces domestic welfare.

This observation is consistent with Ossa (2011) (CES demand) and Melitz and Ottaviano (2008) (Linear demand). Both models feature the monopolistic competition market structure in the differentiated good sectors and a perfect competitive homogeneous sector to balance trade. Both models find that a unilateral import liberalization would cause welfare losses for the liberalizing country. The intuition is exactly as discussed above. As phrased by Ossa (2011), there are two effects after a unilateral import tariff drop. First, there is a production relocation effect. Domestic consumers shift expenditure toward the foreign goods that are cheaper now. This is exactly the loss to domestic producers in our model. Second, there is an import price effect that increases consumer surplus because they now can consume at a lower price. In Ossa (2011) as well as in our model with the price setting assumption, the negative production relocation effect dominates[13].

Next, we study whether the utility increases faster with innovation relative to the benchmark. It suffices to investigate whether $\frac{\partial I^d}{\partial e}$ is positive, because under Lemma 1, there is always under-investment, i.e. $\frac{\partial u}{\partial I^d} > 0$. From equation (4.2), we can take derivative of innovation with respect to entry rate e,

$$\frac{\partial I^d}{\partial e} = -(\pi_1 - \pi_0) + \theta\pi_1$$

The first part of the partial derivative is the negative rent reduction effect, and the second part is the positive escape-competition effect. We can see that, the escape-competition effect decreases as more foreign firms come in with very low production cost, thus there is nowhere to escape for the domestic firms. Therefore, there is a cutoff condition for θ that determines whether the gains from trade are larger under endogenous innovation. We summarize in the following lemma.

[13] Demidova (2017) finds another way of changing the relative magnitude of the relocation effect and the price effect. She shows that when the homogeneous sector is eliminated and wages are allowed to adjust, one obtains the more intuitive prediction that trade is welfare enhancing.

Lemma 3. $\frac{du^{noinnov}}{de} \geq \frac{du^{innov}}{de}$ *iff* $\theta \geq \hat{\theta}$ *where*

$$\hat{\theta} \equiv \frac{\pi_1 - \pi_0}{\pi_1}. \tag{4.10}$$

That is, when the escape-competition effect dominates the rent-destruction effect, there would be an additional utility gain under import liberalization.

4.5 Simulation

In this section, we simulate the cases studied to provide a graphical representation of the results obtained above. Table 4.1 shows the parameter values used.

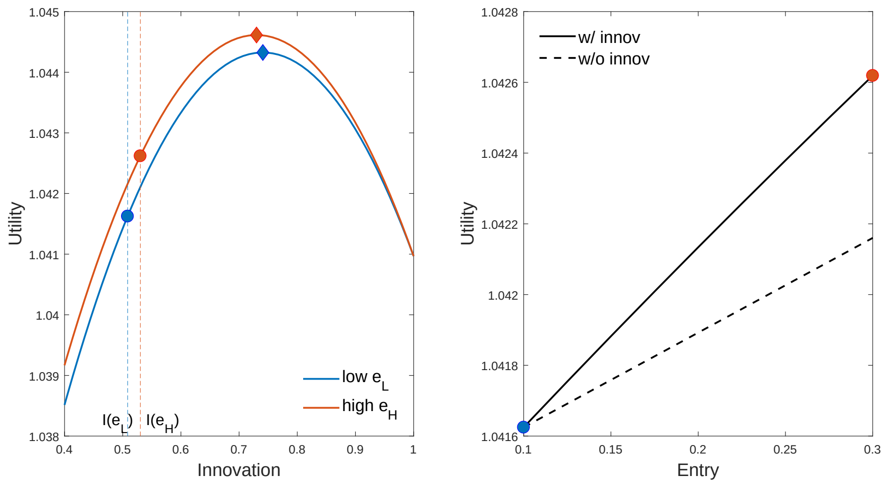
Table 4.1: Parameter values for welfare simulation

Parameter	Description	Value
Demand		
α	Demand shifter	2
γ	Elasticity of substitution	3
Technology		
c	Home firm production cost	1.64
c_0^*	Production cost of foreign good producers	0.29
δ	Cost reduction if innovated successfully	0.7
ϕ	Innovation cost coefficient	10

In Figure 4.1, we show the decentralized innovation and the utilities as the entry rate e increases from 0.1 to 0.3, when the foreign firm is of the same productivity as the home firm, i.e. $\theta = 1$ and $c^* = c$ with probability 1. The quadratic curves on the left panel shows utility as a function of innovation I. The circles denote the decentralized choices of innovation I^d, and the diamonds denote the constrained optimal choices I^{oc}, for $e = 0.1$ and $e = 0.3$. We can see that the circles always lie to the left of the diamonds, which is consistent with under investment.

On the right panel, we show utility as function of e. The solid line shows utility with endogenous innovation. The dashed line shows the benchmark utility setting innovation to the level chosen when $e = 0.1$. Since this case features a strong escape-competition effect, innovation increases as foreign entry increases, and there is an additional positive gain from innovation.

Figure 4.1: Foreign cost is the same as domestic, $\theta = 1$

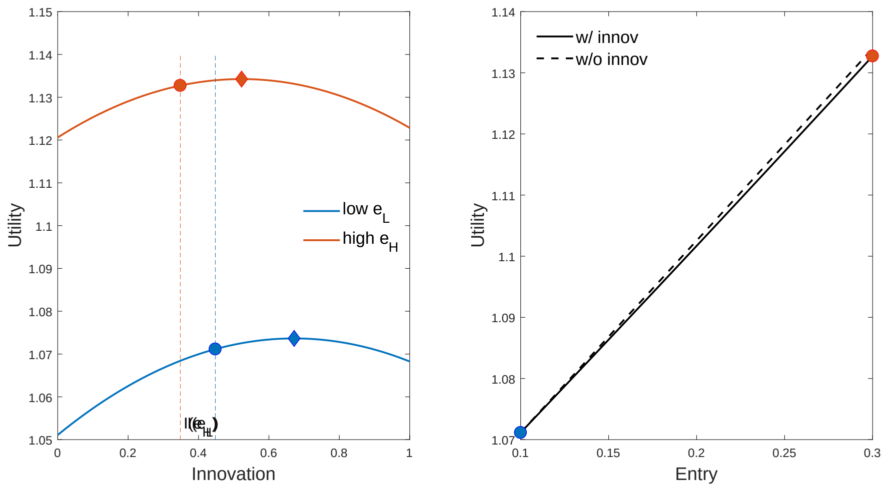


Note: The circles on both panels indicate the decentralized innovation decisions and the resulting utility levels. On the left panel, the diamond symbol marks the constrained optimum level of innovation.

Figure 4.2 shows the case when the foreign firm's cost is much lower than the domestic firms, i.e. $\theta = 0$ and $c^* = c_0^*$ with probability 1. In this case, there is only a rent-destruction effect when innovation is endogenous. Therefore, as shown on the left panel, as e increases, the decentralized innovation effort decreases. And on the right panel, the benchmark utility slopes more steeply than the utility with innovation, indicating a negative gain from endogenous innovation response.

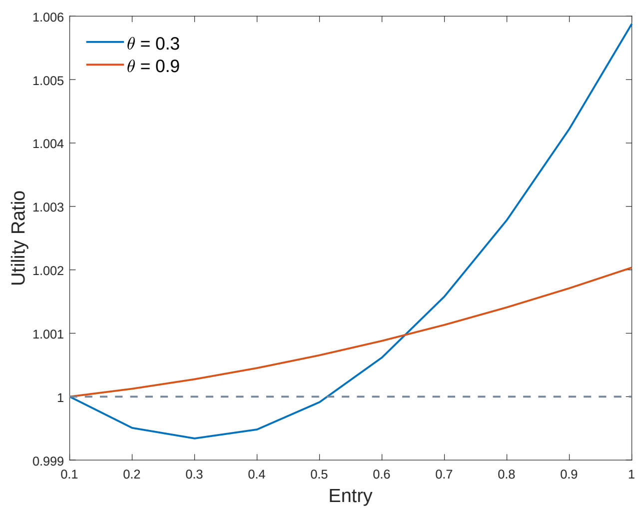
Figure 4.3 shows the ratio of aggregate utilities $\frac{u^{innov}}{u^{noinnov}}$ for different compositions of foreign firms. The red line shows a composition where most of the foreign competitors are close to the domestic firms in productivity ($\theta = 0.9$). In this case, θ is above the cutoff $\hat{\theta}$ defined in equation (4.10) and parameterized in Table 4.1. The escape-competition effect dominates, and the ratio of the utility with innovation to the benchmark utility is always above 1. The blue line shows a composition where most of the foreign competitors are much more productive than the domestic firms ($\theta = 0.3$). In this case, the rent-destruction effect dominates, though its magnitude decreases as e increase. In fact, when $e = 1$, the rent-destruction effect goes to zero. Therefore, the blue line first decreases below 1 then increases to above 1.

Figure 4.2: Foreign cost is much lower than domestic, $\theta = 0$



Note: The circles on both panels indicate the decentralized innovation decisions and the resulting utility levels. On the left panel, the diamond symbol marks the constrained optimum level of innovation.

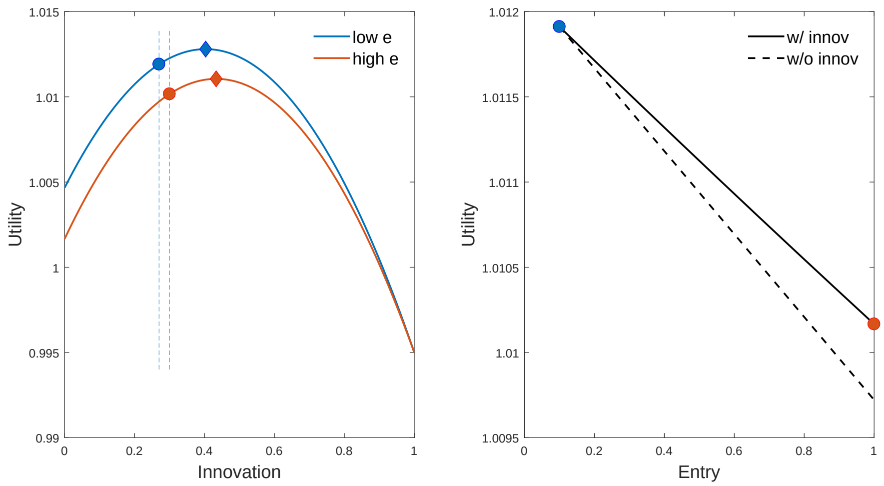
Figure 4.3: Utility ratios $u^{innov}/u^{noinnov}$



Finally, we simulate a case with the assumption that firms need to incur a cost ε to set the price, like in Akcigit et al. (2017), bringing back monopoly pricing. In Figure 4.4, we replicate Figure 4.1 where all foreign firms are of the same productivity as the domestic ones, with the assumption that eliminates limit pricing. As discussed under Lemma 2, the benchmark utility decreases with trade liberalization. See the dashed line on the right panel.

Since in this case, the escape-competition effect dominates the rent-destruction effect (because $\theta > \hat{\theta}$), according to Lemma 3, innovation could contribute to a slower decrease in the aggregate utility.

Figure 4.4: Eliminating limit pricing



Note: The circles on both panels indicate the decentralized innovation decisions and the resulting utility levels. On the left panel, the diamond symbol marks the constrained optimum level of innovation.

Chapter 5

How the Breadth and Depth of Import Relationships Affect the Performance of Canadian Manufacturers

5.1 Introduction

So far in this dissertation, we have looked at the effects induced by changes in import tariff. In this Chapter, we look at another factor that is important to trade — the supplier-buyer relationship. Specifically, we study how the variety of import relationships and the average duration of the relationships affect the performance of Canadian manufacturing importers.

The love of variety forms the basis for the gains from trade in all trade models based on the Armington (1969) assumption or on Dixit-Stiglitz monopolistic competition. It therefore underpins work based on Melitz (2003) and most computable general equilibrium evaluations of trade liberalizations. While the existing evidence focuses on the empirical relevance of the love-of-variety for final goods, there is remarkably little evidence on its implications for intermediate and capital goods purchased by firms, which constitute the bulk of trade.[14,15] With inputs acquired by firms, we rely on the Ethier (1982) theoretical demonstration that the love-of-variety idea can be extended to production functions. Ethier (1982) adopts a parallel version of the Dixit-Stiglitz utility function as an objective for the firm; in this framework, additional inputs increase output in proportion to the total number of products acquired for production. In this chapter, our two proxies for *breadth* are the number of 10-digit products a manufacturing firm imports and the number of supplying firms per imported product. We estimate the elasticities of productivity with respect to both variables.

A complementary view on how import relationships shape firm performance comes from the management literature. In particular, Uzzi (1996) applied Karl Polanyi's idea of *embeddedness* to production networks. He argues that "buyer-supplier networks operate in an embedded logic of exchange that promotes economic performance through inter-firm resource pooling, cooperation, and coordinated adaptation[...]" (Uzzi (1996), p. 675). Using data on New York-based apparel

[14] The groundbreaking work by Broda and Weinstein, 2006 has been the first to structurally estimate the impact of increased variety for welfare. For a recent literature review, see Feenstra (2010).

[15] Miroudot et al. (2009) document that trade in intermediates and capital goods accounted for about 70% of the total Canadian imports in 2006.

firms, he finds that a firm that systematically interacts with a network of suppliers enjoys better outcomes in terms of survival and productivity relative to firms that keep all their transactions at arm's length and do not engage in long-term relationships.[16] Inspired by Uzzi, we use the share of continuous suppliers over the total number of suppliers as our principal measure of relationship depth. It is expected to increase productivity and other performance measures.

Analogously to Kasahara and Rodrigue (2008), we adopt the control function approach of Levinsohn and Petrin (2003) to account for unobserved productivity shocks at the firm level. We further assume that importing decisions are dictated by the presence of fixed costs that are heterogeneous across firms and not perfectly correlated with productivity, so that the effect of importing decisions can be identified. By controlling for intermediate inputs, we also control for the cost reducing channel that a broader variety of inputs could have for importers. Therefore, we try our best to measure the love-of-variety channel using our breadth variables.

Our results show that the number of imported products and the number of suppliers per product increase the size of Canadian importing manufacturers with elasticities of 0.15 and 0.12, respectively. The breadth effects drop to 0.03 and 0.02 after controlling for inputs and including a control function to account for unobserved productivity. We also quantitatively explore how important continuous relationships are to the performance of importing firms. We document that older relationships are more valuable and increase firm size and productivity. An importer that went from using all new suppliers to retaining all the prior year suppliers could increase its productivity by 2.4%. The importance of ongoing relationships is also reflected in our analysis of the size and value of transactions between an importing firm and its long-term partners: both the quantity imported, and the associated unit value are larger. However, after controlling for inputs, the ongoing use of the same suppliers does not have any statistically significant effects on performance in foreign markets. Finally, we analyze the influences of the suppliers' country of origin by including as explanatory variables the share of suppliers from China and the United States. Greater reliance on Chinese suppliers is associated with smaller firm size and has a negative impact on exporting performance; its effect, however, is measured imprecisely, and it is not always significant at the 5% level.

This chapter contributes to the large empirical literature documenting productivity differences across firms differing in their import choices. Data from the United States, Belgium, Italy, Hungary, Colombia, and Chile reveal that importers are bigger in terms of employment, shipments, value added, and TFP if compared with non-importing firms.[17] In fact, firm heterogeneity in importing behavior has important implications for the measurement of the gains from trade, especially when large firms import proportionally more of their inputs.[18]

Our paper also relates to recent work that has emphasized the two-sided nature of trade rela-

[16]Uzzi (1997) and Uzzi (1999) extend these ideas.

