Singapore Management University
Institutional Knowledge at Singapore Management University

Dissertations and Theses Collection (Open Access)
Dissertations and Theses

5-2021

Two essays on innovation and growth in China

Qiugu HE
Singapore Management University

Follow this and additional works at: https://ink.library.smu.edu.sg/etd_coll

Part of the Asian Studies Commons, and the Growth and Development Commons

Citation
HE, Qiugu. Two essays on innovation and growth in China. (2021). 1-78.
Available at: https://ink.library.smu.edu.sg/etd_coll/353

This PhD Dissertation is brought to you for free and open access by the Dissertations and Theses at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Dissertations and Theses Collection (Open Access) by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

TWO ESSAYS ON INNOVATION AND GROWTH IN CHINA

HE Qiugu

A DISSERTATION

IN

ECONOMICS

Submitted to the Singapore Management University in Partial Fulfillment

of the Requirement for the Degreee of PhD in Economics

March 2021

Dissertation Committee

Jianhuan XU (Supervisor/Chair)
Assistant Professor of Economics
Singapore Management University

Pao-li CHANG
Associate Professor of Economics
Singapore Management University

Yuan MEI
Assistant Professor of Economics
Singapore Management University

Bo ZHAO
Assistant Professor of Economics
Peking University

Sinapore Management University 2021
Copyright(2021) Qiugu HE

Singapore Management University

Abstract

School of Economics

Doctor of Philosophy

Two Essays on Innovation and Growth in China

by Qiugu HE

This dissertation studies China's economic growth from a perspective of industry dynamics. In chapter 1, I introduce the background and policies relating to China's economic growth after 1978. In chapter 2, I find that the elasticity of the average R&D expenditure of firms on competition is -0.29 in weak-IPR (intellectual property right) provinces, and -0.06 in strict-IPR provinces. Next, I use the Schumpeterian growth model to explain this finding: When the market becomes more competitive, a firm prefers imitation to innovation to a larger extent, as a means of getting new technology. Due to enforcement of IPR laws, the imitation replaces innovation more slowly in strict-IPR provinces, compared to weak-IPR provinces. In chapter 3, I estimate the TFPs of exporting and importing varieties for 6827 firms from 2002 to 2007 in the garment industry. I present three main channels of the growth of the aggregate TFPR(revenue) of continuous exporters: technology upgrade, reallocation of resources within continuous-exporting products, and switch of products. These three channels explain 27.2%,15.3%, and 9.46% of the aggregate TFPR growth, respectively. From the import side, the adjustment of import counts by firms explains 0.1% of the aggregate TFPR growth.

Acknowledgments

Time indeed flies.

While I still remember the difficulty for me to persuade the director Prof Anthony Tay to offer me a chance of interview during the application to the Phd program in SMU back in 2016, soon I will graduate from SMU and pass on the knowledge I gained in SMU to youngsters in Nankai University, China.

To say I love SMU and its faculty would be an understatement for my strong emotional link and gratitude for them built over the years, as the professors go above and beyond in guiding me throughout my journey here.

I want to extend the most appreciation to my supervisor, Assistant Professor XU Jianhuan. My first encounter with Prof.Xu started way back in 2015, one year before I joined SMU, where I attended a talk about the export pioneer behaviour given by Prof.XU's coauthor. I find the topic was very enlightening as it was related to the life cycle of firms in the field of management, what I was interested in. As a delightful coincidence, Prof XU happened to hold the phd admission interview for me. Subsequently in my year 1 at SMU, I had the opportunity to attend his teaching of ECON602, which brought me a fresh perspective to understand macroeconomics. It just came naturally that I have chosen him as my research supervisor based on these interactions over the years. Being a knowledgeable, patient, calm, friendly mentor, he is willing to teach, guide, and correct me whenever required. He always prompts me to think deeper and improve myself with a supportive manner, which is very important to cultivate my self-confidence despite the challenges I faced along the way. Without the guidance and help from Prof. XU, I would not ever be where I have been today.

I also want to thank the other two respectable committee members, Associate Professor CHANG Pao-Li and Assistant Professor MEI Yuan, where they have rendered me all the support and guidance they could to guide me through this academic journey. Words cannot describe my gratitudes for them and I shall pass on their spirits in my coming teaching post in China.

In addition, I also have benefited from the courses and suggestions from Associate Professor HSU Wen-tai, Assistant Professor Ismail Baydur, Amanda Jakobsson, and LI Jing.

Apart from faculty in SMU, I want to say thanks to Professor John Chiang and ZHAO Bo from Peking University for continuous support in the last 10 years .

Last but not the least, my thanks go to my beloved family for all their supports these years. My parents, younger brother influenced me with a positive attitude of optimism, which is essential for me to overcome all the difficulties over the years. Besides academic

achievement, my takeaways would be the friendship I have gains in Singapore, a thank you to all my friends.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