[17]See Bernard et al. (2007) for the United States; Halpern et al. (2015) for Hungary; Muûls and Pisu (2009) for Belgium; Castellani et al. (2010) for Italy; Kugler and Verhoogen (2009) for Colombia; Kasahara and Rodrigue (2008) and Kasahara and Lapham (2013) for Chile. Episodes of trade liberalizations provide additional evidence on the productivity gains from importing; see, for example, Amiti and Konings (2007), Goldberg et al. (2009), and Topalova and Khandelwal (2011).

[18]See Blaum et al. (2017) and Ramanarayanan (2017).

tionships. Several contributions have analyzed the buyer-supplier margin using export and import transaction data. Bernard et al. (2017) and Carballo et al. (2013) describe the behaviour of Norwegian and South American (Costa Rica, Uruguay and Peru) exporters. More recently, other contributions have focused on the formation of buyer-supplier relationships. Eaton et al. (2015) calibrate a search-and-matching model to match the trade patterns between U.S. buyers and Colombian exporters. Monarch (2014) quantifies the magnitude of frictions between U.S. buyers and Chinese suppliers in finding new partners. Kamal and Sundaram (2016) identify the existence of importer-specific spillovers in the decision of Bangladeshi manufacturers to sell to U.S. importers. Dragusanu (2014) analyzes the matching between buyers and suppliers in a model of sequential production.

A closely related contribution is the paper by Lu et al. (2016), who build a model to analyze the switching behaviour of Colombian importers. Consistent with our findings, they document that Colombian firms importing more products from a larger set of suppliers tend to be larger. While their approach combines productivity and scale effects, our contribution, instead, tries to identify the productivity effects of different dimensions of importing using the control function approach.

The question of the importance of supplier networks for productivity is also the focus in a paper by Bernard et al. (2017), where the authors find a positive effect on productivity and on the number of domestic supplier connections after the opening of high-speed train lines in Japan. Our elasticity estimates, however, are not comparable to theirs because they focus on the reduced form effects in a difference-in-difference strategy; in fact, their identification relies on differences in performance between input intensive firms and labor-intensive firms located close to a new train station relative to firms in locations without a new station, before and after the high-speed train expansion. Our elasticities, instead, are informative of the productivity effects associated with an exogenous change in the breadth and depth variables.

The rest of the chapter is organized as follows. We describe the data in section 5.2; we analyze the main features of the data in subsection 5.2.1. We present our empirical strategy in section 5.3. The results are shown in section 5.4. Section 5.5 concludes.

5.2 Data

The data for our project comes from three sources: The Import Registry, the Annual Survey of Manufactures (ASM)-T2LEAP, and the Export Registry.

The import registry collects transaction data using Form B3 from the Canadian Border Service Agency. Canadian importers are required to fill information on the vendor's name and address, the country of export, the product (HS10 code), the imported value and quantity. Identifiers were created for each supplier from the vendor's name and address.[19] Transaction records with consistent suppliers' identifiers are available from August 2002 to June 2008.[20]

The raw data identifiers are the transaction number, the line number (a particular item in a transaction, often corresponding to a deeper level of disaggregation than a HS10 code), and the

[19]See Appendix C.1 for a summary on the methodology.
[20]Import records at the product-, origin-, and firm-level are available since 1993.

date (month-year). We aggregate the data across transactions to the firm-supplier-HS10-country of origin-year level. The initial dataset contains about 5.5 million observations (corresponding to the firm-supplier-HS10-origin-year combination).

In order to construct firm-level measures of performance, we merge the import customs with firm-level information drawn from the Annual Survey of Manufactures (ASM). The ASM is a survey covering the universe of manufacturing establishments. It includes data on shipments, industry classification (5-digit NAICS codes), employment, salaries and wages, cost of materials, and expenditure on electricity. We enrich the ASM dataset by adding information on assets and investment extracted from the T2-LEAP database. T2-LEAP links two administrative data sources, the Longitudinal Employment Analysis Program (LEAP) and the Corporate Tax Statistical Universal File (T2SUF). Those two sources include all firms that either register a payroll deduction account with the Canada Revenue Agency (CRA) or file a T2 tax return with the CRA. The capital/investment data reported in T2-LEAP encompass manufacturing and non-manufacturing activities of each firm; we therefore allocate capital/investment to the individual manufacturing establishments using the share of the establishment revenues in manufacturing over the total firm sales.

We merge the import registry with firm-level characteristics and we collapse the information on import choices at the firm-year level. This creates our final dataset with 93,386 observations (here an observation is a firm-year combination).

Export-related information on Canadian firms comes from the Canadian Export Customs. The custom data include export records at firm-, product (HS8 code)-, and destination-level for the universe of exporters located in Canada.

5.2.1 Import Network Characteristics: Breadth and Depth

This subsection explores the main features of the Canadian import registry. We focus our discussion on cross-sectional and dynamic characteristics of the importers' distribution. Table 5.1 summarizes the main cross-sectional aggregates by sector in 2007.

Columns (1)–(2) describe the intensive import margin: column (1) shows the total import value for each sector, while column (2) reports the share of imports out of total manufacturing sales. Although chemical and oil imports are the largest industries in terms of value, other sectors–namely, Computing, Apparel, and Transportation Equipment–are relatively more dependent on foreign products. Some sectors, such as Beverages & Tobacco and Apparel, display import shares that are larger than our estimates of the share of materials in production (see tables C.4 to C.6). This finding may be due to *carry-along trade*, the fact that firms tend to import both intermediate inputs and final consumption goods.[21]

Columns (3)–(7) focus on the extensive import margin: they show the number of countries, products (HS10 codes), Canadian buyers, foreign suppliers and buyer-supplier relationships. Each sector imports a large number of products (from 9% of all HS10 codes in Leather to 46% in Machinery) from a large number of countries (the median sector imports from 81 countries). The

[21]See Bernard et al. (2017) for a detailed theoretical and empirical analysis of carry-along trade.

large scope of the imported products raises concerns on secondary wholesale activities. While we focus on firms in the manufacturing sectors, this classification requires that the majority of firm revenues comes from manufacturing activities; thus, we cannot exclude that those firms may include plants whose industry code is in wholesale or in other non-manufacturing sectors. A similar caveat applies to firm-level statistics (see column (4) in table 5.2). In the empirical analysis, we rely on firm fixed effect to capture time-invariant differences in activity classifications across firms. Looking across columns (5)–(7), we note that the number of relationships is mainly driven by the number of suppliers. This fact suggests that Canadian firms tend to adopt a multi-sourcing strategy, as micro-level statistics will confirm.

Table 5.1: Aggregate Statistics by 3-digit industry, 2007

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Industry	Imp. Value[1]	Imp. Share	Countries	Products	Firms	Suppliers	Relations
Food	7.90	0.09	115	6067	1396	18684	29624
Bev. & Tob.	2.01	0.34	71	2422	145	3641	4619
Text. Mills	0.61	0.58	55	2332	197	3251	4197
Text. Prod	0.55	0.55	55	2717	301	3947	4760
Apparel	1.18	0.67	80	3326	723	10217	14549
Leather	0.13	0.62	48	1518	146	1801	2190
Wood	1.99	0.11	72	3750	1052	11462	16401
Paper	3.75	0.22	74	3580	383	8989	12997
Printing	0.78	0.18	54	3077	941	6971	9875
Petrol	18.22	0.19	57	2392	86	3153	3840
Chemical	18.82	0.57	104	7345	948	22158	33878
Plastics	7.21	0.43	83	5964	1218	19204	28477
Mineral	2.23	0.25	72	4307	713	8361	11598
Metals	11.85	0.33	90	3756	325	8839	11436
Met. Prod	5.54	0.31	89	6987	2980	28155	40804
Machinery	11.23	0.48	118	7760	2436	40088	61318
Computing	9.84	0.80	115	5309	1008	30605	47549
Electrical	4.13	0.65	89	4385	588	13630	17644
Tran. Eq.	74.75	0.63	123	7143	1029	40726	65539
Furniture	3.27	0.29	87	5377	1104	12704	17651
Miscel.	3.68	0.63	102	6542	1648	17713	21608
n/a	0.16	1.32	62	4316	2065	5883	6594
Total Mfg	189.83	0.62	194	16721	21432	233718	467148

[1] Values in millions.
Notes: Aggregate import statistics by sector. The last row reports the totals for all manufacturing.

Table 5.2 takes a closer look at the importing behavior of firms, with a focus on 2007. The first two columns report the firm-level average import value and import share across sectors, confirming the patterns shown in columns (1)-(2) of Table 5.1. Oil companies are the biggest importers, although their share of imports out of total sales is small compared with firms in other industries.

Columns (3)-(5) focus on the extensive margin. The quasi-median firm sources its inputs from 2 countries and imports multiple products from a large set of suppliers.[22] This evidence confirms a strong multi-sourcing nature of the Canadian import relationships.[23]

Table 5.2: Firm-level statistics on importing, 2007

Industry	(1) Import value[2]	(2) Import share	(3) Sources[2] /firm	(4) Products[2] /firm	(5) Supps[2] /firm	(6) Avg age
Food	366.57	0.09	2	12	9	1.4
Bev. & Tob.	296.02	0.13	2	12	10	1.2
Text. Mills	484.16	0.39	3	13	11	1.5
Text. Prod	212.10	0.33	2	13	10	1.6
Apparel	299.75	0.38	4	17	12	1.3
Leather	165.65	0.36	3	11	9	1.7
Wood	170.38	0.07	1	7	5	1.5
Paper	638.63	0.21	1	11	10	1.7
Printing	82.53	0.06	1	7	6	1.3
Petrol	5140.36	0.14	2	41	25	1.6
Chemical	518.09	0.24	2	20	14	1.5
Plastics	344.65	0.17	2	14	11	1.6
Mineral	189.25	0.13	2	12	8	1.7
Metals	800.08	0.20	2	14	14	1.6
Met. Prod	162.28	0.11	1	10	7	1.6
Machinery	257.82	0.18	2	15	11	1.6
Computing	384.29	0.33	3	23	16	1.4
Electrical	363.42	0.31	3	16	14	1.5
Tran. Eq.	547.98	0.24	2	25	17	1.7
Furniture	137.99	0.10	2	11	8	1.5
Miscel.	125.33	0.21	2	9	8	1.5
n/a	19.32	-	1	3	3	1.9

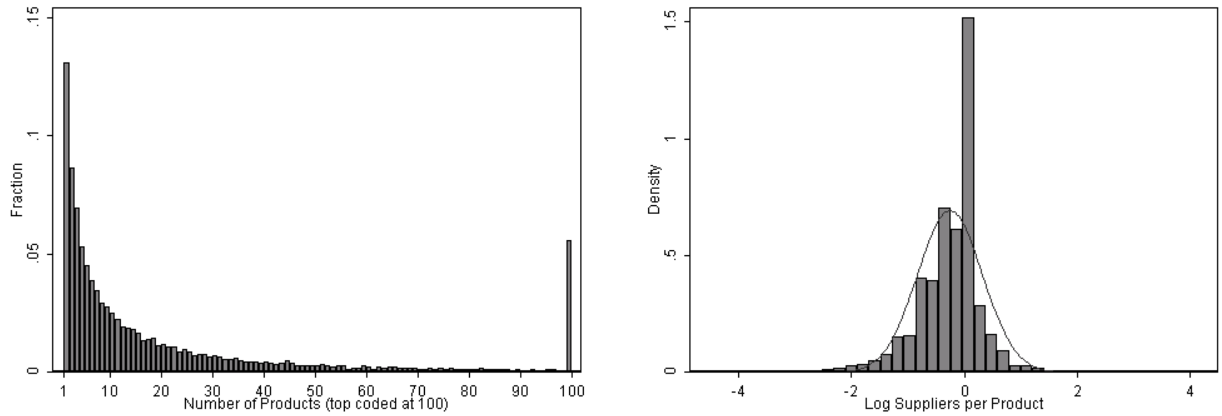
[1] Values in thousands of dollars of imports per firm.
[2] Quasi-medians: means of 10–11 observations around the median.

The firm-level statistics in table 5.2 hide a large degree of heterogeneity across suppliers, products, and countries. Figure 5.1 offers more details on the distributions of products (top panel) and suppliers per product (bottom panel). The modal firm imports one product from one supplier; however, while the product distribution is right-skewed, the distribution of log-suppliers per product is slightly negatively-skewed. Therefore, across sectors the median supplier-per-product ratio is smaller than 1 (in log-scale smaller than zero), suggesting that searching for a supplier might be more costly than searching for a product.

[22] Quasi-median are calculated as the average of 10/11 observations around the true median. This procedure is required to maintain data confidentiality.

[23] Blum et al. (2010) find that Chilean manufacturers import 11.9 HS8 products from 3.2 countries, roughly consistent with our findings.

Figure 5.1: Distribution of number of products and supplier per products



(a) Products imported (top coded at 100)

(b) Log suppliers per product imported

We highlight the geographical distribution of the import network in table 5.3. This table shows the top 10 country of origin for suppliers in 2003 and 2007. The United States is the top source of foreign suppliers in both 2003 and 2007. However, the share of U.S. suppliers decreased from 75% to 69% over the five-year period, with the bulk of the change absorbed by a larger presence of Chinese suppliers. China had already reached the top 2 position in 2003 but consolidated its margin over Germany by 2007. The rest of the distribution remained almost unchanged between 2003 and 2007; only India and Mexico swapped their positions in the ranking.

Table 5.3: Top 10 Country Distribution, 2003 and 2007

2003		2007	
Country	Share of Suppliers	Country	Share of Suppliers
US	74.80%	US	68.53%
China	2.99%	China	7.32%
Germany	2.68%	Germany	3.05%
Italy	2.41%	Italy	2.42%
Great Britain	2.17%	Great Britain	2.07%
Hong Kong	1.68%	Hong Kong	1.98%
Taiwan	1.29%	Taiwan	1.51%
France	1.28%	France	1.36%
India	0.77%	Mexico	0.94%
Mexico	0.76%	India	0.93%

Moving back to the firm-level analysis, we emphasize a dynamic dimension of the import network in the last column of table 5.2, the average age across supplier relationships for a given firm. Martin et al. (2017) suggests that the longer duration of buyer-supplier transactions might be explained by the *specificity* of the relationship, due to the cost of switching to new suppliers. In our data, we set "Age" equal to 0 if a firm starts importing from a particular supplier in a given year and has never imported from the same supplier before; the "Age" variable is equal to 1 if the relationship with the supplier existed in the previous year and so on. In the data we used to build table 5.2, the longest relationships are of age 5. Column (6) reveals that, after a relationship is established, firms tend to keep their suppliers for additional 1.5 years; if we include the initial year in which the relationship is formed, the average duration of buyer-supplier relationships totals 2.5 years.

Figure 5.2: Older relations are less frequent but more valuable

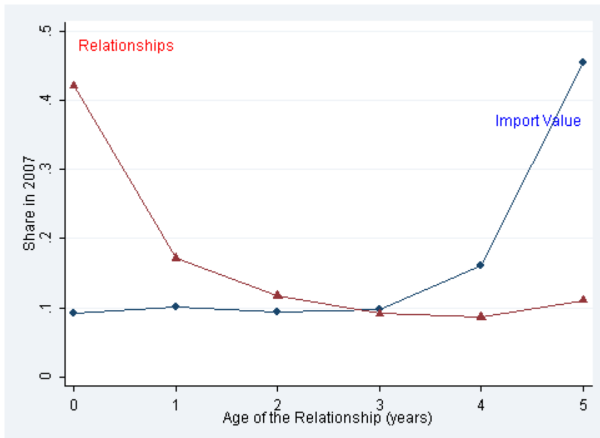


Figure 5.2 explores two characteristics of import relationships along the age dimension; in particular, we look at the number of relationships and the value share over the age distribution of buyer-supplier relationships. Figure 5.2 plots the shares for 2007, where the oldest observed relationship is 5 years. In the appendix, figure C.1 extends our results to the partial year of 2008.[24] Both graphs reveal that older relationships tend to be much less common but much more valuable. Relationships of 5 or more years account for only 10% of the total number of relationships but capture 40% of Canadian firms' total imports. Monarch and Schmidt-Eisenlohr (2016) document a similar finding for U.S. import relationships.

Table 5.4: Import value decomposition by type of relationship

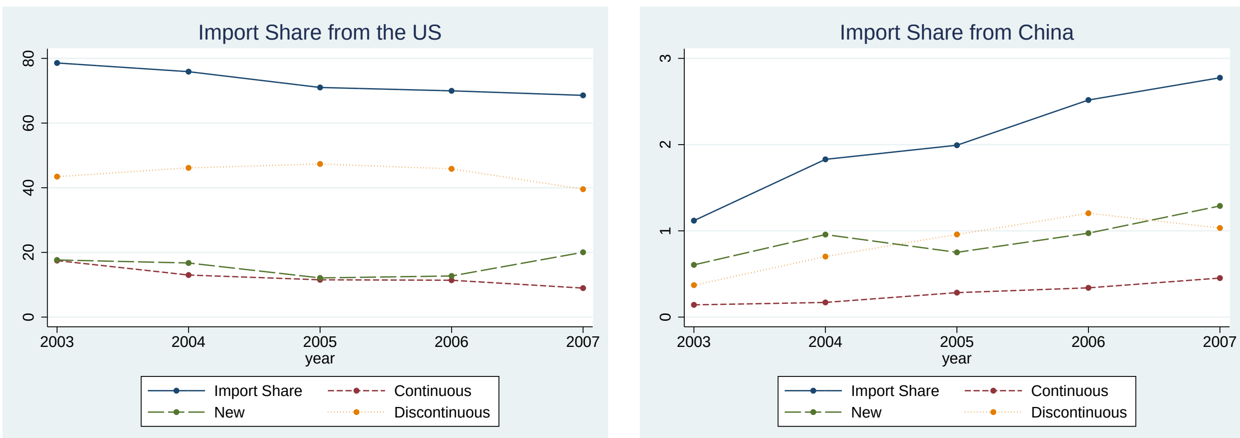
year	Continuous	New	Discontinuous
2003	18.88	24.88	56.24
2004	14.59	24.19	61.22
2005	15.56	22.83	61.61
2006	16.97	20.52	62.52
2007	12.50	31.95	55.55

Table 5.4 looks further into the dynamic import margin. While we rely on the age distribution in our cross-sectional analysis in figures 5.2 and C.1, extending such concept over time would be ardu-

[24]The results are robust across 2007 and 2008; 2007 is our preferred year as the Custom Registry data for 2008 are available only through June.

ous due to changes in the composition of the different age groups over time. Thus, table 5.4 develops a time-series concept of import dynamics by decomposing the imported value across continuous, new, and discontinuous relationships, where relationships are defined at the supplier-product level. A relationship is considered to be "continuous" if a firm imported the same product from the same supplier at least the year before. A relationship is considered "new" if the firm imports a product from a supplier for the first time in a given year (either the firm has never imported the product from that supplier or it has never imported any product from that supplier). We classify all other relationships as "discontinuous". In each year, around one quarter of total imports comes from new suppliers, more than half from discontinuous, and less than 20% from continuous relationships.

Figure 5.3: Decomposition of imports by length of relationship



(a) United States

(b) China

Figure 5.3 applies a similar decomposition to the import shares for the top 2 Canadian partners over 2003 to 2007; the y-axis in both panels indicates the percent of total imports. Overall, the U.S. import share decreased from around 80% in 2003 to 70% in 2007; the Chinese import share, instead, more than doubled over the same period, raising to 2.9% in 2007 from 1% in 2003. Decomposing the import values by type of relationship reveals different patterns in the two countries. While continuous suppliers account for one quarter of the total value imported from the United States, continuous relationships with Chinese suppliers represent a much smaller fraction (around one tenth), with a contribution slowly growing over time. A second point of contrast lies in the contribution of new and discontinuous suppliers: while discontinuous suppliers dominate in U.S.-Canada trade, new Chinese suppliers seem to be as important as discontinuous ones. This fact suggests that Canadian importers tend to experiment more in the Chinese market.

We will now proceed with our investigation of the impact of the breadth and depth of import relationships on firm performance.

5.3 Estimation Framework

In this section, we lay out a simple estimation framework that clearly identifies the conditions under which we can measure the effect of decisions related to the breadth and depth of import relationships. The primary challenge we face is to disentangle the effect of import decisions from that of underlying and unobserved firm productivity. The timing is similar to the one adopted by Kasahara and Rodrigue (2008), which in turn modifies the standard assumptions in Olley and Pakes (1996) and Levinsohn and Petrin (2003) (henceforth, referred to as OP/LP).

Establishment i starts each period t with a stock of capital K_{it} and productivity ω_{it}. It subsequently chooses all variable inputs of production (labor, materials, electricity) and decides next period's capital $K_{i,t+1}$. At this point the firm also makes all decisions relating to importing, like the number of products to be imported and from how many suppliers, which we summarize here by d_{it} and discuss in detail later. The production function in logs is as follows:

$$y_{it} = \beta_0 + \beta_d d_{it} + \beta_l l_{it} + \beta_e e_{it} + \beta_m m_{it} + \beta_k k_{it} + \beta_a \text{Age}_{it} + \omega_{it} + \delta_{st} + \alpha_i + \varepsilon_{it} \tag{5.1}$$

where y_{it}, k_{it}, l_{it}, e_{it}, m_{it} are the logarithm of, respectively, the value of output, capital, labor, electricity and material costs; Age_{it} is the age dummy of firm i and year t; δ_{st} is a sector-time dummy, α_i is the firm fixed effect and ε_{it} is an unexpected shock to firm output after all input and import decisions have been made.

The coefficient of interest throughout this chapter is β_d which measures the effect of importing decisions on output—*holding firm productivity and all other inputs constant.* The main challenge that we face in identifying β_d is the endogeneity of importing decisions, which virtually any model would link to the unobserved productivity shock ω_{it}. To address this issue, we adopt the control function approach in OP/LP. The specific assumption in Levinsohn and Petrin (2003) is that material input choices are a function of capital, age of the firm, and productivity shock ω_{it}. We can therefore write:

$$m_{it} = f(k_{it}, \text{Age}_{it}, \omega_{it}) \tag{5.2}$$

and, under standard monotonicity assumptions, we can invert the function to find ω_{it}:

$$\omega_{it} = f^{-1}(k_{it}, m_{it}, \text{Age}_{it}). \tag{5.3}$$

We can then substitute equation (5.3) into (5.1) and collect all terms for k_{it}, Age_{it} and m_{it} into the function $\varphi(\cdot)$ to obtain

$$y_{it} = \beta_0 + \beta_d d_{it} + \beta_l l_{it} + \beta_e e_{it} + \varphi(k_{it}, m_{it}, \text{Age}_{it}) + \delta_{st} + \alpha_i + \varepsilon_{it}, \tag{5.4}$$

where $\varphi()$ is a second-degree polynomial in capital, age and materials:

$$\varphi\left(k_{it}, m_{it}, \mathrm{Age}_{it}\right) = \beta_1 k_{it} + \beta_2 m_{it} + \beta_3 \mathrm{Age}_{it} + \beta_4 k_{it}^2 + \beta_5 m_{it}^2 + \beta_6 \mathrm{Age}_{it}^2 + \beta_7 k_{it} m_{it} + \beta_8 m_{it} \mathrm{Age}_{it} + \beta_9 \mathrm{Age}_{it} k_{it},$$

and all the β's are 3-digit industry specific parameters.

The OP/LP procedure would then entail a second stage to estimate the capital, material, and age coefficients, but we omit discussion of this portion of the estimation because we are not directly interested in these parameters. The second stage coefficients are industry-specific, so we only report them in the industry-specific regressions shown in Appendix C.2.

As our coefficient of interest is β_d, we now discuss under which conditions this coefficient can be identified. The key condition for identification is that d_{it} is not uniquely determined by the productivity shock ω_{it}. Take the case, for example, in which d_{it} represents the number of suppliers from which the firm imports. Those suppliers could be firms with which firm i already interacted in the past. Alternatively, firm i may choose to establish new relationships with suppliers it never collaborated with. We assume that these decisions entail a fixed cost that may depend on the number of relationships and whether those relationships are established or new but does not depend on the quantity imported. Furthermore, we assume these fixed costs are heterogeneous across firms and not perfectly correlated with the firm's productivity ω_{it}. This type of assumption has become commonplace in the literature that explores various outcomes associated with the export status (see, for example, Helpman et al., 2016) and is typically justified by the fact that, controlling for productivity, various outcomes such as firm-level wages are still correlated with the firm export status. In our context, it is plausible to assume that a firm's TFP does not uniquely determine its fixed cost of establishing and maintaining relationships, a cost which could depend, for example, on the skills of the accounting, purchasing and legal departments of the company. Moreover, those costs could also depend on the history of past relationships, a factor that varies from firm to firm.

The assumption of fixed cost heterogeneity breaks the perfect collinearity that would otherwise arise between all variable inputs and the importing decisions. In this sense, our assumption addresses the concern raised by Ackerberg et al. (2015) (ACF) in the context of production function estimation.[25] To reiterate the point, if we did not make the assumption that heterogeneous fixed costs affected importing decisions, then our coefficient of interest could not be estimated because of the functional dependence problem pointed out by ACF: once we control for all variable input choices, there would be no independent variation left in the choice of d_{it} to estimate β_d. It is worth emphasizing that it does not matter for identification whether the fixed costs of importing are positively or negatively correlated with productivity shock ω_{it} as long as the correlation is not perfect. If material purchases are all made after this productivity shock, then the control function approach will account for ω_{it} and β_d will identify the causal effect of importing on output.

[25] ACF point out that in the OP/LP framework in the absence of further productivity shocks, labour and other variable inputs are perfectly collinear because they are all determined by ω_{it}. This problem prevents the identification of the labour elasticity.

Let us now turn to the different components of d_{it}, a variable that so far has stood in for all importing decisions. In particular, we are going to focus on three sets of variables (for summary statistics see Section 5.2.1):

- Breadth$_{it}$: in this category we include two variables. The first one is ln Products$_{it}$ which represents the variety of imported inputs a firm decides to access. The second variable is the ratio of suppliers to products, $\ln$ Supp/Prod, which measures the number of different suppliers from whom firm i decides to import a given variety.

- Depth$_{it}$: we adopt one variable, the share of continuous relationships, Continuous$_{it}$ to identify the depth of the import network.

- Origin$_{it}$: the variables US Share$_{it}$ and CN Share$_{it}$ measures the degree to which imports by firm i come from the top two source countries, i.e. the United States and China.

To summarize, writing our preferred specification (5.4) in explicit form:

$$y_{it} = \beta_0 + \beta_{d1}\text{Breadth}_{it} + \beta_{d2}\text{Depth}_{it} + \beta_{d3}\text{Origin}_{it} + \text{Input Controls}_{ist} + \delta_{st} + \alpha_i + \varepsilon_{it} \qquad (5.5)$$

where Input Controls$_{ist} \equiv \beta_{s,l} l_{it} + \beta_{s,e} e_{it} + \varphi_{s,t}(k_{it}, m_{it}, \text{Age}_{it})$. Notice that the coefficients in the Input Controls$_{ist}$ function are sector s specific to allow the production function to differ across sectors (3-digit NAICS codes in the regressions). Our coefficients of interest are $(\beta_{d1}, \beta_{d2}, \beta_{d3})$. Ethier (1982) suggests that $\beta_1 > 0$ if the number of HS10 codes and the ratio of suppliers to products induce productivity gains from breaking-up production into multiple stages; a similar mechanism applies to products imported from different countries of origin ($\beta_{d3} > 0$) if those products are imperfect substitutes. Finally, we expect $\beta_{d2} > 0$, that is the share of continuous suppliers to be positively correlated with firm productivity; a positive correlation emerges in Uzzi (1996), which argues that firms within a network benefit from continuing partnerships with their suppliers. We'll explore the source of productivity gains in continuous relationships in more details in section 5.4.3. Our causal interpretation of the results relies on the ability of the input control function, sector-time and firm dummies to capture all factors other than productivity shocks that may simultaneously affect firm importing decisions and sales.

5.4 Results

Table 5.5 shows the results for specification (5.5). Columns (1)–(4) report the coefficients of interest from a restricted version of this specification that excludes the Input Controls. The final column (5) includes Input Controls. The number of imported products and the number of suppliers per product increase firm size with elasticities of 0.15 and 0.12, respectively. The import breadth elasticities drop to 0.03 and 0.02 after controlling for inputs and including the control function. Having continuous relationship with suppliers has also a positive effect on firm productivity; the coefficient on the share of continuous suppliers is positive and significant across all specifications.

The origin of suppliers shows a somewhat unexpected effect on firm sales. While the share of U.S. suppliers has no significant impact on the dependent variable, the share of Chinese suppliers shows a negative and significant coefficient that persists in column (5). One possible explanation to rationalize the negative effect of Chinese suppliers is that firms sourcing from China are aware that their initial supplier draws are likely to be poor, but they expect to find better matches through search and continued experience.

Table 5.5: Firm size and productivity regressions

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: ln Sales				
ln Products	0.154^a	0.154^a	0.153^a	0.153^a	0.030^a
	(0.005)	(0.005)	(0.005)	(0.005)	(0.003)
$\ln \frac{\text{Supp}}{\text{Prod}}$	0.122^a	0.122^a	0.121^a	0.121^a	0.023^a
	(0.007)	(0.007)	(0.007)	(0.007)	(0.003)
Continuous	0.037^a	0.038^a	0.040^a	0.040^a	0.024^a
	(0.008)	(0.008)	(0.008)	(0.008)	(0.004)
US share		0.015		-0.001	0.003
		(0.012)		(0.012)	(0.006)
China share			-0.207^a	-0.207^a	-0.051^c
			(0.044)	(0.045)	(0.024)
Input Controls*	n	n	n	n	y
Firm Fixed Effects	y	y	y	y	y
Sector-Year FEs	y	y	y	y	y
Obs.	93,386	93,386	93,386	93,386	93,386
R^2	0.036	0.036	0.037	0.037	0.717

ln *Products*: log number of imported products (HS10).
$\ln \frac{\text{Supp}}{\text{Prod}}$: log number of foreign suppliers per imported products.
Continuous: share of suppliers from which the buyer purchased for at least the previous year.
US Share: number of U.S. suppliers divided by total foreign suppliers.
CN Share: number of Chinese suppliers divided by total foreign suppliers.
* Input Controls include employment, electricity, and quadratic in capital, materials and age. All controls are also interacted with 3-digit NAICS code dummies.
Notes: Firm FE regression, years 2002–2008. A sector represents a 3-digit NAICS code. Robust standard errors, clustered at the firm level, in parentheses. Significance thresholds are 0.1% (*a*), 1% (*b*), 5% (*c*). The last column implements our preferred specification with input controls as shown in equation (5.5).