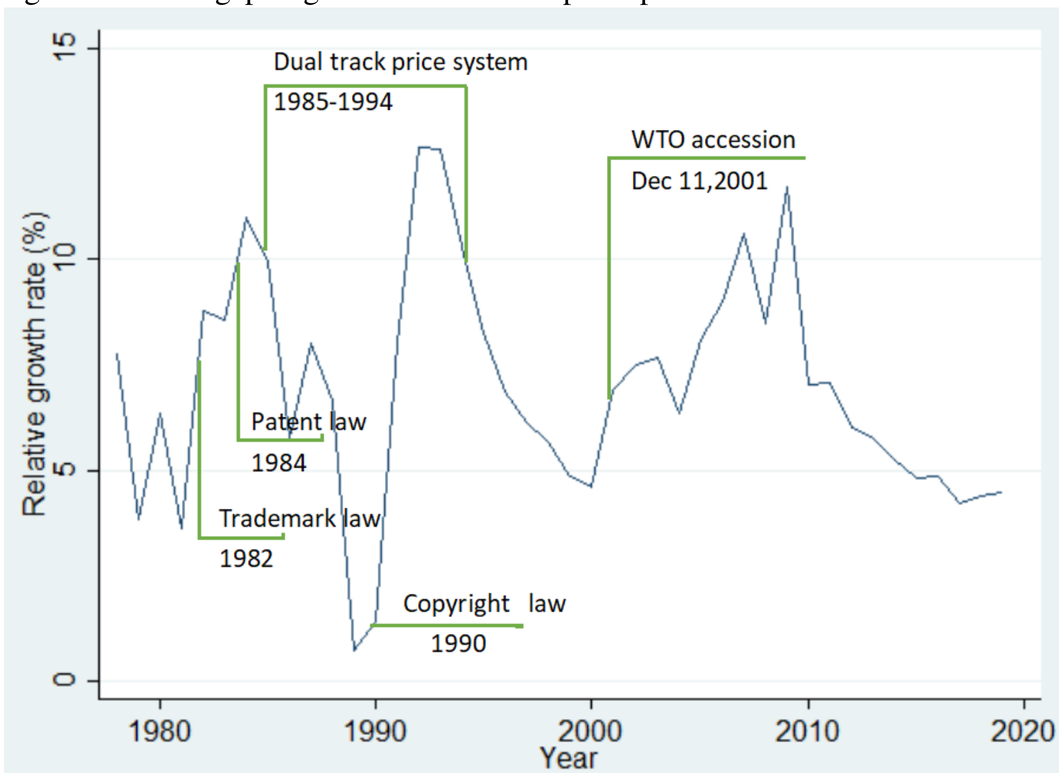
From 1978 to 2019, China experienced unprecedented growth in economy, where the average annual growth rate of GDP per capita was 8.42%, while the number for the rest of the world was about 1.46%. Consistent with the rapid growth in China, GDP per capita in China jumped from 4.98% of the world average to 74.6% of it. The aggregated GDP of China also increased in a similarly huge magnitude: While China's GDP only constituted 1.74% of the world's GDP in 1978, it accounted for 16.3% of it in 2019, with United States 24.4% and Japan 5.7% in 2019. China growth has already changed the situation of the world economy.

The rapid growth brings both benefits and potential threats to other countries. On the up side, more than a billion residents from China join the international industrial chain, which provides cheap goods to the world due to its relatively low wage and rich labor supply. As the productivity in China starts from a lower base, China firms quickly gained access to advanced technology, industry knowledge, efficient equipments, huge capitals and management experience through their integration to the world economy. From this aspect, China has the latecomer advantage to catch up with the developed economies. With the shrinking gap in efficiency between China and the developed economies, the world average productivity also increases at the same time. Apart from the benefits from its economic growth, China has been blamed for its technology theft, job destruction, counterfeiting, and unfair competition. In addition, the increasing interests of China squeezes the powers of some other countries. For example, the GDP share for OECD members to the world's total decreased from 80.1% to 61.1%, which is quite comparable to China's increased share 14.6%. In recent years, China's economic miracle has attracted attentions from economists, sociologists and policy makers. As a result, China's economic power impacts its growth in military, politics, and science.

1.1 Policies, endowments and performance relating to China's growth

The existing research in general attribute China's success to the government policy, institutional design, rich endowments, positive spillover of technology and knowledge, and the active engagement with the world's demand. The relevant policies include but are not limited to the promotion tournament, the dual-track price system during 1985 to 1994, the continuous reform of state-owned enterprise after 1978, the fiscal decentralization, the enact of laws related to intellectual property right, and the decision to join WTO [Yao, 2014]. A time line of the policies is given in Figure 1.1.

Figure 1.1: The gap of growth rate of GDP per capita between China and the world



The aforementioned policies are effective as it has taken advantage of national endowments and resources. National endowment refers to rich labor resource, long coastlines, and suitable natural environment. As the huge population size creates large scale of demand in goods and services, which drives business to invest more to scale up its production, and this helps to overcome high fixed cost of investment. This is especially obvious in industries which need huge investment,i.e. the high-end hardware industry. For example, the cost for building a Wafer foundry is around a billion dollars in year 2020. The revenue from the huge market makes it possible for China firms to finish the project. From the labor supply side, the rich labor resource from agriculture provides a rather inelastic labor supply curve for employers from cities. Then, the low labor wage transforms to cheap good price, which makes Chinese goods quite competitive compared to other countries. Apart from the labor resources, the long coastline in East China makes it convenient for firms from the inland to participate in international trade. The lengths of the coast line is around 14,500 kilometres, ranking 10th in the world, based on the World Factbook. The number of coastal provincial-level prefectures is 13, including Hongkong and Taiwan. The long coastline leads to a low level of natural trade iceberg cost. Furthermore, the fast development in infrastructures, including air traffic, railway and road networks, reduces the trade cost of goods. Moreover, the rich natural resources including freshwater, minerals, plants, animals, and energy provides rich materials input to industries.

The appropriate combination of policies and endowments results in good economic performance. The impacts of these policies include macroeconomic stability, urbanization, human capital accumulation and growth in productivity. China has done a good job in maintaining macroeconomic stability as it has successfully converted to a market-driven economy from a command economy without much volatility. One evidence is that China's average annual inflation rate has been stable at 4% in the last 40 years, comparing to the rates in other transition economies like Russia (112%), Ukraine (219%), and Vietnam (47%), and the world average at 5.83%. In addition to the stable economic environment, China adopts the compulsory education policy and higher education policy, which aims to improve younger generations' knowledge and help them meet the diversified demand in skills from various industries. The result of this policy is the fast accumulation of human capital: the number of graduates from high school was 0.165 million in 1978 while it reached 6.24 million in 2012. At the same time, workers and students from rural areas moved to the urban areas and settled down, which resulted in the fast increase in scales of large cities. For instance, urban population ratio rose from 17.9% in 1978 to 60.3% in 2019. Another indicator for better economy performance is the rising productivity, which is measured by TFP (Total factor productivity). During 1978-2005, the average growth rate of TFP is between 3.72% and 4.27%, which accounts for 42% of GDP growth [Zheng et al., 2009].