The smaller elasticities in column (5) are just what we would expect from a more complete model of the firm's behavior. Suppose an increase in breadth variables lead to 1% productivity improvement. Holding factor prices constant, this should lead to an $\eta - 1$ percent expansion in the value of sales (pq) of the firm, where η is the local (absolute) price elasticity of demand. The ln Products coefficients in columns (4) and (5) are consistent with firm own-elasticities of about six ($\eta - 1 \approx 0.15/0.03$) whereas the corresponding supplier per product elasticities imply $\eta \approx 7$. Both seem on the high side of the values found in the literature but not unreasonably so. In a recent paper, Antràs et al. (2017) report a lower trade elasticity (around 5); their estimate, however, is based on a model that features only the extensive margin of importing at the country level. Thus, with additional (within-country) margins of adjustment at the product and at the supplier level, it is reasonable to expect higher elasticity estimates than in Antràs et al. (2017). [26]

Table 5.6: Summary Statistics for variables used in regressions

	Mean	Std Deviation
Explanatory variables		
ln Products (no. of HS10 imported)	2.18	1.44
ln $\frac{\text{Supp}}{\text{Prod}}$ (suppliers per product)	-0.21	0.57
Continuous share	0.15	0.09
US share	0.73	0.23
CN share	0.16	0.32
Dependent variables		
ln Sales	14.93	1.67
Productivity (Levinsohn-Petrin residuals)	5.37	1.31
ln Exports	13.11	2.71
Export Status	0.62	0.49
ln Number of Destinations	0.72	0.79
ln Exported Products	1.37	1.07

How big are the breadth and depth effects we have estimated in Table 5.5? Perhaps the most natural thought experiment for the breadth effects is to double the number of products or suppliers per product. This would lead to a $2^{0.03}$ =2.1% increase in productivity for doubling products whereas doubling suppliers per product would yield a 1.6% productivity boost. These effects seem somewhat modest. Raising the Continuous share from 0 to 100% would lead to a 2.4% productivity improvement. These hypothetical shocks may not be considered realistic. Another popular way to quantify results is to express them in terms of standard deviations of the explanatory variables. Using Table 5.6 to obtain the standard deviations, we see that a one-standard-deviation increase in ln Products implies a productivity gain by 2.6% of a standard deviation (sd); a one-standard-deviation increase in the number of suppliers, keeping the product margin constant, improves productivity by 0.8% of a sd. Continuous relationship are also associated with small productivity

[26] When disentangling the "micro" elasticity of substitution among alternative suppliers from the "macro" elasticity of substitution between domestic and foreign suppliers, Feenstra et al. (2017) find that micro elasticity estimates tends to be larger than macro estimates.

gains: a one-standard-deviation increase in Continuous raises firm productivity by 0.1% of a sd. The coefficient on the share of Chinese suppliers implies, instead, a sizable negative effect on productivity: a one-standard-deviation increase in the share of Chinese suppliers is associated with a 1% of a sd drop in productivity.

The effects that we document are smaller than the firm-level productivity gains documented by Amiti and Konings (2007) and Topalova and Khandelwal (2011) (12% in the case of Indonesia, 4.8% for India for a 10% reduction in input tariffs); however, while the estimates in those papers reveal the aggregate effect on productivity, our estimates aim at identifying specific channels for the realization of those gains.

5.4.1 Robustness and sectoral estimates

The additional results in section C.2 show that the panel fixed effects results are mainly robust when we instead estimate the regressions in long differences. Table C.3 considers the variation in sales between 2003 and 2007. One notable difference is that the we no longer obtain negative effects of the Chinese share on productivity (after controlling for the U.S. share and inputs). While input variety and dynamic variables remain positive and significant with similar magnitudes to those documented in Table 5.5, the negative sign on the share of Chinese suppliers fades in the specification with the full set of controls (column (5)).

Tables C.4–C.6 and C.7 in Appendix C.2 show the results when we estimate the productivity specification (5) for each sector. The first three tables show Levinsohn-Petrin estimates of the breadth, depth, and country-of-origin effects along with the four factor input elasticities. These regressions include the second-stage coefficients for regressions based on the same identifying assumption as presented in Table 5.5. Table C.7, instead, uses the Olley-Pakes approach in which investment is part of the control function. This approach requires us to drop firms with zero investment which accounts for the sample attrition. Levinsohn and Petrin (2003) motivate their method in part by warning that such attrition could be non-random. In general, we do not detect systematic differences. Often the coefficients are very similar, but the higher standard errors in OP lead to less statistically significant results. For example, Transport Equipment has a typical product breath elasticity of 0.029 in the LP specification with a standard error of 0.011 (Table C.6). In the OP version, shown in Table C.7, the coefficient is 0.027 with a standard error of 0.015. Overall the LP and OP results both support the near ubiquity of productivity gains from importing more variety.

Importing more products has a positive impact on productivity across all industries; the coefficient is significant in most cases. The product import margin seems to be particularly relevant in industries using larger share of differentiated inputs (e.g., Computing, Transportation Equipment and Machinery). Conditioning on the number of imported products, the supplier margin is also associated with a significant productivity increase in about one third of all sectors. The productivity effect of additional suppliers seems to be particularly relevant in Metals and Metallic Products, and across other industries making larger use of homogeneous inputs.

Continuous relationships with suppliers tend to have a positive impact on productivity across all sectors; the effect is significant only in Computing, Paper, Apparels, Metallic Products and Chemicals. As for the countries of origin, the share of U.S. suppliers does not display any effect on productivity; the sign of the coefficient on US Share varies across sectors although the variable is never significant. The share of Chinese suppliers, instead, tend to be associated with lower productivity in sectors with larger Chinese penetration (Apparel and Other Manufacturing Activities); however, firms in Textiles and Petrol that have more Chinese suppliers tend to have bigger sales, controlling for input usage.

The input elasticities reported in Tables C.4-C.6 and C.7 are in line with the estimates by Halpern et al. (2015). In particular, they find that the capital share in production is around 0.04, which is equal to our average capital share estimate across sectors. Moreover, while their labour elasticity estimate (0.2) is in line with our results, their share of materials (0.75) appear significantly larger than ours. We believe that this difference may be due to the fact that we separately control for electricity.

5.4.2 Impact of import relationships on export performance

Table 5.7 investigates how the characteristics of the import network affect export performance. Past research has shown that the majority of firms do not export and, among the exporters, the modal firm exports a single product to a single destination.[27] In standard models of heterogeneous firms, more productive firms can cover fixed costs associated with exporting. Thus, to the extent that our breadth and depth variables trigger productivity gains, we expect them to raise export performance. We consider 4 measures of export performance: total exports (columns 1 and 2), the number of products (HS8 codes) exported (columns 3 and 4), whether a firm exports to any country (5 and 6), and the number of export destinations (7 and 8). We set the number of destinations equal to 1 for non-exporters (this can be thought of as home as the first destination). The even-numbered columns adopt a specification similar to column (5) of Table 5.5, where we add controls for inputs and age and the LP quadratic function.

We find that firms importing more products from more suppliers are more likely to be exporters, export more, and sell more products to more destinations. The imported product and supplier elasticities imply similar magnitudes for the effects on performance. Considering the coefficient on ln Products, a one-standard-deviation increase in the number of imported products increases exports by 15% of a sd, raises the number of exported products by 14% of a sd, increase the number of export destination by 7% of a sd and increases the probability of exporting by 3 percentage points.

[27] See Bernard et al. (2007) for the United States and Mayer and Ottaviano (2007) for some European countries.

Table 5.7: How import relationships affect export performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln Exports		ln Exp. Products		Export Status		ln Destinations	
ln Products	0.253^a	0.298^a	0.110^a	0.105^a	0.026^a	0.024^a	0.069^a	0.038^a
	(0.016)	(0.020)	(0.007)	(0.009)	(0.003)	(0.003)	(0.003)	(0.004)
ln $\frac{\text{Supp}}{\text{Prod}}$	0.265^a	0.260^a	0.082^a	0.070^a	0.029^a	0.026^a	0.061^a	0.040^a
	(0.025)	(0.025)	(0.011)	(0.011)	(0.004)	(0.005)	(0.006)	(0.006)
Continuous	-0.207^a	0.050	0.063^a	0.006	-0.017^b	-0.012	0.079^a	-0.011
	(0.036)	(0.041)	(0.015)	(0.017)	(0.006)	(0.007)	(0.007)	(0.007)
US share	0.100^c	0.057	-0.033	-0.014	0.006	0.006	-0.029^b	-0.015
	(0.050)	(0.048)	(0.022)	(0.021)	(0.009)	(0.009)	(0.011)	(0.011)
CN share	-0.701^a	-0.381^c	-0.268^a	-0.280^a	-0.043	-0.024	-0.045	-0.066^c
	(0.199)	(0.189)	(0.068)	(0.069)	(0.033)	(0.033)	(0.031)	(0.031)
Input Controls*	n	y	n	y	n	y	n	y
Firm FE	y	y	y	y	y	y	y	y
Sector-Year	y	y	y	y	y	y	y	y
Obs.	44,939	44,939	44,939	44,939	67,184	67,184	67,184	67,184
R^2	0.017	0.091	0.012	0.052	0.003	0.032	0.012	0.059

ln *Products*: log number of imported products (HS10).
ln $\frac{\text{Supp}}{\text{Prod}}$: log number of foreign suppliers per imported products.
Continuous: share of suppliers from which the buyer purchased for at least the previous year.
US Share: number of U.S. suppliers divided by total foreign suppliers.
CN Share: number of Chinese suppliers divided by total foreign suppliers.
* Input Controls include employment, electricity, and quadratic in capital, materials and age. All controls are also interacted with 3-digit NAICS code dummies.
Notes: Firm FE regression, years 2002–2008. A sector represents a 3-digit NAICS code. Robust standard errors, clustered at the firm level, in parentheses. Significance thresholds are 0.1% (*a*), 1% (*b*), 5% (*c*). The even-numbered columns implement our preferred specification with input controls as shown in equation (5.5).

Neither the share of continuous relationships nor the U.S. import share has a robust effect on export outcomes. We find that Chinese suppliers tend to have a negative impact on export performance. Having more Chinese suppliers is associated with lower exports, fewer exported products, and fewer destinations; the coefficient on the likelihood of becoming an exporter is negative but not significant. The surprisingly negative effect of relationship with Chinese suppliers on total exports are quite big. Consider a firm that goes from 0% Chinese suppliers to 100% Chinese suppliers. The column 1 coefficient of -0.7 implies that its exports will fall by half ($\exp(-0.7) = 0.496$). This is, of course, a radical and unrealistic change but even looking at one standard deviation changes, we find big effects from increased usage of Chinese suppliers. A one-standard-deviation larger share of Chinese suppliers reduces exports by 4.5% of a sd, lowers the number of exported products by 8.4% of a sd and the number of export destination by 2.7% of a sd.

5.4.3 The Dynamics of Import Relationships

What is the source of the productivity gains arising in continuous relationships? A possible explanation is that the buyers and suppliers in continuous relationships tend to exchange products better *tailored* to the production process of the buyer. In order to provide support to this mechanism, we estimate a specification relating the type of relationship to import outcomes,

$$\text{Import Outcome}_{ijpt} = \beta_0 + \beta_1 \cdot \text{Relationship Type}_{ijpt} + D_{pt} + \varepsilon_{ijpt} \tag{5.6}$$

The dependent variable is either the import value, the quantity imported, or the unit value in the transaction of product p between firm i and supplier j at time t. Relationship Type_{ijpt} includes continuous, new, and discontinuous relationships. We also consider how *unique* relationships—supplier-product combinations that are linked to a unique buyer—are related to import outcomes. It is possible that when a Canadian firm is the only buyer of a foreign product, it is because that product has been customized for that firm and that such customization might be reflected in the price paid for the imported product. The excluded category covers buyer-supplier-product relationships that are discontinuous and not unique. The specification also includes HS2 dummies, unit of measure dummies, as well as 3-digit NAICS-year dummies.

Whether we rely on Uzzi's idea of embeddedness or on a model with search and matching, we expect similar predictions. In fact, following Uzzi (1996), a firm embedded in a production network would have longer-lasting relationships and better-customized products. Similarly, in a framework in which searching for a trade partner is costly and agents' learn about their partner's productivity over time, better matches tend to last longer and generate larger surplus, which translates into larger pay-offs for all participants in the relationships. In particular, we expect that firms in continuous relationships tend to import larger values, not only because of bigger quantities, but also because they pay higher unit values.

Table 5.8: Import Relationships

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln Import Value		ln Imp. Val.		ln Imp. Quant.		ln Unit Value	
Continuous	1.121^a	0.116^a	0.317^a	0.039^a	1.107^a	0.079^a	-0.002	0.017^a
	(0.028)	(0.005)	(0.013)	(0.004)	(0.031)	(0.006)	(0.016)	(0.004)
New	-0.229^a	-0.159^a	0.008	-0.038^a	-0.325^a	-0.180^a	0.102^a	0.010^c
	(0.030)	(0.005)	(0.016)	(0.004)	(0.031)	(0.007)	(0.017)	(0.005)
Unique	-0.409^a	0.093^a	-0.149^a	0.027^a	-0.397^a	0.076^a	-0.034^c	0.007
	(0.023)	(0.005)	(0.011)	(0.004)	(0.029)	(0.006)	(0.014)	(0.004)
ln Quantity	n	n	y	y	n	n	n	n
Rel. FE	n	y	n	y	n	y	n	y
Sector×Year	y	y	y	y	y	y	y	y
Observations	5.5mn	5.5mn	3mn	3mn	3mn	3mn	3mn	3mn
R^2	0.164	0.144	0.677	0.668	0.343	0.107	0.505	0.012

Continuous: dummy equal to one if a firm imported the same product from the same supplier at $t-1$.
New: dummy equal to one if a firm imports a product from a supplier for the first time.
Unique: dummy equal to one if a supplier sells a product only to one firm at t.
Notes: The odd-numbered columns report pooled OLS regressions, while the even-numbered columns report relationship (defined as firm-product-supplier dummies) fixed-effect regressions. In all columns, we also control for log sales, log export, HS2 product dummies, and dummies for the unit of measurement. A sector stands for a 3-digit NAICS code. Robust standard errors, clustered at the firm level, in parentheses. Significance thresholds are 0.1% (a), 1% (b), 5% (c).

Table 5.8 reports the OLS and Relationship FE regression results for specification (5.6).[28] All specifications control for the characteristics of the Canadian firm in its output market, i.e. the log of total sales and the log of total exports, so that we can compare firms with equal sales that adopt different strategies regarding the duration or exclusivity of their relationships. Firms in continuous relationships import larger values than in discontinuous connections; the effect on value comes both from larger quantities and higher unit values (columns (3)-(4) and (8)). New relationships, instead, involve lower import values; this outcome seems to be primarily a quantity rather than a price effect. Evidently, buyers are reluctant to place large orders from firms they have no prior experience with.

Finally, let us consider the behavior of unique supplier-product combinations. Exploiting both the cross-sectional and time variation, unique relationships seem to be associated with lower import values, resulting both from lower quantities and lower unit values; however, suppliers becoming the unique provider of a certain good (columns (2), (4), (6) and (8)) export larger values, larger quantities and sell their products at a higher unit value (the coefficient on Unique in column (8) is positive but not significant). We believe that our dummy for unique relationships captures attempts

[28]We include D_{pij} fixed effects in the even numbered columns of table 5.8.

of buyers to find the best inputs compatible with their production process.

5.5 Conclusion

In this chapter, we have explored the productivity effects of the breadth and depth of firms' import relationships. With the caveat that our identification strategy relies on the control function approach to partial out unobserved productivity shocks, we find significant and economically relevant breadth effects. Both the number of varieties imported and the number of suppliers per variety raise productivity. These results support the theoretical foundation in Ethier (1982) and are consistent with a wider literature in which we see that reductions in the costs of imported inputs (via tariff cuts or changes in transport access) lead to productivity improvements. These results on breadth have many other counterparts in the literature on gains from variety in final consumer goods.

We also find novel and promising effects of import relationship depth. The share of continuous importing relationships the firm is engaged also appears to raise firm performance. In addition, we find that firms engaged in continuous import relationships with the same suppliers systematically feature transactions that are larger and, to a lesser extent, have higher value. We are not aware of any model that can fully explain these findings, but we hypothesize that it could be the result of a search and matching process whereby only the most successful matches survive. Only a firm's best supplier relationships carry on and because they are better matches, they take up a larger share of the firm's total imports. We have only laid out a possible theoretical interpretation of these novel results, but we are optimistic that they could help a better understanding of where the productivity gains of importing come from. They come not only from wider variety of inputs, but also from a deeper pool of suppliers in which the firm can find an *ideal* partner.

Our results point to several important policy implications. First, import tariff reductions on intermediate inputs are likely to help Canadian productivity and boost the performance of Canadian firms in international markets. This is consistent with evidence from less developed countries but was not previously known for a country like Canada with a well-developed manufacturing sector. Secondly, since the United States provides the majority of the suppliers used by Canadian firms, it would be helpful to shrink the fixed costs of adding and maintaining suppliers. It is not obvious how to achieve that but travel and visa facilitation are probably valuable. There may also be gains from harmonization of technical standards. The most general policy implication of all is that even if trade policy makers are focused on export markets, they should not neglect that Canadian firms' success in selling abroad is very much predicated upon their ability to use a broad and deep roster of foreign suppliers.

Chapter 6

Conclusion

In this dissertation, we study the impact of two important aspects of the importing market on firm performance: import competition and the buyer-supplier relationship.

In studying the effect of import competition on firm performance (Chapters 2-4), we found a robust empirical relationship for China during the period around the WTO accession, that an increase in import competition would raise innovation among the most productive firms. While the effect is not significantly different from zero for the less productive firms. We develop a model with monopolistic competition as in Melitz and Ottaviano (2008) across varieties, and neck-and-neck competition within varieties. The model stresses the two opposing forces that drive firm's innovation incentives when there is a change in import competition: the escape-competition effect and the rent-reduction effect. When competition increases, the positive escape-competition effect dominates for the more productive firms. In the aggregate economy, if the escape-competition effect dominates overall, there will be an additional gains from trade under a unilateral trade liberalization.

Our empirical study is among the first that focus on the relationship between import competition and innovation in the context of a developing country. Our model is the first to combine the neck-and-neck competition in a classic trade framework. It aims to isolate the competition channel from technology diffusion or the market size effect, and suggest that, a unilateral trade liberalization may be harmful to domestic innovation if the home firms are too lagged behind and competition would only bring about rent destruction. In the case of China, it seems at least for the more productive firms, they were well equipped with the potential to improve when China entered the WTO in 2001. The next step in our study is to calibrate the model to the Chinese data and calculate the model-implied welfare changes.

In Chapter 5, we explore how the structures of import relationships could affect the performance of Canadian importers. Under the control-function identification strategy, we find that the breadth and depth of firms' import relationships have significant and economically relevant effects on firm performance. Firms grow bigger and more productive if they import a wider variety of products, source from more suppliers in each product, and was able to find relationship that could last longer.

In the past twenty years, the trade literature has focused much on the gains from exporting or symmetric trade liberalization. In this dissertation, we put our focus on the importing market, and showed empirically and theoretically, that even a unilateral trade liberalization can bring gains, especially when firms are encouraged to innovate more, and are facilitated to form better import relationships.

Bibliography

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

Aghion, P., A. Bergeaud, M. Lequien, and M. Melitz (2017). The impact of exports on innovation: Theory and evidence. Technical report, working paper.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). Competition and Innovation: an Inverted-U Relationship. *The Quarterly Journal of Economics 120*(2), 701–728.

Aghion, P., R. Blundell, R. Griffith, P. Howitt, and S. Prantl (2009). The effects of entry on incumbent innovation and productivity. *Review of Economics and Statistics 91*(1), 20–32.

Aghion, P., J. Cai, M. Dewatripont, L. Du, A. Harrison, and P. Legros (2015). Industrial policy and competition. *American Economic Journal: Macroeconomics 7*(4), 1–32.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001). Competition, imitation and growth with step-by-step innovation. *Review of Economic Studies 68*(3), 467–492.

Akcigit, U., S. T. Ates, and G. Impullitti (2017). Innovation and trade policy in a globalized world. Working paper.

Akcigit, U., J. Grigsby, and T. Nicholas (2017). The rise of american ingenuity: Innovation and inventors of the golden age. Technical report, National Bureau of Economic Research.

Amiti, M. and J. Konings (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *The American Economic Review*, 1611–1638.

Ang, J. S., Y. Cheng, and C. Wu (2014). Does enforcement of intellectual property rights matter in China? Evidence from financing and investment choices in the high-tech industry. *Review of Economics and Statistics 96*(2), 332–348.

Antràs, P., T. C. Fort, and F. Tintelnot (2017). The margins of global sourcing: theory and evidence from US firms. *American Economic Review 107*(9), 2514–2564.

Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *Staff Papers-International Monetary Fund*, 159–178.

Arora, A., S. Belenzon, and A. Patacconi (2015). Killing the golden goose? The decline of science in corporate R&D. Working paper, National Bureau of Economic Research.

Atkeson, A. and A. T. Burstein (2010). Innovation, firm dynamics, and international trade. *Journal of Political Economy 118*(3), 433–484.

Autor, D., D. Dorn, G. H. Hanson, G. Pisano, and P. Shu (2017). Foreign competition and domestic innovation: Evidence from US patents. Working paper.

Aw, B. Y., M. J. Roberts, and D. Yi Xu (2011). R&D investment, exporting, and productivity dynamics. *American Economic Review 101*(4), 1312–1344.

Bernard, A. B., E. J. Blanchard, I. Van Beveren, and H. Y. Vandenbussche (2017). Carry-along trade. *Review of Economic Studies* (forthcoming).

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003). Plants and productivity in international trade. *American Economic Review 93*(4), 1268–1290.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007). Firms in International Trade. *Journal of Economic Perspectives 21*(3), 105–130.

Bernard, A. B., A. Moxnes, and Y. U. Saito (2017). Production networks, geography and firm performance. *Journal of Political Economy* (forthcoming).

Bernard, A. B., A. Moxnes, and K. H. Ulltveit-Moe (2017). Two-sided heterogeneity and trade. *The Review of Economics and Statistics* (forthcoming).

Blaum, J., C. Lelarge, and M. Peters (2017). The gains from input trade with heterogeneous importers. *AEJ Macro* (forthcoming).

Bloom, N., M. Draca, and J. Van Reenen (2016). Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. *Review of Economic Studies 83*(1), 87–117.

Blum, B. S., S. Claro, and I. Horstmann (2010). Facts and figures on intermediated trade. *The American Economic Review*, 419–423.

Blundell, R. and J. L. Powell (2003). Endogeneity in nonparametric and semiparametric regression models. *Econometric Society Monographs 36*, 312–357.

Bøler, E. A., A. Moxnes, and K. H. Ulltveit-Moe (2015). R&D, international sourcing, and the joint impact on firm performance. *American Economic Review 105*(12), 3704–3739.

Bombardini, M., G. Orefice, and M. D. Tito (2015). Does exporting improve matching? evidence from french employer-employee data. Technical report, National Bureau of Economic Research.

Brandt, L. and P. M. Morrow (2016). Tariffs and the organization of trade in China. Working paper.

Brandt, L., J. Van Biesebroeck, L. Wang, and Y. Zhang (2017). WTO accession and performance of Chinese manufacturing firms. *American Economic Review 107*(9), 2784–2820.

Brandt, L., J. Van Biesebroeck, and Y. Zhang (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics 97*(2), 339–351.

Broda, C. and D. E. Weinstein (2006). Globalization and the gains from variety. *The Quarterly journal of economics 121*(2), 541–585.

Buera, F. J. and E. Oberfield (2017). The global diffusion of ideas. Working paper.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms. *American economic review 101*(1), 304–340.

Carballo, J., G. I. Ottaviano, and C. Volpe Martincus (2013). The buyer margins of firms' exports. *CEPR Discussion Paper No. DP9584*.

Castellani, D., F. Serti, and C. Tomasi (2010). Firms in international trade: Importers' and exporters' heterogeneity in Italian manufacturing industry. *The World Economy 33*(3), 424–457.

Coe, D. T. and E. Helpman (1995). International R&D spillovers. *European Economic Review 39*(5), 859–887.

De Loecker, J. and F. Warzynski (2012). Markups and firm-level export status. *American Economic Review 102*(6), 2437–71.

Demidova, S. (2017). Trade policies, firm heterogeneity, and variable markups. *Journal of International Economics 108*, 260–273.

Dragusanu, R. (2014). Firm-to-firm matching along the global supply chain. Technical report, Citeseer.

Eaton, J., M. Eslava, D. Jinkins, C. J. Krizan, and J. Tybout (2015). A search and learning model of export dynamics.

Eaton, J. and S. Kortum (1999). International technology diffusion: Theory and measurement. *International Economic Review 40*(3), 537–570.

Ethier, W. J. (1982). National and international returns to scale in the modern theory of international trade. *The American Economic Review*, 389–405.

Fang, L. H., J. Lerner, and C. Wu (2017). Intellectual property rights protection, ownership, and innovation: Evidence from China. *Review of Financial Studies 30*(7), 2446–2477.

Feenstra, R. C. (2010). *Product variety and the gains from international trade.* MIT Press Cambridge, MA.

Feenstra, R. C., P. A. Luck, M. Obstfeld, and K. N. Russ (2017). In search of the armington elasticity. *Review of Ecoomics and Statistics*.

Fieler, A. C., M. Eslava, and D. Y. Xu (2016). Trade, quality upgrading, and input linkages: Theory and evidence from Colombia.

Fieler, A. C. and A. Harrison (2018). Escaping import competition and downstream tariffs. Technical report.

Gilbert, R. (2006). Looking for Mr. Schumpeter: Where are we in the competition–innovation debate? *Innovation Policy and the Economy 6*, 159–215.

Goldberg, P., A. K. Khandelwal, N. Pavcnik, and P. B. Topalova (2009). Trade liberalization and new imported inputs. In *American Economic Review, Papers and Proceedings*, Volume 99, pp. 494–500.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature 28*(4), 1661–1707.

Grossman, G. M. and E. Helpman (1991). Quality ladders in the theory of growth. *Review of Economic Studies 58*(1), 43–61.

Hall, B., C. Helmers, M. Rogers, and V. Sena (2014). The choice between formal and informal intellectual property: A review. *Journal of Economic Literature 52*(2), 375–423.

Halpern, L., M. Koren, and A. Szeidl (2015). Imported inputs and productivity. *American Economic Review 105*(12).

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. J. Redding (2016). Trade and inequality: From theory to estimation. *The Review of Economic Studies*, 357–405.

Holmes, T. J., E. R. McGrattan, and E. C. Prescott (2015). Quid pro quo: technology capital transfers for market access in China. *Review of Economic Studies*, rdv008.

Hsieh, C.-T., P. J. Klenow, and I. Nath (2018). A global view of creative destruction.

Hsieh, C.-T. and Z. M. Song (2015). Grasp the large, let go of the small: The transformation of the state sector in China. Working paper, National Bureau of Economic Research.

Hu, A. G. and G. H. Jefferson (2009). A great wall of patents: What is behind China's recent patent explosion? *Journal of Development Economics 90*(1), 57–68.

Hu, A. G., G. H. Jefferson, and Q. Jinchang (2005). R&D and technology transfer: Firm-level evidence from Chinese industry. *Review of Economics and Statistics 87*(4), 780–786.

Jiang, K., W. Keller, L. D. Qiu, and W. Ridley (2018). International joint ventures and internal vs. external technology transfer: Evidence from china. Technical report, National Bureau of Economic Research.

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

Kasahara, H. and B. Lapham (2013). Productivity and the decision to import and export: Theory and evidence. *Journal of International Economics 89*(2), 297–316.

Kasahara, H. and J. Rodrigue (2008). Does the use of imported intermediates increase productivity? Plant-level evidence. *Journal of development economics 87*(1), 106–118.

Khandelwal, A. K., P. K. Schott, and S.-J. Wei (2013). Trade liberalization and embedded institutional reform: Evidence from Chinese exporters. *American Economic Review 103*(6), 2169–2195.

Kugler, M. and E. Verhoogen (2009). Plants and imported inputs: New facts and an interpretation. *The American Economic Review*, 501–507.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies 70*(2), 317–341.

Lim, K., D. Trefler, and M. Yu (2017). Trade and innovation: The role of scale and competition effects. Work in progress.

Lu, D., A. Mariscal, and L.-F. Mejia (2016). How firms accumulate inputs: Evidence from import switching. In *University of Rochester Working Paper*.

Lu, Y. and L. Yu (2015). Trade liberalization and markup dispersion: Evidence from China's WTO accession. *American Economic Journal: Applied Economics 7*(4), 221–253.

MacGarvie, M. (2006). Do firms learn from international trade? *Review of Economics and Statistics 88*(1), 46–60.

Martin, J., I. Méjean, and M. Parenti (2017). Relationship specificity: Measurement and applications to trade.

Mayer, T. and G. Ottaviano (2007). *The Happy Few: The Internationalisation of European firms.* Bruegel Blueprint Series.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Melitz, M. J. and G. I. Ottaviano (2008). Market size, trade, and productivity. *Review of Economic Studies 75*(1), 295–316.

Miroudot, S., R. Lanz, and A. Ragoussis (2009). Trade in intermediate goods and services. *OECD Trade Policy Papers* (93).

Monarch, R. (2014). "It's not you, it's me": Breakups in US-China trade relationships. *US Census Bureau Center for Economic Studies Paper No. CES-WP-14-08*.

Monarch, R. and T. Schmidt-Eisenlohr (2016). Learning and the value of relationships in international trade.

Moser, P. (2013). Patents and innovation: Evidence from economic history. *Journal of Economic Perspectives 27*(1), 23–44.

Muûls, M. and M. Pisu (2009). Imports and exports at the level of the firm: Evidence from Belgium. *The World Economy 32*(5), 692–734.

Nagaoka, S., K. Motohashi, and A. Goto (2010). Patent statistics as an innovation indicator. *Handbook of the Economics of Innovation 2*, 1083–1127.

Olley, S. and A. Pakes (1996). The dynamics of productivity in the telecomunications equipment industry. *Econometrica 64*, 1263–97.

Ossa, R. (2011). A "new trade" theory of GATT/WTO negotiations. *Journal of Political Economy 119*(1), 122–152.

Ottaviano, G., T. Tabuchi, and J.-F. Thisse (2002). Agglomeration and trade revisited. *International Economic Review*, 409–435.

Perla, J., C. Tonetti, and M. E. Waugh (2015). Equilibrium technology diffusion, trade, and growth. Working paper, National Bureau of Economic Research.

Qian, Y. (2007). Do national patent laws stimulate domestic innovation in a global patenting environment? a cross-country analysis of pharmaceutical patent protection, 1978–2002. *Review of Economics and Statistics 89*(3), 436–453.

Ramanarayanan, A. (2017). Imported inputs and the gains from trade.

Rivera-Batiz, L. A. and P. M. Romer (1991). Economic integration and endogenous growth. *The Quarterly Journal of Economics 106*(2), 531–555.

Teshima, K. (2008). Import competition and innovation at the plant level: Evidence from Mexico.

Topalova, P. and A. Khandelwal (2011). Trade liberalization and firm productivity: The case of India. *Review of economics and statistics 93*(3), 995–1009.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American sociological review*, 674–698.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative science quarterly*, 35–67.