1.2 The turning point

After the WTO accession, the significant turning point appeared in year 2009, when the Subprime mortgage crisis occurred. After that, the trend of growth rate has been declining without a rebound, as Figure 1.1 indicates: the relative growth rate has been decreasing continuously from 11.7% in year 2009 to 4.4% in year 2019. Since I define the relative growth rate as the gap of growth rate between China and the world, the decreasing trend is unlikely to be explained by the global business cycle. In addition, the duration of the decreasing trend lasted 10 years, which is longer than the other two decreasing periods, 1985-1989 and 1995-2000. There must be some fundamental elements, which supported the previous growth, no longer supports the current growth. Wei et al. [2017] believe that the fundamental elements are labor supply, and investment. The labor supply begins to face shortage and the ratio of investment to GDP has been too large. As the utilization of production factors has reached a high level already, the previous experience of labor policy and investment policy unlikely support high growth in the future. Apart from labor and investment, TFP also plays a role in production function. The potential sources for TFP growth could be from advanced management experience, technology upgrade, resource reallocation and knowledge spillover. It is wise for policy makers to move their eyes to innovation, which is the main source for growth of developed countries. From this point of view, innovation related policies, including intellectual property right policy and the subsidy policy to R&D, become essential for future growth. Another potential growth comes from resource reallocation. In the thesis, I measure the growth due to technology upgrade, and resource reallocation.

1.3 Research about China's growth

The existing research in China relies on several limited database heavily, with the data from government agencies. These agencies include the National Bureau of Statistics (NBS), the General Administration of Customs China (GACC), the Supreme People's Court , the China National Intellectual Property Administration (CNIPA), the Ministry of Finance (MFP), the State Administration of Taxation (STA), the People's Bank of China (PBC), and the State Administration of Foreign Exchange (SAFE).[1] These agencies collect information when they perform their duties and save information in the database. Finally, economists transform the database into economic research.

By using these data, the existing research has achieved fruitful findings in understanding China's growth. Economists and policy makers understand the relationship between policies and economic performance based on their experience, communication and observation. Researchers use economics theory to explain the relationship. Based on the evidence from the data, the researcher testify the hypothesis. Constrained by limited database of economy performance, it is not easy to find direct links between objects. Thus, most existing research focuses on reduced-form empirical research to check the impacts of policy on economy, which in fact misses the direct link. In fact, economists usually use international panel, provincial panel, or city panel to compares the result of similar policies. The advantage of the research is to enrich our understanding of China growth and extend the range of topics. In addition, the reduced form empirical research offers the coefficient for independent variables, which is at least a method to quantify the impacts, though not perfect. As statistics are not efficient in causality identification, the problem left for economists is how to identify the clear causality between policies and economic performance. Different from microeconomic topics, like how a policy impacts individual choices clearly, the logic chain in the topics of macroeconomics is much longer and more complex. In addition, the logic of the most existing stories is too simplified, and many details, which are not recorded by data, are ignored. To summarize, the deeper problem is that most of the current research about China growth ignores the mechanisms and focuses on the correlation between objects. Without understanding the interaction of different objects, any prediction might be problematic.

To overcome the shortage of simplified analysis, we may introduce structural models to analyze the mechanisms, which provide the causality assumed by the modeler. The advanced theory behind structural models provides a framework how objects interact with each other. It offers the detailed logic chain of how the policy affects individual consumer and firm's decision. Further on, I can aggregate the individual decision into a macro level. The aggregation depends on the individual level data or the theoretical distribution of variables of firms' and personal characteristics. To show how the theory analysis beyond reduced form research work, I use several widely accepted models as example. Klette and Kortum [2004] provides a framework with the features above with the distribution of firm size, which makes quantification in the topic become an easy job. One interesting variation of the Schumperterian model is the step-by-step model, like Aghion et al. [2005]. In the model, the logic of innovation distinguishes from frog-swap innovation, where the firm

[1] Before 2018, the previous name of CNIPA is the State Intellectual Property Office (SIPO).

lacking frontier technology can become the leader of technology once the discovery of new technology succeeds. In the field of industry organization, Olley and Pakes [1996] estimate firm-level TFP by considering simultaneity and selection problem. In addition, the aggregation of TFP is widely used to explain economic growth at the country level. It offers a method to check the impact of policy on firm's growth.

More publication of micro-level data also benefit research adopting structural models. With more and more newly published micro-level database, we can find shorter logic link between data and the related policies. By using the micro data, it is meaningful to apply models to connect the policy, endowments, and economic performance. It offers a clearer quantitative framework to guide data flow in the economic system.

1.4 The aim of the thesis

I revisit two topics in China's growth, make new discoveries, and apply models to fit micro data. New structural models are designed to explain several new channels of China growth. In chapter 2, I build a Schumpeterian growth model to explain a new finding that strict IPR enforcement protects firms' R&D incentives better in the more competitive market. It is because the market failure caused by intellectual property right infringement is a more serious problem in a more competitive market. In chapter 3, it provides the first ever estimation regarding the TFPs of exporting and importing varieties within firms in China, which has not been studied much in the existing research field. My coauthor and I could open up the black box in understanding what have contributed to the growth of surviving firms. I find three new channels leading to the growth of surviving firms, namely, technology upgrade, reallocation of resources within existing products, and switch of products, have contributed to the growth of China's exporting firms by 27.2%,15.3%, and 9.46% respectively, during WTO accession. In addition, the import of varieties only explains 0.1% of the growth.

Chapter 2

IPR enforcement, competition and innovation: evidence from China

2.1 Introduction

The chapter documents two novel findings on the relationship between competition policy, IPR enforcement and innovation across China provinces and industries. In the group of provinces showing stricter IPR records from lawsuit data, including Shanghai and Jiangsu, the elasticity of average R&D expenditure per firm on competition is -0.06, while the elasticity is -0.29 in the weaker-IPR province group.

Competition and IPR protection are two of the most important drivers for growth. On one hand, the government encourages competition to reduce product price and stimulate innovation; on the other hand, IPR policy encourages innovation as it protects post-innovation profit. However, the relationship between IPR and market structure are mixed: (i). Competition policy, like encouraging firm entry, changes the market structure; (ii). IPR protection contributes to monopoly market. This paper studies not only how the interaction of IPR enforcement and competition policy affect innovation and welfare, but also how the two kinds of policies shapes market structure.

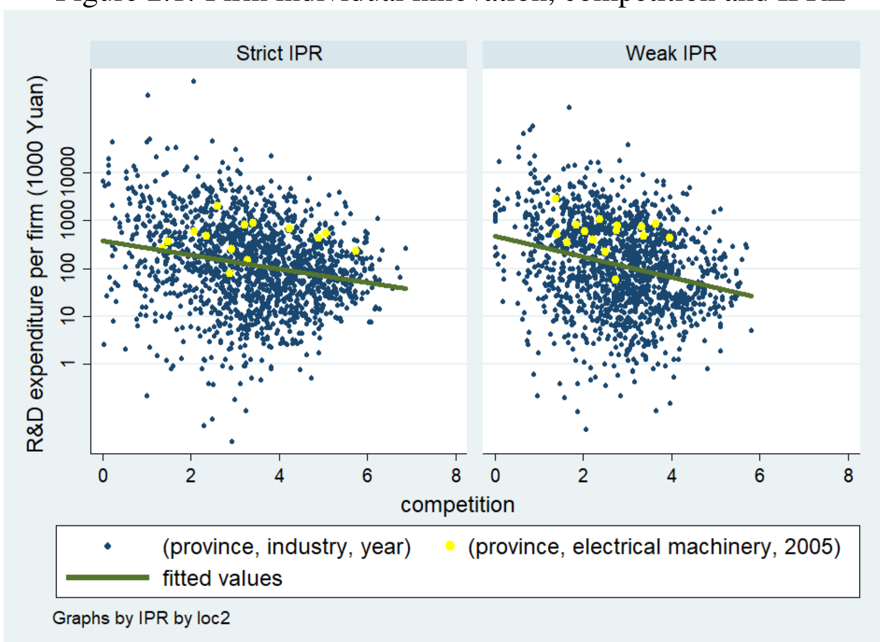
To address the research question, I take China as the example. In the last twenty years, competition brought by entry and exit of firms has been the most important source to drive economic growth in China. From 2000 to 2010, the number of medium and large sized firms has grown from 162,885 to 344,875.[1] During the same period, different provinces in China have seen uneven levels of progress in IPR protection. For instance, based on sample lawsuits involving medium and large sized firms, the probability of a plaintiff winning an IPR lawsuit is 32% in coastal provinces while the probability is 26% in inland provinces.[2] With this rich variation in competition and IPR enforcement across provinces, China can be used as a good example to understand the research question.

[1]the Annual Survey of Industrial Enterprises in China published by NBS (National Bureau of Statics) records firms with annual revenue above 5 million RMB.

[2]Since the defendant and plaintiff in IPR related cased can be firm, government, personnel and other agencies, the error, brought by the heterogeneity of organization, can be diminished by limiting to a sample. Since the medium and large sized firms play an important role in economy, it is appropriate to take it to reflect the IPR enforcement.

The paper aims to fill the gap between empirical and theoretical research. First, I construct an index of IPR enforcement using China Judgements Database (see Appendix A.1.1), where lawsuits during 2010 to 2017 are publicized. Second, I use firm level NBS (National Bureau of Statistics) data to construct an index of competition and calculate R&D figures at the provincial industrial level. In addition, I create a new dataset at the provincial industrial level by merging the aforementioned two datasets. Using the new dataset, I find that the positive correlation between competition and aggregate R&D expenditure is stronger in the group of provinces with stricter IPR enforcement, while this group witnesses a less negative correlation between competition and firm-level R&D expenditure. The two findings are recorded in Figure 2.1. After controlling the potential endogeniety problem of causality, competition and IPR enforcement impact R&D as the findings indicate.

Figure 2.1: Firm individual innovation, competition and IPRE



This paper is an extension of Klette and Kortum [2004] and adds the dimension of infringement. Imitation behavior is introduced to my model while the original model characterizes creative destruction of products. In this paper, not only does the firm make decisions in production, innovation, but also imitate. The infringed firm puts effort to sue the imitator. The stricter government IPR enforcement improves the efficiency of the innovator's legal action, leading to a higher rate of the infringement termination.[3]

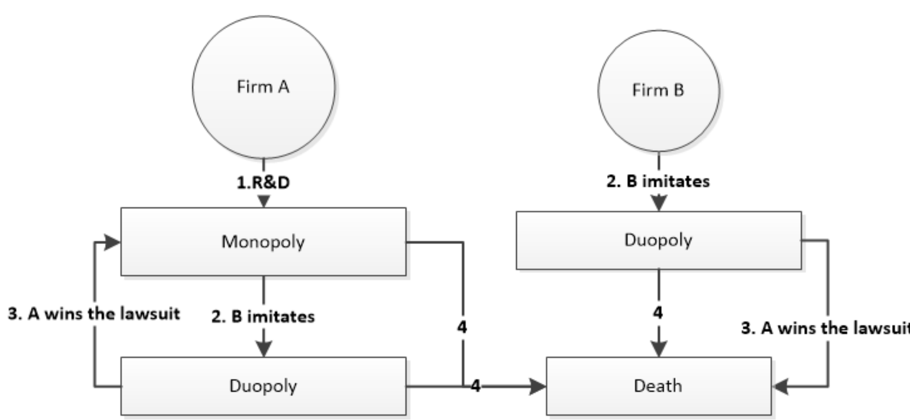
The model explains the findings as follow. Both competition and IPR affect innovation through externalities of technology. Competition policy impacts innovation via two channels. The first channel is due to the **"innovator stealing"**, where a firm's success in innovation leads to the destruction of an existing product. The entry of more firms, encouraged by competition policy, reduces the expected value of an innovation activity as the new

[3] In my model, the innovators' legal action follows a suing production function. The infringed innovator invests in collecting evidences and employing lawyers to increase the probability of winning the lawsuit.

product faces a higher rate of replacement. The second channel is called **"imitator stealing"**: firms take imitation as the method to extend the product portfolio, which reduces the innovator's profit. In addition, the second channel is also relating to dissemination of knowledge: imitators master the latest technology through imitation, which helps them innovate in the future. Strict IPR enforcement improves the intensive margin of innovation by punishing the infringing firm. However, it also hinders the extensive margin of innovation as less firms master the latest technology. Since competition policy and IPR policy both affects the extensive margin and intensive margin of innovation, there is a synergy from the two policies on aggregate innovation. The conclusion holds for welfare as it is a linear function of aggregate innovation in the model.