Uzzi, B. (1999). Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *American sociological review*, 481–505.

Wei, S.-J., Z. Xie, and X. Zhang (2017). From "made in China" to "innovated in China": Necessity, prospect, and challenges. *Journal of Economic Perspectives 31*(1), 49–70.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT press.

Xie, Z. and X. Zhang (2015). The patterns of patents in China. *China Economic Journal 8*(2), 122–142.

Appendix A

Appendix to Chapter 2

A.1 Productivity Estimation

In this section, we describe our production function estimation procedure and report the estimated coefficients.

We assume that the final output of a firm is produced by labor, capital, and intermediate material input following a Cobb-Douglass production function.

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \beta_m m_t + \omega_t + \varepsilon_t \tag{A.1}$$

where y_t, l_t, k_t and m_t denote the log gross output, labor, capital and material inputs, respectively.

We deal with the classical endogeneity problem, that the unobserved productivity ω is correlated with inputs, following the two-step control function procedure in Levinsohn and Petrin (2003) and Ackerberg et al. (2015). Incur the usual assumptions of the control function approach as stated in Ackerberg et al. (2015): (i) productivity shocks ω_{it} are First Order Markov, (ii) both labor and capital are quasi-fixed, (iii) firm's investment or intermediate input decisions are only determined by one unobservable that is ω_{it}, and (iv) the decision function is invertible. We can re-write the production function:

The first stage involves regressing the log gross output on the same set of fourth order polynomials as in the value added specification, where the higher orders are already transformed into the Chebyshev polynomials.

$$y_t = \phi_t(k_t, l_t, m_t, i_t) + \varepsilon_t \tag{A.2}$$

$$y_t = \delta_0 + \sum_{i=0}^{4} \sum_{j=0}^{4-i} \sum_{r=0}^{4-i-j} \sum_{z=0}^{4-i-j-r} \delta_{ij} k_t^i m_t^j l_t^r \mathrm{inv}_t^z + \varepsilon_t \tag{A.3}$$

In the first stage, for all approaches, we estimate $\hat{\phi}_{it}$ as a fourth order polynomial in $\tilde{k}_{it}, \tilde{l}_{it}, \tilde{m}_{it}, \widetilde{\mathrm{inv}}_{it}$, where inv_{it} denotes log investment, and $\tilde{x} = \frac{x-\mu_x}{\sigma_x}$ denotes the normalized variables. To ensure orthogonality of the polynomial basis, we also apply the Chebyshev polynomial transformation[29].

[29] The Chebyshev polynomial transformation $T(\cdot)$ acts as follows:

$$T_0(x) = 1$$
$$T_1(x) = x$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

In the second stage, the law of motion for productivity is approximated by an AR(1) process:

$$\hat{\omega}_t = \rho_0 + \rho_1 \omega_{t-1} + \eta_t \tag{A.4}$$

Substitute into the production function,

$$\hat{\phi}_t = \beta_l l_t + \beta_k k_t + \beta_m m_t + \rho_0 + \rho_1(\hat{\phi}_{t-1} - \beta_l l_{it} - \beta_k k_{t-1} - \beta_m m_{t-1}) + \eta_t \tag{A.5}$$

We use the NLS method to estimate the second stage equation, which is equivalent to the following moment condition:

$$E\left[\begin{pmatrix} l_t - \rho_1 l_{t-1} \\ k_t - \rho_1 k_{t-1} \\ m_t - \rho_1 m_{t-1} \\ \hat{\phi}_{t-1} - \beta_l l_{it} - \beta_k k_{t-1} - \beta_m m_{t-1} \\ 1 \end{pmatrix} \times \eta_t\right] = 0 \tag{A.6}$$

Table A.1 shows the estimation results. Columns(1)-(3) shows the coefficient estimates. The output elasticity of intermediate good is averaged around 0.81, being as high as 0.91 in the Chemical industry (code 28) and as low as 0.68 in the furniture manufacturing industry (code 21). Column (4) shows the estimate of the auto regressive coefficient for the productivity shock process. The average estimate is 0.66. Out of the 29 industries, 18 industries have the auto regressive coefficient above 0.6. The electric equipment and machinery industry (code 39) has the lowest persistence parameter (0.28) which implies the firms are more subject to transitory shocks. Column (5) shows the returns to scale. With the gross output production function specification, the return to scale estimate is quite reasonable, being very close to 1. Column (6) shows the correlation between the estimated TFP from the ACF approach and a simple OLS regression within each two digit sector. The average sectoral correlation is 0.98. The aggregate correlation is 0.86.

Alternatively, we can also formulate the production function using value-added output, or using the translog specification[30].

[30] The production function for the value-added output is

$$v_t = \beta_0 + \beta_l l_t + \beta_k k_t + \omega_t + \varepsilon_t$$

And the production function for the translog specification is

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \beta_m m_t + \beta_{ll}\tilde{l}_t^2 + \beta_{kk}\tilde{k}_t^2 + \beta_{mm}\tilde{m}_t^2 + \beta_{lk}\tilde{l}_t\tilde{k}_t + \beta lm\tilde{l}_t\tilde{m}_t + \beta_{km}\tilde{k}_t\tilde{m}_t + \beta_{lkm}\tilde{l}_t\tilde{k}_t\tilde{m}_t + \omega_t + \varepsilon_t$$

Table A.1: Production function estimation coefficients by sector

code	industry name	(1) l	(2) k	(3) m	(4) ρ_1	(5) rts	(6) corr	(7) no. obs
13	Food Processing	0.09	0.09	0.80	0.61	0.99	0.97	32,407
14	Food Production	0.07	0.08	0.90	0.67	1.05	0.97	13,980
15	Beverage Production	0.09	0.13	0.83	0.78	1.05	0.93	9,705
16	Tobacco Industry	0.06	0.14	0.81	0.78	1.01	0.93	1,063
17	Textile Industry	0.07	0.06	0.84	0.58	0.98	0.98	51,961
18	Garment and Other Fiber Products	0.12	0.08	0.76	0.66	0.96	0.98	28,549
19	Leather, Furs, Down & Related Products	0.10	0.06	0.81	0.49	0.98	0.99	13,991
20	Timber and Bamboo Processing	0.14	0.10	0.75	0.81	0.98	0.96	10,489
21	Furniture Manufacturing	0.15	0.08	0.68	0.89	0.91	0.91	6,395
22	Papermaking and Paper Products	0.08	0.08	0.79	0.79	0.95	0.93	19,298
23	Printing & Record Medium Reproduction	0.12	0.20	0.80	0.79	1.12	0.88	13,210
24	Cultural, Educational & Sports Goods	0.12	0.06	0.79	0.48	0.97	0.99	8,415
25	Petroleum refining and Coking	0.03	0.05	0.86	0.80	0.94	0.99	5,412
26	Raw Chemical materials/Products	0.05	0.06	0.87	0.45	0.97	1.00	46,224
27	Medical and Pharmaceutical Products	0.10	0.12	0.78	0.64	1.00	0.96	15,152
28	Chemical Fiber	0.03	0.03	0.91	0.71	0.97	0.99	2,989
29	Rubber Products	0.10	0.11	0.74	0.61	0.96	0.94	7,274
30	Plastic Products	0.12	0.11	0.71	0.70	0.95	0.94	28,045
31	Nonmetal Mineral Products	0.09	0.09	0.78	0.86	0.95	0.90	52,202
32	Ferrous Metal Mining and Dressing	0.05	0.04	0.88	0.59	0.97	0.99	13,492
33	Nonferrous Metal Mining and Dressing	0.08	0.04	0.82	0.87	0.94	0.97	10,543
34	Metal Products	0.07	0.07	0.84	0.36	0.97	0.99	31,068
35	Ordinary Machinery	0.05	0.07	0.87	0.52	0.98	0.99	46,172
36	Special Purposes Equipment	0.03	0.08	0.88	0.56	0.99	0.98	23,440
37	Transport Equipment	0.08	0.10	0.84	0.56	1.02	0.98	30,134
39	Electric Equipment and Machinery	0.06	0.06	0.87	0.28	0.99	1.00	37,231
40	Electronic and Telecommunications	0.13	0.10	0.76	0.80	0.99	0.96	20,960
41	Instruments, Cultural act Machinery	0.09	0.08	0.80	0.87	0.96	0.96	8,607
42	Other Manufacturing	0.08	0.04	0.85	0.57	0.98	0.99	10,993

Columns(1)-(3) shows the coefficient estimates for the Cobb-Douglas gross output production function estimation using the ACF method. Column (4) shows the estimate of the autoregressive coefficient for the productivity shock process. Column (5) shows the returns to scale. Column (6) shows the correlation between the estimated TFP from the ACF approach and a simple OLS regression within each two digit sector. The overall correlation is 0.86.

Table A.2 shows the correlation between our baseline estimation and the two alternative specifications. The correlation between the gross output (GO) and translog (TL) production function is quite high. The average correlation within the 29 two-digit industries is 0.83, while this average correlation is 0.66 between the GO and VA approaches, and 0.54 between the TL and VA approaches. The main reason that the VA approach has lower correlation with the other two approaches could be that the output measure is log value added, which is different from the gross output in the other two methods.

Table A.2: Correlation between estimated productivity, level

code	industry name	(1) GO vs. TL	(2) GO vs. VA	(3) TL vs. VA	(4) obs
13	Food Processing	0.87	0.67	0.49	118,852
14	Food Production	0.95	0.45	0.27	47,434
15	Beverage Production	0.84	0.61	0.45	32,974
16	Tobacco Industry	0.94	0.72	0.53	2,459
17	Textile Industry	0.81	0.66	0.39	163,459
18	Garment and Other Fiber Products	0.80	0.74	0.49	92,560
19	Leather, Furs, Down and Related Products	0.78	0.69	0.43	45,916
20	Timber and Bamboo Processing	0.91	0.75	0.61	42,224
21	Furniture Manufacturing	0.79	0.83	0.70	22,326
22	Paper making and Paper Products	0.81	0.77	0.57	57,825
23	Printing & Record Medium Reproduction	0.89	0.56	0.41	40,348
24	Cultural, Educational & Sports Goods	0.76	0.70	0.46	25,274
25	Petroleum refining and Coking	0.90	0.54	0.32	17,410
26	Raw Chemical materials/Products	0.85	0.56	0.28	140,679
27	Medical and Pharmaceutical Products	0.83	0.73	0.56	40,574
28	Chemical Fiber	0.80	0.62	0.24	9,798
29	Rubber Products	0.56	0.76	0.56	22,970
30	Plastic Products	0.81	0.77	0.60	89,491
31	Nonmetal Mineral Products	0.86	0.78	0.62	169,019
32	Ferrous Metal Mining and Dressing	0.85	0.56	0.27	46,650
33	Nonferrous Metal Mining and Dressing	0.92	0.56	0.35	35,027
34	Metal Products	0.69	0.61	0.36	103,632
35	Ordinary Machinery	0.78	0.59	0.35	146,181
36	Special Purposes Equipment	0.85	0.61	0.37	80,205
37	Transport Equipment	0.84	0.61	0.38	92,060
39	Electric Equipment and Machinery	0.78	0.52	0.25	113,064
40	Electronic and Telecommunications	0.83	0.72	0.52	61,852
41	Instruments, Cultural act Machinery	0.85	0.70	0.51	26,548
42	Other Manufacturing	0.89	0.62	0.32	37,732
	minimum	0.56	0.45	0.24	
	maximum	0.95	0.83	0.70	
	mean	0.83	0.66	0.44	
	aggregate correlation	0.77	0.37	0.28	

A.2 Robustness

Table A.3: Output tariff and patenting, interaction with quartiles

	(1)	(2)	(3)	(4)
Dep. var: Patent application counts				
Output competition				
$\tau^{output}_{s,t-2} \times$ Top Quartile$_{is,t-2}$	-4.027**	-3.366*	-2.414	-2.827*
	(2.018)	(1.996)	(2.239)	(1.452)
$\tau^{output}_{s,t-2} \times$ 3rd Quartile$_{is,t-2}$	-0.825	-0.304	0.659	0.234
	(1.854)	(1.860)	(2.058)	(1.272)
$\tau^{output}_{s,t-2} \times$ 2nd Quartile$_{is,t-2}$	0.452	1.175	1.997	1.678*
	(1.646)	(1.647)	(1.988)	(0.964)
$\tau^{output}_{s,t-2} \times$ Bottom Quartile$_{is,t-2}$	-1.376	-0.306	0.079	0.002
	(1.795)	(1.808)	(1.942)	(1.602)
Export control				
$D^{exporter}_{ist-2}$		1.278***		0.909***
		(0.333)		(0.329)
$E^{demand}_{s,t-2}$		0.063***		0.066***
		(0.021)		(0.021)
$E^{demand}_{s,t-2} \times D^{exporter}_{ist-2}$		0.026		0.023
		(0.025)		(0.025)
Import control				
$D^{importer}_{ist-2}$			1.799***	0.984***
			(0.188)	(0.168)
$\tau^{input}_{is,t-2}$			-0.336	-4.936
			(11.268)	(10.972)
$\tau^{input}_{is,t-2} \times D^{importer}_{ist-2}$			-11.314***	-7.385**
			(3.848)	(3.639)
obs	800,292	800,292	800,292	800,292

Notes: The Top dummy equals to 1 if the firm is above the 75th percentile productivity in industry s at time $t-2$. All columns control for four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. The four sectors are: chemicals and petroleum, computers and electronics, machinery and equipment sector, and others. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Industry specification controlling for FDI

	(1)	(2)	(3)
Dep. var: Patent application counts			
FDI variable	ln(foreign equity)	ln(foreign_HMT equity)	foreign equity share
Output competition			
$\tau_{s,t-2}^{output} \times Top_{is,t-2}$	-3.521**	-3.549**	-3.542**
	(1.477)	(1.466)	(1.464)
$\tau_{s,t-2}^{output}$	2.904	2.367	1.709
	(1.946)	(1.929)	(1.936)
$Top_{is,t-2}$	1.207***	1.208***	1.206***
	(0.147)	(0.147)	(0.147)
FDI control			
$FDI_{s,t-2}$	-0.093*	-0.065	0.247
	(0.049)	(0.049)	(0.340)
obs	797,958	799,335	800,276

Notes: The Top dummy equals to 1 if the firm is above the 75th percentile productivity in industry s at time $t-2$. All columns control for four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. The four sectors are: chemicals and petroleum, computers and electronics, machinery and equipment sector, and others. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Two-stage control function estimation

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var: Patent application counts						
Specification	OLS	CF	OLS	CF	OLS	CF
Output competition						
$\ln M^{output}_{s,t-2} \times Top_{is,t-2}$	0.102**	0.121***	0.107***	0.126***	0.104**	0.112**
	(0.040)	(0.045)	(0.039)	(0.045)	(0.042)	(0.047)
$\ln M^{output}_{s,t-2}$	0.023	0.476	0.035	0.443	0.018	0.219
	(0.025)	(0.322)	(0.028)	(0.318)	(0.025)	(0.370)
$Top_{is,t-2}$	0.768***	0.750***	0.787***	0.770***	0.740***	0.749***
	(0.088)	(0.090)	(0.087)	(0.089)	(0.088)	(0.090)
Export control						
$\mathrm{D}^{\mathrm{exporter}}_{ist-2}$			1.586***	1.588***		
			(0.102)	(0.102)		
$\ln X^{exp}_{s,t-2}$			-0.184***	0.034		
			(0.046)	(0.066)		
$\ln X^{export}_{s,t-2} \times D^{exporter}_{ist-2}$			0.206***	0.217***		
			(0.041)	(0.044)		
Import control						
$\mathrm{D}^{\mathrm{importer}}_{ist-2}$					1.756***	1.731***
					(0.250)	(0.251)
$\ln M^{input}_{is,t-2}$					-4.439	0.392
					(11.949)	(14.453)
$\ln M^{input}_{is,t-2} \times D^{importer}_{ist-2}$					-11.713**	-11.607**
					(5.161)	(5.202)
First Stage						
Endogenous var.		$\ln M^{output}_{s,t-2}$		$\ln X^{export}_{s,t-2}$		$\ln M^{input}_{s,t-2}$
Instruments	$\tau^{output}_{s,t-2}$	-5.544***	$\tau^{export}_{s,t-2}$	0.406***	$\tau^{input}_{s,t-2}$	-12.975***
		(1.641)		(0.021)		(2.299)
obs	802,410	802,410	802,410	802,410	802,410	792,963