This model allows endogenous market structure. When a new innovating product faces one infringement, the monopoly market governed by the innovator becomes duopoly. The oligopoly market becomes more competitive when facing more imitators or infringements. Figure 2.1 shows the dynamics of market structure change with firms' decisions. All participants in the market can determine to expand the product portfolio by innovation or imitation. When the benefit from innovation is lower than imitation, a firm prefers less innovation when market structure becomes more competitive. That is because the payoff of innovation falls down faster than imitation. In the stricter-IPR environment, firms face less benefit from imitation, so that they still prefer innovation when market becomes quite competitive. The explanation supports the negative relationship between innovation and competitive market structure, and the relationship is more significant in strict-IPR environment.

Figure 2.2: The dynamics of market structure



Notes: Firm A is an innovator. Firm B is a counterfeiter. Event 4, 5 or 6 occurs passively when a new technology replace the current technology innovated by innovator A.

In the equilibrium, I show how to solve the distribution in an iterative method. Given the measure of a firm with a given portfolio, I use a adjacency matrix equation to solve the distribution of the neighbours(the firms having one more product compared to the given portfolio). With knowing the measure of entrant, firstly I solve the measure of firms with 1 products. Then, the measure of firms of other statuses can be solved.

The parameters are estimated based on the distribution of firm size, the distribution of profit rate, the ratio of innovating firm, and growth information in China in the year of 2007. The parameters are estimated by the SMM (Simulated Method of Moments). In the original model, the innovation intensity, entrants innovation determines the distribution of firms. To identify the imitation, my paper uses the index of competition(profit rate). The intuition is that perfect IPR protection leads to monopoly, but imitation makes the market more competitive. My estimated model implies that the imitation efficiency is 25 times of the innovation. The suing efficiency is 0.4 of imitation efficiency. In simulation, only 10% of all product markets do not suffer infringement, and the mean number of firms of all markets is 7, given the average size of the market is 25 million RMB. In the future, I will conduct a counterfactual analysis to check how innovation or welfare changes with competition and IPR enforcement.

The paper contributes to the existing literature in two ways.

First, it contributes to the growth literature and address the question of how the interaction of IPR enforcement and competition has a synergy impact on growth. Most existing growth models have not been able to answer this question because they do not combine infringement and firm entry together. In my model, it fills this gap: the model distinguishes innovators who invest in innovation technology, and suing technology from counterfeiter who invest in stealing technology. Quoting from Romer [1990], "The interesting case for growth theory is the set of goods that are nonrival yet excludable. ... technology is a nonrival input. ... improvement in the technology must confer benefits that are at least partially excludable".[4] The analysis of the IPR enforcement in the model is based on the the two attributes of technology, on one side, rivalry of technology determines the wide and cheap dissemination of knowledge; on the other side, the degree of excludability varies in IPR enforcement. The intrinsic nature of the interaction of competition and IPR enforcement is the interaction of the two attributes.

Second, the paper leads to a better understanding in how IPR enforcement could influence innovation and welfare. In the growth literature, knowledge spillovers, stealing effects, and monopoly are sources of technology externalities [Lentz and Mortensen, 2016]. In the IP literature, impact of strict IPR improves the innovation of developed countries but harms the welfare of developing countries by reducing knowledge spillover [Deardorff, 1992, Helpman, 1993]. Compared to these literatures, my paper illustrates how the government controls the "counterfeiter stealing" through legal proceedings.

This paper is organized as follows. Section 2 discusses the related literature and Section 3 provides the reduced-form evidence and data. Section 4 highlights the suing and infringement technology functions and Section 5 is a discourse on the model that I use in the study. Section 6 provides the estimates of the model. Section 7 provides counterfactual analysis result. Finally Section 8 concludes.

---

[4] Romer [1990] points out: Rivalry is a purely technological attribute. A purely rival good has the property that its use by one firm or perion precludes its use by another; a purely nonrival good has the property that its use by one firm or person in no way limits its use by another. Excludability is a function of both the technology and the legal system. A good is excludable if the owner can prevent others from using it.

2.2 Related Literature

There are several other related papers in the literature. First, The model in this paper extends the Schumpeterian growth model.[5] Klette and Kortum [2004] develop a dynamic general equilibrium model to describe how creative destruction is caused by entrants' innovation and incumbents' innovation. Lentz and Mortensen [2008] use the model to analyse how the exit of low productivity firm contributes to Denmark growth. I extend this model and add new elements of infringement so that I could provide a framework to discuss the impacts of IPR enforcement on growth by a quantitative method. Compared to this literature, the interaction in infringement and legal action between the innovator and the counterfeiter is new.

Second, the paper relates to the "competition and innovation" literature. Schumpeter [1943] finds a negative relationship between competition and innovation. Nickell [1996], Blundell et al. [1999], Hashmi [2013] and Aghion et al. [2015] all find a positive relationship, while Scherer [1967] and Aghion et al. [2005] discover a significant inverted-U shape. Different from their findings, my research finds that average firm-level R&D expenditure decreases with competition while aggregate R&D increases with competition. It implies that competition's effect on extensive margin of innovation dominates the intensive margin, which is because of the fact that more firms benefit from knowledge spillover. The relationship between innovation and competition could be adjusted by IPR enforcement, as government could control the "counterfeiter stealing" through lawsuit.

Third, a large numbers of paper emphasizes the impact of national patent law and IPR enforcement in innovation. Early papers focus on the national patent law. Deardorff [1992], Helpman [1993], and Grossman and Lai [2004] study theoretically the impact of strict IPR law and find that it improves the innovation of developed countries but harms the welfare of developing countries. Empirically, Qian [2007] use cross-country data to show that the strict patent law contributes to domestic innovation only in countries with higher economic development. Ang et al. [2014] show that provincial governments' attitude to IPR can be identified through keywords from the Communist Party's official newspaper. My research merges China Judgements Database and firm-level data to construct variables at province-industry level, which makes research about IPR enforcement across provinces and industries within a country becomes possible.

2.3 Data and motivating facts

This section describes data source, summarizes key variables at the beginning, and shows the procedure of constructing variables measuring competition and IPR enforcement. Then it offers two motivating facts: (1) aggregate-level of province-industrial R&D increases with competition; (2) In provinces with stricter IPR enforcement, the increment is larger; (3) firm-level of R&D decreases with competition; (4) In provinces with stricter IPR enforcement, the decrement is smaller.

[5] Compared to other literature about innovation, the Schumpeterian growth offers a beautiful dynamic framework of multi-product firm, and innovation decision, production decision and firm size distribution can be solved in close forms.

2.3.1 Data source

Our analysis draws on two data sources from China. These data sources can be merged through industry and location. Below I briefly describe these data and sample selection.

The first data source is the Annual Survey of Industrial Enterprises in China (ASIEC). The survey includes all industrial firms of various ownerships with annual revenue above 5 million RMB. It reports information on revenue, production cost, R&D expenditure, employment, four-digit industry classification and six-digit geographical identifier. The second data source is the China Judgements Database. Lawsuits from 2010 to 2017 are publicized through China Judgements Online aimed at information disclosure. These data contain information on plaintiff, defendent, cause of civil cases, date, court and a context including a detailed description of the case.

The process of merging the two databases is given in Appendix A.1.1. Sample selection is given in A.1.2.

2.3.2 The measurement of IPR enforcement and competition

The Chinese data allows me to construct province-industry IPR enforcement and degree of competition. While the geographical feature of IPR of China has been well documented by Ang et al. [2014], how competition and IPR enforcement interact within a nation is still a question which needs to be answered.

First, to measure IPR enforcement, I use *compensation share* to reflect the degree of the plaintiff being favoured by the court. I define plaintiff i's *compensation share* as the ratio of the compensation determined by judge against the plaintiff's claim, e_i. Then I calculate the average *compensation share* belonging to a case of action a in a province $\tilde{p}$,

$$e_{a\tilde{p}} = \frac{\sum_{i \in \Omega_{a\tilde{p}}} e_i}{N_{a\tilde{p}}}$$

where $\Omega_{a\tilde{p}}$ denotes the set of IPR cases related to a case of action a and proceeding in a court in province $\tilde{p}$ and $N_{a\tilde{p}}$ denotes the number of lawsuits in this set.

Different cases of action vary in the mean or standard error compensation share. A case of action with larger sample size in one province will dominates the measure if I do not control the size. Considering it, I give equal weight to each case of action,

$$e_{\tilde{p}} = \frac{\sum_{a \in A} e_{a\tilde{p}}}{N_a} \tag{2.1}$$

where $e_{\tilde{p}}$ measures the government's attitude to IPR protection, and N_a denotes the number of cases of action.[6]

There is an inherent assumption in our definition of the index for IPR enforcement: our understanding of the real attitude of the local government towards IPR enforcement

[6]To avoid bias caused by cases of action of a small frequency, I only consider cases of action appear in a significant number. There are 8 kinds in my sample: copyright contract disputes, trademark contract disputes, patent contract disputes, technology contract disputes, franchise contract disputes, copyright ownership disputes, trademark ownership disputes and, patent ownership disputes.

is increasing with the sample size (An outline of the enforcement is offered in Appendix A.1.4).

Second, I use the Herfindahl–Hirschman Index (HHI) to measure competition. The HHI is a measurement of firm market power. I use firm-level market share to calculate it

$$HHI_{\tilde{p}jt} \equiv \sum_{i \in \Omega_{\tilde{p}\tilde{j}t}} s^2_{i\tilde{j}t}$$

My measure of competition $c_{\tilde{p}\tilde{j}t}$ is negatively related to HHI,

$$c_{\tilde{p}\tilde{j}t} \equiv -log(HHI_{\tilde{p}\tilde{j}t}) \tag{2.2}$$

where i indexes firms, j indexes industry, $\tilde{p}$ indexes province, t indexes time, Ω_{jpt} denotes the set of firms in industry $\tilde{j}$ of province $\tilde{p}$ in year t. This index is negatively correlated to the average profit within an industry of a province in a linear way. The value of 0 indicates absolute monopoly.

2.3.3 Motivating facts

The analysis in Section 2.3.2 shows that the enforcement varies across geography and industry. This finding raises a question of whether, and how, the IPR enforcement impacts on innovation in China. To answer these questions, I extend Aghion et al. [2015], who provide evidence that innovation benefits from competition reform and this effect only exists in nations with better patent rights. They find that the response of innovation intensity to competition is positive in industries of countries with strong patent rights, but not so in industries of countries with weak patent rights. I apply a similar identification strategy to the China setting with the goal of examining whether this relationship exists between IPR enforcement and competition on innovation. In addition, I use the same identification strategy in Appendix A.1.6 as a robustness check.

To do so, consider the following regression:

$$y_{\tilde{p}\tilde{j}t} = \beta_0 + \beta_1 \times c_{\tilde{p}\tilde{j}t} + \beta_2 \times c_{\tilde{p}\tilde{j}t} \times H_{\tilde{p}} + F_t + F_{\tilde{j}} + F_{\tilde{p}} + \epsilon_{\tilde{p}\tilde{j}t} \tag{2.3}$$

where $y_{\tilde{p}jt}$ denotes two measures of R&D: logarithmic form of R&D expenditure and R&D intensity, respectively. Group indicator $H_{\tilde{p}}$ equates 1 if $e_{\tilde{p}}$ is greater or equal to the mean of e_p; otherwise, H_p is 0. $F_t, F_{\tilde{j}}$ and $F_{\tilde{p}}$ denote the fixed effects of year, industry and province. When I use IPR enforcement as the independent variable, $F_{\tilde{p}}$ is dropped to avoid col-linearity. $\epsilon_{\tilde{p}\tilde{j}t}$ is the error term.

Table A.5 and Table A.6 in Appendix A.1.5 report the summary statistics and correlation of variables and Table 2.1 reports the coefficient estimates. The estimation result shows that both measures of R&D increase with the interaction term of competition and IPR enforcement. In other words, IPR enforcement reduces the loss of R&D caused by fierce competition. However, R&D investment may enlarge the demand for protection for IPR, which leads to an endogenous problem of reverse causality. To mitigate this concern, I use the enforcement observables of the neighbour provinces as instrument variables. Specifications III and IV in Table 2.1 use the group indicator by the value of the neigh-

Table 2.1: Reduced form evidence

Regression Model	*OLS*	*OLS*	*IV*
Dependent Variable	Log firm-level R&D expenditure		
	I	II	III
Competition	-0.196^a	-0.417^a	-0.291^a
	(0.0326)	(0.0379)	(0.0372)
Competition $\times$ H IPRE	0.142^a	0.0792^b	0.228^a
	(0.041)	(0.0412)	(0.0407)
H IPRE	-0.331^b		
	(0.137)		
Province effect	no	yes	yes
Industry effect	yes	yes	yes
Year effect	yes	yes	yes
Observations	3053	3053	3053
R-squared	0.60	0.67	0.60