The even columns show results using the control function approach (CF). It is a two-stage procedure. The First stage regresses the endogenous variables on the respective instruments. The second stage regresses patent counts on the endogenous variables, controlling for the error terms from the first stage.
The Top dummy equals to 1 if the firm is above 75 percentile in industry s at time $t-2$.
All columns control for four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. The four sectors are: chemicals and petroleum, computers and electronics, machinery and equipment sector, and others.
Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Alternative measures of innovation

Dep var	(1) Granted application *Poisson*	(2) Citation weighted *Poisson*	(3) Patent dummy *OLS*	(4) Patent count *OLS*	(5) $\ln(Pat+1)$ *OLS*
Output competition					
$\tau^{output}_{s,t-2} \times Top_{is,t-2}$	-2.750*	-2.836**	-0.057***	-0.917***	-0.096***
	(1.483)	(1.303)	(0.008)	(0.229)	(0.014)
$\tau^{output}_{s,t-2}$	2.732	-0.454	-0.019	-0.002	-0.020
	(2.370)	(2.094)	(0.015)	(0.167)	(0.021)
$Top_{is,t-2}$	1.126***	1.048***	0.012***	0.165***	0.019***
	(0.142)	(0.122)	(0.001)	(0.032)	(0.002)
Export control					
$\text{D}^{\text{exporter}}_{ist-2}$	0.928***	0.875***	0.042***	0.293***	0.055***
	(0.159)	(0.141)	(0.003)	(0.047)	(0.004)
$E^{demand}_{s,t-2}$	1.477**	1.001	0.015**	0.238***	0.029***
	(0.626)	(0.639)	(0.007)	(0.071)	(0.009)
$E^{demand}_{s,t-2} \times D^{exporter}_{ist-2}$	-0.227	-1.143	-0.275***	-2.063***	-0.374***
	(1.954)	(1.679)	(0.025)	(0.373)	(0.035)
Import control					
$\text{D}^{\text{importer}}_{ist-2}$	1.599***	1.415***	0.049***	0.273***	0.063***
	(0.284)	(0.240)	(0.004)	(0.047)	(0.005)
$\tau^{input}_{is,t-2}$	-0.114	0.120	0.240***	0.569	0.257***
	(19.535)	(13.923)	(0.046)	(0.459)	(0.062)
$\tau^{input}_{is,t-2} \times D^{importer}_{ist-2}$	-16.220**	-11.554**	-0.556***	-3.599***	-0.739***
	(6.997)	(5.686)	(0.077)	(0.911)	(0.105)
R^2			0.038	0.008	0.037
obs	762,640	770,316	770,978	770,978	770,978

Notes: The Top dummy equals to 1 if the firm is above 75 percentile in industry s at time $t-2$. All columns control for four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. The four sectors are: chemicals and petroleum, computers and electronics, machinery and equipment sector, and others. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1

Table A.7: Drop 2-digit sectors, one at a time

	(1)	(2)	(3)	(4)	(5)
Dep. var: Patent application counts					
Sector Dropped	**Food**	**Drinks**	**Tobacco**	**Textile**	**Furniture**
firm $\tau_{s,t-2}^{output} \times Top_{is,t-2}$	-2.417*	-2.653*	-2.943**	-2.919**	-2.890**
	(1.415)	(1.416)	(1.412)	(1.387)	(1.362)
firm $\tau_{s,t-2}^{output}$	-1.095	-0.723	0.377	1.183	-0.465
	(2.071)	(2.218)	(2.231)	(2.025)	(1.965)
$Top_{is,t-2}$	1.099***	1.104***	1.117***	1.112***	1.116***
	(0.133)	(0.133)	(0.132)	(0.131)	(0.131)
obs	738657	790877	803282	681173	778623
Sector Dropped	**Paper**	**Chemical**	**Stone**	**Metal**	**Machinery**
firm $\tau_{s,t-2}^{output} \times Top_{is,t-2}$	-2.852**	-3.217**	-2.874**	-2.656*	-3.151**
	(1.377)	(1.400)	(1.383)	(1.428)	(1.334)
firm $\tau_{s,t-2}^{output}$	0.117	-1.629	-0.466	-0.102	-0.272
	(2.062)	(2.051)	(1.999)	(1.939)	(2.077)
$Top_{is,t-2}$	1.108***	1.212***	1.120***	1.038***	1.226***
	(0.133)	(0.150)	(0.131)	(0.140)	(0.134)
obs	752471	665091	732278	724141	692330
Sector Dropped	**Transportation**	**Computers**	**Other**		
firm $\tau_{s,t-2}^{output} \times Top_{is,t-2}$	-2.083	-4.202***	-2.884**		
	(1.585)	(1.181)	(1.371)		
firm $\tau_{s,t-2}^{output}$	-1.360	-0.687	-0.372		
	(2.112)	(1.656)	(1.960)		
$Top_{is,t-2}$	1.097***	0.985***	1.114***		
	(0.142)	(0.128)	(0.131)		
obs	764506	731702	791453		

Notes: The Top dummy equals to 1 if the firm is above 75 percentile in industry s at time $t-2$. All columns include the export and import controls, the four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. The four sectors are: chemicals and petroleum, computers and electronics, machinery and equipment sector, and others. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: Firm output tariff and patenting, OLS

	(1)	(2)	(3)	(4)
Dep. var: Patent application counts				
Output competition				
firm $\tau^{output}_{s,t-2} \times Top_{is,t-2}$	-3.276***	-3.381***	-4.851***	-5.049***
	(0.922)	(0.953)	(1.557)	(1.605)
firm $\tau^{output}_{s,t-2}$	0.027	0.084	0.180	0.294
	(0.248)	(0.265)	(0.445)	(0.475)
$Top_{is,t-2}$	0.573***	0.586***	0.828***	0.852***
	(0.139)	(0.143)	(0.224)	(0.230)
Export control				
firm $\tau^{export}_{is,t-2}$		0.024***		0.036***
		(0.006)		(0.011)
Import control				
firm$\tau^{input}_{is,t-2}$			0.320	0.338
			(0.598)	(0.612)
R^2	0.017	0.018	0.019	0.020
obs	141,823	136,031	75,460	73,139

Notes: OLS regression instead of Poisson. The Top dummy equals to 1 if the firm is above 75 percentile in industry s at time $t-2$. All columns control for four-digit CIC industry fixed effects, as well as four-sector by year fixed effects. Standard errors are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1

Appendix B

Appendix to Chapter 4

Proof of Lemma 2. The partial derivative of utility with respect to foreign entry rate e at given innovation I is

$$\begin{aligned}\frac{\partial u}{\partial e} &= I\left[\theta\left(\pi_1 + \mathrm{CS}_m\left(\delta c\right)\right) + \left(1-\theta\right)\mathrm{CS}_m\left(c_0^*\right) - \pi\left(\delta c\right) - \mathrm{CS}_m\left(\delta c\right)\right] \\ &\quad + \left(1-I\right)\left[\theta \mathrm{CS}_l\left(c\right) + \left(1-\theta\right)\mathrm{CS}_m\left(c_0^*\right) - \pi\left(c\right) - \mathrm{CS}_m\left(c\right)\right] \\ &= I\left(1-\theta\right)\Delta_1 + \left(1-I\right)\Delta_0\end{aligned}$$

where

$$\begin{aligned}\Delta_1 &\equiv \mathrm{CS}_m\left(c_0^*\right) - \mathrm{CS}_m\left(\delta c\right) - \pi\left(\delta c\right) \\ \Delta_0 &\equiv \theta \mathrm{CS}_l\left(c\right) + \left(1-\theta\right)\mathrm{CS}_m\left(c_0^*\right) - \mathrm{CS}_m\left(c\right) - \pi\left(c\right) \\ \mathrm{CS}_m\left(c\right) &= \frac{\left(\alpha - c\right)^2}{8} \\ \mathrm{CS}_l\left(c\right) &= \frac{\left(\alpha - c\right)^2}{2} \\ \pi\left(c\right) &= \frac{\left(\alpha - c\right)^2}{4}\end{aligned}$$

The expression for firm profits $\pi\left(\cdot\right)$ incurs Assumption 2 and 4, which ensure that as long as the firm is producing, it could obtain the monopoly profit.

To prove that $\frac{\partial u}{\partial e} > 0$, it suffices to show that $\Delta_1 > 0$ and $\Delta_0 > 0$.

For the proportion of Home firms that succeeded in innovating, the marginal gain from foreign entry for consumers, $\mathrm{CS}_m\left(c_0^*\right) - \mathrm{CS}_m\left(\delta c\right)$, is always larger than the marginal loss in profits for Home firms. To show this, from Assumption 3,

$$\begin{aligned}&2\alpha + c_0^* < \alpha + 2\delta c \\ \Rightarrow &2\alpha - 2\delta c < \alpha - c_0^* \\ \Rightarrow &4\left(\alpha - \delta c\right)^2 < \left(\alpha - c_0^*\right)^2 \\ \Rightarrow &\frac{1}{8}\left[\left(\alpha - c_0^*\right)^2 - 3\left(\alpha - \delta c\right)^2\right] > 0 \\ \Rightarrow &\Delta_1 = \mathrm{CS}_m\left(c_0^*\right) - \mathrm{CS}_m\left(\delta c\right) - \pi\left(\delta c\right) > 0 \qquad \text{(B.1)}\end{aligned}$$

Next, we show that for the proportion of Home firms that did not innovate, the marginal gain for consumers is also larger than the marginal loss of producers due to foreign entry.

$$
\begin{aligned}
&\mathrm{CS}_l\,(c) - \mathrm{CS}_m\,(c) - \pi\,(c) \\
=&\frac{(\alpha - c)^2}{8} > 0 \\
&\mathrm{CS}_m\,(c_0^*) - \mathrm{CS}_m\,(c) - \pi\,(c) \\
>&\mathrm{CS}_m\,(c_0^*) - \mathrm{CS}_m\,(\delta c) - \pi\,(\delta c) > 0
\end{aligned}
$$

where the last line comes from equation (B.1). Therefore,

$$\Delta_0 > 0.$$

□

Appendix C

Appendix to Chapter 5

C.1 Coding Supplier Identifiers

Transaction records are collected from Form B3 of the Canadian Border Service Agency. Importers are required to report the vendors' name on the form among the other information. The vendor's name is transformed into a consistent identifier according to a procedure articulated into 3 steps.[31] The first step creates the basic vendor identifier according to the following stages:

1. Remove stop words, like *ltd*, *corp*, *inc* etc.; we will refer to the output of this stage as the *standard name*.
2. Remove punctuation but leave spaces into the vendor's name; this generates the *clean name*.
3. Replace French characters with English characters.
4. Remove other irrelevant words not integrated in the vendor's name, e.g. *and*, *the*, *of*, a, etc.
5. Remove vowels from the name.
6. Assign the basic vendor identifier.

The second stage of the procedure tries to propagate identifiers across records likely to represent the same firm:

- Generate a second identifier using the first two words of the clean name, if the first two words are not blank and standard name contains at least 6 characters. Firms whose name has the same first and second words are assigned the same identifier.
- Construct a third identifier based on the clean name, if the first non-blank word does not contain more than 16 characters.
- Generate a fourth identifier based on the first 3 words from the vendor's name.
- Construct a fifth identifier based on the ZIP code and the first three words of the vendor's name.
- Generate a sixth identifier attributed to vendors exporting to the same Canadian firm the same product and with the same first word.

[31]The matching algorithm was developed by Statistics Canada employees, inspired by the SIMILE project procedure.

The second identifier is selected as the preferred identifier; if such identifier could not be created, the third identifier would be used and so on. Finally, the third stage constructs a measure to characterize the quality of the identifiers. The quality is measured over 9 levels:[32]

- Level 0 is assigned if the vendor's name and its address are consistent across observations carrying the same identifier.
- Level 1 is assigned if the clean name and the address are consistent across observations carrying the same identifier.
- Level 2 is assigned if the vendor's name is consistent across observations carrying the same identifier.
- Level 3 is assigned if the clean vendor's name is consistent across observations carrying the same identifier.
- Level 4 is assigned if the distance between the vendor's and the clean name normalized by their length is less than 10, the first word and the address match across observations carrying the same identifier.
- Level 5 is assigned if the normalized distance between the names is less than 6, the basic identifier and the first word match across observations carrying the same identifier.
- Level 6 is assigned if the normalized distance between the names is less than 6, the Canadian Business Number and the HS10 product-code imported from the vendor match across observations carrying the same identifier.
- Level 7 is assigned if the normalized distance between the names is less than 3.
- Level 8 is assigned if the normalized distance between the names is less than 10.

Let us work through an example. Consider three fictional vendor's names

- Great Oranges and Nuts, Corporation
- Great Oranges and Néwton
- Great Oranges

Following the first steps of the algorithm, we would be able to generate the basic identifiers

1. Remove Corp./ Corporations
    - Great Oranges and Nuts,
    - Great Oranges and Néwton

[32] The presence of a match quality indicator is very important as it allows to run robustness checks over groups of different match quality.

    - Great Oranges

2. Remove Punctuation
    - Great Oranges and Nuts
    - Great Oranges and Néwton
    - Great Oranges
3. Remove French Characters
    - Great Oranges and Nuts
    - Great Oranges and Newton
    - Great Oranges
4. Remove stop words
    - Great Oranges Nuts
    - Great Oranges Newton
    - Great Oranges
5. Remove Vowels
    - Grt Orngs Nts
    - Grt Orngs Nwtn
    - Grt Orngs
6. Assign the vendor basic identifier
    - 123
    - 456
    - 789

Following the second step of the procedure, preferred identifiers are based on the matching the first two words of the clean vendor's name.

- 123
- 456
- 123

In the third step firms with equal identifiers from the second step are assigned a measure of the quality of the match. In our example, the two firms with identifier 123 have a match quality of 4 if the address is the same. In case the two observations do not share the same address, the match quality would be 8.

C.2 Supplemental Empirical Results

Table C.1: Average Market Share by sector, 2002–2008

NAICS	Industry	Domestic Mkt Share
311	Food	20.91%
312	Bev. & Tob.	1.80%
313	Text. Mills	0.44%
314	Text. Prod.	0.32%
315	Apparel	0.64%
316	Leather	0.07%
321	Wood	4.00%
322	Paper	4.91%
323	Printing	1.41%
324	Petrol	9.42%
325	Chemical	7.02%
326	Plastics	3.83%
327	Mineral	1.88%
331	Metals	6.82%
332	Met. Prod.	4.06%
333	Machinery	4.12%
334	Computing	2.90%
335	Electrical	1.54%
336	Trans. Eq.	21.30%
337	Furniture	1.54%
339	Miscel.	1.01%

Table C.2: Summary Statistics from Import Registry

Variable	Mean	Std Deviation
ln Import Value	8.07	2.80
ln Unit Value	3.41	2.55
Continuous (Indicator)	0.30	0.46
Unique	0.78	0.42
New	0.65	0.48

Figure C.1: Relationship age in extended sample ending in June 2008

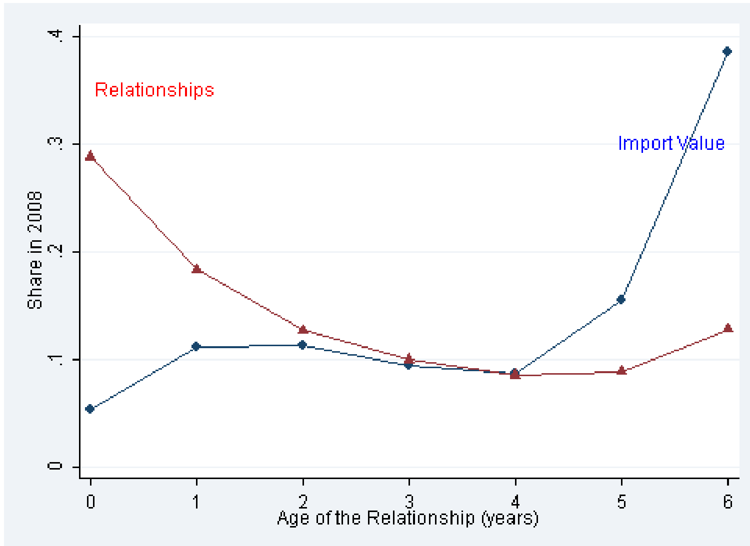


Table C.3: Long-difference (2003–2007) estimates

Variable	(1)	(2)	(3) ln Sales	(4)	(5)
ln Products	0.278[a]	0.280[a]	0.278[a]	0.279[a]	0.046[a]
	(0.013)	(0.013)	(0.013)	(0.013)	(0.007)
$\ln \frac{\text{Supp}}{\text{Prod}}$	0.175[a]	0.174[a]	0.174[a]	0.173[a]	0.025[b]
	(0.018)	(0.018)	(0.018)	(0.018)	(0.008)
Continuous	0.141[a]	0.151[a]	0.148[a]	0.154[a]	0.028[c]
	(0.025)	(0.026)	(0.025)	(0.026)	(0.014)
US share		0.089[b]		0.067	0.032
		(0.033)		(0.035)	(0.018)
CN share			-0.280[c]	-0.220	0.043
			(0.132)	(0.138)	(0.053)
Input Controls	n	n	n	n	y
Firm Effects	y	y	y	y	y
Sector Effects	y	y	y	y	y
Obs.	28,769	28,769	28,769	28,769	28,769
R^2	0.085	0.086	0.086	0.087	0.787

ln Products: log number of imported products (HS10).
ln Supp/Prod: log number of suppliers per imported products.
Continuous: share of suppliers from which the buyer purchased for at least the previous year.
US Share: share of U.S. suppliers.
CN Share: share of Chinese suppliers.
Notes: Long-difference regressions, years 2003 and 2007. The last column implements the first stage of a specification à la Levinsohn and Petrin (2003); see equation (5.4).

Table C.4: Sales Regressions by Sector (NAICS 31)

	ln Sales					
Variable	Food	Bev.	Text. M	Text. P	App.	Leath.
ln Products	0.017^c	0.010	0.052^b	0.027	0.035^a	0.028
	(0.007)	(0.034)	(0.016)	(0.015)	(0.011)	(0.024)
$\ln \frac{\text{Supp}}{\text{Prod}}$	0.028^b	0.002	-0.006	0.025	0.026	-0.020
	(0.009)	(0.041)	(0.025)	(0.020)	(0.014)	(0.032)
Continuous	0.023	-0.068	0.072	0.032	0.066^c	-0.042
	(0.014)	(0.090)	(0.038)	(0.028)	(0.026)	(0.044)
US share	-0.005	-0.053	0.071	-0.000	0.023	0.018
	(0.015)	(0.077)	(0.038)	(0.036)	(0.027)	(0.058)
CN share	-0.019	-0.058	0.275^c	-0.118	-0.092^c	0.070
	(0.081)	(0.331)	(0.122)	(0.066)	(0.042)	(0.215)
Log Empl	0.161^a	0.151^a	0.300^a	0.254^a	0.297^a	0.178^a
	(0.008)	(0.040)	(0.020)	(0.017)	(0.014)	(0.024)
Log Elect	0.120^a	0.239^a	0.074^a	0.064^a	0.081^a	0.038^c
	(0.007)	(0.034)	(0.013)	(0.013)	(0.012)	(0.017)
Log K	0.052	-0.015	0.068	0.084	-0.059	0.116^a
	(0.020)	(0.090)	(0.032)	(0.047)	(0.031)	(0.034)
Log Mat	0.561^a	0.265^a	0.443^a	0.379^a	0.509^a	0.492^a
	(0.043)	(0.049)	(0.034)	(0.035)	(0.024)	(0.039)
Obs.	6,462	572	1,156	1,699	3,887	709
R^2	0.737	0.643	0.769	0.668	0.714	0.729

ln Products: log number of imported products (HS10).
ln Supp/Prod: log number of suppliers per imported products.
Continuous: share of suppliers from which the buyer purchased for at least the previous year.
US Share: share of U.S. suppliers.
CN Share: share of Chinese suppliers.
Notes: Firm FE regression, years 2002–2008. Sector represents 3-digit NAICS. Robust standard errors, clustered at the firm level, in parenthesis. Significance thresholds are 0.1% (a), 1% (b), 5% (c). Each column implements a two-stage Levinsohn-Petrin specification.

Table C.5: Sales Regressions by Sector (NAICS 32)

	ln Sales						
Variable	Wood	Paper	Print.	Oil	Chem.	Plast.	Min.
ln Products	0.027[a]	0.033[b]	0.026[a]	0.032	0.059[a]	0.015[c]	0.018
	(0.007)	(0.010)	(0.006)	(0.030)	(0.011)	(0.006)	(0.011)
ln $\frac{\text{Supp}}{\text{Prod}}$	0.009	0.046[b]	0.011	0.106[c]	0.040[c]	0.008	0.016
	(0.009)	(0.014)	(0.009)	(0.051)	(0.018)	(0.009)	(0.015)
Continuous	0.003	0.050[c]	0.010	0.043	0.062[c]	0.014	0.011
	(0.013)	(0.023)	(0.013)	(0.081)	(0.026)	(0.013)	(0.021)
US share	0.006	-0.018	-0.014	0.016	0.020	-0.027	0.024
	(0.020)	(0.027)	(0.016)	(0.105)	(0.032)	(0.015)	(0.029)
CN share	-0.046	-0.000	-0.064	4.436[c]	-0.024	-0.031	0.153
	(0.066)	(0.201)	(0.062)	(1.974)	(0.120)	(0.051)	(0.079)
Log Empl	0.258[a]	0.228[a]	0.358[a]	0.104[a]	0.213[a]	0.294[a]	0.333[a]
	(0.010)	(0.012)	(0.012)	(0.026)	(0.011)	(0.009)	(0.013)
Log Elect	0.120[a]	0.095[a]	0.143[a]	0.017	0.093[a]	0.117[a]	0.097[a]
	(0.007)	(0.007)	(0.010)	(0.021)	(0.007)	(0.006)	(0.007)
Log K	0.031	0.017	0.028	0.105[a]	0.024	0.021	0.019
	(0.033)	(0.010)	(0.023)	(0.041)	(0.060)	(0.017)	(0.016)
Log Mat	0.557[a]	0.595[a]	0.416[a]	0.577[a]	0.338[a]	0.507[a]	0.522[a]
	(0.024)	(0.016)	(0.031)	(0.033)	(0.044)	(0.016)	(0.026)
Obs.	4,719	1,914	4,178	353	4,654	6,276	3,381
R^2	0.817	0.892	0.769	0.716	0.612	0.797	0.775

ln Products: log number of imported products (HS10).
ln Supp/Prod: log number of suppliers per imported products.
Continuous: share of suppliers from which the buyer purchased for at least the previous year.
US Share: share of U.S. suppliers.
CN Share: share of Chinese suppliers.
Notes: Firm FE regression, years 2002–2008. Sector represents 3-digit NAICS. Robust standard errors, clustered at the firm level, in parenthesis. Significance thresholds are 0.1% (*a*), 1% (*b*), 5% (*c*). Each column implements a two-stage Levinsohn-Petrin specification.

Table C.6: Sales Regressions by Sector (NAICS 33)

	ln Sales							
Variable	Met.	Met. P.	Mach.	Comp.	Elect.	Tr. Eq.	Furn.	Misc.
ln Products	0.073^a	0.024^a	0.018^b	0.098^a	0.007	0.029^c	0.020^b	0.036^a
	(0.016)	(0.005)	(0.006)	(0.015)	(0.013)	(0.011)	(0.006)	(0.007)
$\ln \frac{\text{Supp}}{\text{Prod}}$	0.086^a	0.025^a	0.015	0.035	0.012	0.039^c	0.003	0.030^b
	(0.025)	(0.007)	(0.008)	(0.020)	(0.018)	(0.016)	(0.009)	(0.009)
Continuous	0.044	0.021^c	0.011	0.090^c	0.033	0.043	-0.007	0.028^c
	(0.034)	(0.009)	(0.013)	(0.039)	(0.029)	(0.026)	(0.013)	(0.014)
US share	0.061	0.003	0.014	-0.026	0.005	0.006	0.005	0.000
	(0.041)	(0.013)	(0.016)	(0.039)	(0.032)	(0.039)	(0.016)	(0.018)
CN share	-0.095	-0.011	0.108	-0.056	-0.147	-0.047	-0.037	-0.136^b
	(0.148)	(0.046)	(0.075)	(0.208)	(0.087)	(0.140)	(0.036)	(0.052)
Log Empl	0.305^a	0.313^a	0.313^a	0.314^a	0.249^a	0.238^a	0.294^a	0.287^a
	(0.017)	(0.006)	(0.007)	(0.015)	(0.013)	(0.014)	(0.012)	(0.010)
Log Elect	0.068^a	0.100^a	0.094^a	0.121^a	0.096^a	0.140^a	0.104^a	0.105^a
	(0.014)	(0.005)	(0.006)	(0.011)	(0.010)	(0.012)	(0.009)	(0.007)
Log K	0.022	0.053^a	0.056^a	0.099^a	0.027	0.060	0.021	0.024
	(0.030)	(0.016)	(0.016)	(0.041)	(0.026)	(0.040)	(0.020)	(0.022)
Log Mat	0.438^a	0.467^a	0.469^a	0.265^a	0.458^a	0.499^a	0.569^a	0.454^a
	(0.039)	(0.017)	(0.016)	(0.023)	(0.024)	(0.027)	(0.032)	(0.020)
Obs.	1,428	14,363	12,026	5,110	2,903	4,852	5,232	7,512
R^2	0.754	0.708	0.717	0.567	0.732	0.716	0.772	0.683

ln Products: log number of imported products (HS10).
ln Supp/Prod: log number of suppliers per imported products.
Continuous: share of suppliers from which the buyer purchased for at least the previous year.
US Share: share of U.S. suppliers.
CN Share: share of Chinese suppliers.
Notes: Firm FE regression, years 2002–2008. Sector represents 3-digit NAICS. Robust standard errors, clustered at the firm level, in parenthesis. Significance thresholds are 0.1% (a), 1% (b), 5% (c). Each column implements a two-stage Levinsohn-Petrin specification.

Table C.7: Log Sales Regressions by Sector (OP)

Variable	Food	Bev.	Text. M	Text. P	App.	Leath.
ln Products	0.030^b	0.013	0.061^c	0.043^c	0.015	0.011
	(0.009)	(0.038)	(0.026)	(0.018)	(0.012)	(0.029)
$\ln \frac{\text{Supp}}{\text{Prod}}$	0.038^b	0.009	0.020	0.027	-0.009	-0.032
	(0.012)	(0.036)	(0.032)	(0.024)	(0.016)	(0.032)
Log Empl	0.188^a	0.157^b	0.335^a	0.267^a	0.306^a	0.146^a
	(0.018)	(0.056)	(0.050)	(0.027)	(0.020)	(0.039)
Log Elect	0.099^a	0.240^a	0.042	0.094^b	0.045^b	0.043^c
	(0.014)	(0.050)	(0.022)	(0.029)	(0.017)	(0.020)
Log Mat	0.511^a	0.286^a	0.335^a	0.414^a	0.464^a	0.501^a
	(0.045)	(0.057)	(0.077)	(0.037)	(0.023)	(0.057)
Obs.	6069	530	998	1504	3174	598
R^2	0.714	0.605	0.688	0.663	0.685	0.692

	Wood	Paper	Print.	Oil	Chem.	Plast.	Min.
ln Products	0.026^a	0.032^c	0.026^a	0.037	0.072^b	0.008	0.013
	(0.008)	(0.016)	(0.010)	(0.066)	(0.026)	(0.010)	(0.015)
$\ln \frac{\text{Supp}}{\text{Prod}}$	0.003	0.031	0.018	0.090	0.046	0.009	0.017
	(0.009)	(0.014)	(0.009)	(0.051)	(0.018)	(0.009)	(0.015)
Log Empl	0.257^a	0.194^a	0.357^a	0.113^c	0.239^a	0.293^a	0.317^a
	(0.017)	(0.021)	(0.021)	(0.048)	(0.018)	(0.021)	(0.024)
Log Elect	0.111^a	0.094^a	0.135^a	0.004	0.102^a	0.118^a	0.099^a
	(0.014)	(0.025)	(0.017)	(0.026)	(0.014)	(0.013)	(0.013)
Log Mat	0.544^a	0.623^a	0.349^a	0.559^a	0.346^a	0.500^a	0.489^a
	(0.028)	(0.047)	(0.033)	(0.046)	(0.041)	(0.026)	(0.032)
Obs.	4359	1758	3810	327	4310	5836	3102
R^2	0.801	0.868	0.759	0.682	0.518	0.792	0.754

	Met.	Met. P.	Mach.	Comp.	Elect.	Tr. Eq.	Furn.	Misc.
ln Products	0.066^b	0.019^a	0.018^c	0.076^a	0.008	0.027	0.024^b	0.026^b
	(0.021)	(0.005)	(0.009)	(0.020)	(0.018)	(0.015)	(0.008)	(0.009)
$\ln \frac{\text{Supp}}{\text{Prod}}$	0.080^c	0.017^c	0.014	0.009	0.003	0.023	0.001	0.033^b
	(0.033)	(0.007)	(0.011)	(0.026)	(0.020)	(0.017)	(0.011)	(0.010)
Log Empl	0.309^a	0.308^a	0.315^a	0.333^a	0.249^a	0.248^a	0.296^a	0.291^a
	(0.028)	(0.012)	(0.014)	(0.024)	(0.022)	(0.021)	(0.027)	(0.018)
Log Elect	0.073^a	0.090^a	0.085^a	0.110^a	0.092^a	0.126^a	0.090^a	0.086^a
	(0.020)	(0.009)	(0.011)	(0.017)	(0.014)	(0.020)	(0.017)	(0.011)
Log Mat	0.401^a	0.426^a	0.440^a	0.278^a	0.481^a	0.443^a	0.498^a	0.405^a
	(0.035)	(0.012)	(0.015)	(0.023)	(0.028)	(0.034)	(0.036)	(0.018)
Obs.	1,332	13,321	10,961	4,606	2,592	4,460	4,686	6,594
R^2	0.743	0.683	0.703	0.554	0.739	0.671	0.720	0.653

All regressions have firm and year fixed effects, Robust standard errors, clustered at the firm level, in parenthesis. Significance thresholds are 0.1% (*a*), 1% (*b*), 5% (*c*). Each column shows the first stage results from an Olley-Pakes specification.