Note: a and b indicate significant at the 1% and 5% level, respectively. In parentheses, I report robust standard errors for coefficients. I include year, industry fixed effects in regression I. I include year, province, industry fixed effects in regression II and III. IPRE is short for intellectual property right enforcement. H IPRE is a dummy, which equates to 1 if the observation belongs to the group with higher value of IPRE. Here, I use *winning rate* to measure for IPR enforcement. In regression I, the mean of $Competition$ index is 3.03, so in a strict-IPR province with the mean competition index, the RD expenditure is 9.925%(3.03*0.142-0.331) higher than the counterpart province in the weak-IPR group.

bour province's *compensation share*. Enforcement in neighbour provinces' is close to local province enforcement because of the proximity of culture and economies. The neighbour's enforcement hardly affects local enterprises' R&D decision. The two specifications with instruments show that addressing the endogeneity does not affect the result.

Taken together, the results in Table 2.1 indicates that the elasticity of firm individual R&D expenditure on innovation is -0.063 (-0.291+0.228) in strict IPR enforcement group while the figure is -0.291 in the weak-IPR group. The closest finding is Aghion et al. [2015], who show patent rights protection strengthens the positive effect from competition policy and on country-level R&D expenditure. In their setup, competition index reflects the profitability. Different from them, my definition of competition is the market structure of an industry, which is positively correlated to firm number.

2.4 Firm, product and technology

The economy consists of one general type of multi-product firms. A firm is defined by the statuses of products in its portfolio. The status of a product is determined by the upgrade version, market structure, and the method to achieve the producing technology. An active firm owns at least one product line. Two kinds of decisions are made to achieve new technology: innovation creates new technology in the economy; imitation brings the latest technology to a firm. The productivity of product j of technology generation k, $z(j,k)$, varies with technology upgrades. In addition, when an innovator faces infringement, it determines how much effort to sue the infringers through suing function. Government attitude of IPR protection is shown in the efficiency of the suing function.

2.4.1 Product market

The measure of a continuum of products is fixed at one in the economy. Each product is indexed by $j \in [0,1]$. There is a market for each product. The status of the market is determined by the quality version and market structure together. The quality $z(j,k)$ of product j depends on the times of technology upgrade, k. I assume the step size of quality upgrade q is fixed. Thus, $z(j,k) = q^k$. The market structure is indicated by the number of imitators in the market(n). For example, there is zero imitator in a monopoly market, and one imitator in a duopoly market.

2.4.2 Product portfolio and firm

The status of a product of a firm is determined by the technology upgrade version, the source of the technology, and the market structure of the product market. The quality version is determined by the aggregate frequency of innovation on the product, k. In addition, there are two source of the technology: creation or learning from other firms. The two sources are denoted by I and H, respectively. At the same time, the market structure of the product market is the number of imitators, n.

A firm is determined by the held product portfolio, Ω. The portfolio is composed of three types of products: 1. The portfolio of innovating products which monopolies the product market, denoted by Ω^I; 2. the portfolio of the innovating products which suffers infringement, denoted by $\Omega^{\bar{m}}$; 3. the imitating products portfolio, Ω^m. I use the order in the set to indicate the market structure, and the value to show the count of products with the market structure. For example, there is one element in Ω^I. The element is the number of monopoly innovating products held by the firm. For the second portfolio, $\Omega^{\bar{m}} = \{0,1,0,3\}$ means: Zero product suffers one infringement, one product suffers two infringement, zero product suffers three infringement, three products suffer four infringements. For the infringement portfolio, $\Omega^m = \{0,2,1\}$ means: zero infringing product in a duopoly market, two products infringes in markets with two imitators, and one product in a market with three imitators.

2.4.3 Innovation technology

We assume that an innovator's innovation rate depends on both R&D expenditure and its *knowledge capital*. Following Klette and Kortum [2004], the size of knowledge capital is defined as the number of products produced by the firm. The innovation production function can be written into a cost function, $c^I(\lambda) = \gamma_1^I \lambda^{\gamma_2^I}$, where λ indicates the Poisson arrival rate of the new innovation for one piece of knowledge capital.

2.4.4 Imitation technology

The imitation rate also depends on imitation expenditure and knowledge capital. The cost function for a piece of knowledge capital to imitate is $c^H(h) = \gamma_1^H h^{\gamma_2^H}$, where h denotes the Possion arrival rate of newly successful imitation.

2.4.5 Suing technology

A firm can sue the infringers. The cost function of suing one plaintiff is $c^{sue}(\sigma) = \gamma_1^{sue} \sigma^{\gamma_2^{sue}}$, where σ denotes the Possion arrival rate of stopping the infringement. A higher value of γ_1^{sue} indicates lower IPR protection from the government.

2.5 A Schumpeterian model featuring IPR enforcement

In this section, I introduce the setup of the model. This section is organized as follows: I first describe the consumer and firm problem and then define the equilibrium. Second, I show how the matrix equation solve the distribution firms, as there is no closed form for the distribution. Then I use the model to explain the motivating facts.

2.5.1 Household

Setup The population is L. Individuals are homogeneous. Every individual is endowed with 1 unit of labor. In each date, a household consumes a Cobb-Douglas composite of goods and gets wage and dividend. Across years, the discount rate is ρ.

The household has a Cobb-Douglas preference over each good $j \in [0, 1]$ at each date $t \geq 0$

$$U = \int_0^\infty e^{-\rho t} ln C_t dt$$

$$ln C_t = \int_0^1 ln(X_{jt}) dj$$

where X_{jt}^a is the aggregate consumption of good j, adjusted by quality, at date t .

All versions of each product can be consumed. A version is determined by the generation, technology source, and market structure together. Note that different versions of the same product are perfect substitutes if quality is considered.

$$X_{jt} = \sum_{k=-1}^{J_t(j)} \left(x_{jt}^I(k) + \sum_{n=1}^{n^{max}} \left(x_{jt}^\Sigma(k,n) + x_{jt}^H(k,n) \right) \right) z_{jt}(k)$$

where $x_{jt}^I(k), x_{jt}^\Sigma(k,n)$ and $x_{jt}^H(k,n)$ denote the aggregate consumption of product j in different states at date t. I and Σ indicates the innovative product, the difference is that Σ

indicates the infringed status. H indicates the technology is achieved through imitation. A product experiences $J_t(j)$ times of upgrade through innovation from date 0 to date t. The market structure of I type product is monopoly. The quality version set of product j at time t is $\{-1, 0, ..., J_t(j)\}$. The quality of version k of product j is $z(j, k)$.

2.5.2 Firms and competition

There is a continuum of multi-product firms. They can acquire new technology through two methods. The first one is innovation. It brings new technology to the society and contributes to growth. The second is imitation, which helps the imitator forming its knowledge capital and steal the innovator's profit. Its contribution to growth is undetermined. When a firm innovates successfully, it prefers Bertrand competition as it masters the leading technology and monopolies the market by charging low price. When new imitators share the same advanced technology, the old innovator cannot adopts price competition because they will have zero profit. Through cooperation, they engage in Cournot competition to enjoy positive profit. The infringed innovator makes a decision of endeavor on lawsuit to the infringing imitator, considering the local government's attitude to IPR protection. A firm loses a product line if the technology is replaced by a new innovation, and exits if it loses all product lines.

Competition

Case 1. A new technology is developed. Compared to the existing active firms in the product market, whose cost is c, the quality of the new version grows by upgrade size, $q(>1)$. To occupy the whole market, the innovator charges price at the follower's cost c. Adjusted by quality, the cost of the new product is $\frac{c}{q}$. Then the innovator monopolies the market and gets profit rate $1 - \frac{1}{q}$.

Case 2. Competition with n imitators. When there are at least one imitator acquiring the advanced technology. The innovator does not adopts price competition because of zero profit as a result. To achieve positive profit, firms cooperate in the market. The market becomes oligarchy. Based on the Cobb-Douglas utility, the demand curve is $q = \frac{E}{p}$, where E denotes the consumption. I solve the equilibrium market share is $\frac{1}{1+n}$, and the profit rate is $\frac{1}{1+n}$. The detail is referred to Appendix A.2.

An implicit assumption of Klette and Kortum [2004] model is: the firm freely masters the outdated technology. The assumption guarantee the leaders adopts Bertrand competition. Otherwise, infinite followers will leave a very small market for the innovator. Note that imitation of the latest technology is costly.

The Firm's Problem

Different from Klette and Kortum [2004], in my paper a firm owns not-infringed portfolio Ω^I, an infringed portfolio $\Omega^{\bar{m}}$, and an imitating portfolio Ω^m. The creation of Ω^I and the destruction of the three portfolios depend on the creative destruction process. The dynamics of Ω^I, $\Omega^{\bar{m}}$ and Ω^m also depend on the imitation decision, the infringement rate, and the suing decision.

Due to the Cobb-Douglas utility function, the consumption expenditure on every product market is the same. The market share and profit of a firm's product are determined by the market structure.

Creative destruction. Innovation destructs the existing product technology: (1) To earn more profit, an innovator invests in R&D expenditure, which leads to the Poisson rate λ at which the new technology arrives. (2) A firm's product line faces the probability of destruction because other firm's new technology randomly replaces the existing ones. The Poisson hazard rate of destruction, faced by every existing product, is $\mu^I > 0$. It is called the *destruction rate*, which is exogenous to the individual firm.

Infringement and the shock Infringement decision results in infringement shocks to innovating products: (1) Firms make decision to imitate other firms randomly with Poisson arrival rate h. (2) The innovating products randomly suffer imitation. The Poisson hazard rate of this event, per product market, is μ^H. The parameter $\mu^H(> 0)$ is called the *infringement rate*, which is exogenous to the individual firm.

Suing decision and the shock The suing decision of an innovator results in shocks to its imitators: (1)Facing imitation, a firm determines an effort to sue the imitator with the Poisson hazard rate of winning, σ. The winner will reduces one competitor. (2) The imitator faces two shocks, (i) it will exit the market if the innovator sues him and wins the lawsuit, (ii) it enjoys higher profit if the innovator sues other firms and wins.

As both innovation and imitation rely on knowledge capital, a firm should decide the usage of the knowledge capital. To model the choice between innovation and imitation, I assume the payoff of the two strategies meets independent shocks respectively. The shock follow the extreme value distribution, $\epsilon \in exp(-\theta\epsilon)$. In addition, the firm decides not to use the knowledge capital if the payoff of the two strategies is lower than a cutoff, o. It is because firm manager may enjoy less management cost if they do not organize a new research team. So the firm should choose one of the three usage of knowledge capital: innovation, imitation, or non-utilization.

The value function of a general firm is given by

$$
\begin{aligned}
rV(\Omega^I, \Omega^{\bar{m}}, \Omega^m) = & \max_{\lambda,h,\sigma_j} \underbrace{\sum_{j\in\Omega^I} \pi_j(0) + \sum_{j\in\Omega^{\bar{m}}} \pi_j(n) + \sum_{j\in\Omega^m} \pi_j(n)}_{\text{profits}} \\
& + (|\Omega^I| + |\Omega^{\bar{m}}| + |\Omega^m|) \times \max\left\{ \left(-\underbrace{wc^I(\lambda)}_{\text{R\&D}} + \lambda \Delta V^{RD} \right) \epsilon_1, \left(-wc^H(h) + h\Delta V^{Imitation} \right) \epsilon_2, o\epsilon_3 \right\} \\
& + \sum_{j\in\Omega^{\bar{m}}} \left\{ \underbrace{-wc^{sue}(\sigma_j)}_{\text{suing cost}} + \sigma_j \Delta V_j^{sue} \right\} + \underbrace{\sum_{j\in\Omega^I} \mu^I \Delta V_j^{I,dest}}_{\text{destruction}} + \underbrace{\sum_{j\in\Omega^{\bar{m}}} \mu^I \Delta V_j^{\bar{m},dest}}_{\text{destruction}} + \underbrace{\sum_{j\in\Omega^m} \mu^I \Delta V_j^{m,dest}}_{\text{destruction}} \\
& + \underbrace{\sum_{j\in\Omega^I} \mu^{\bar{m}} \Delta V^{I,infrig}}_{\text{infringement}} + \underbrace{\sum_{j\in\Omega^{\bar{m}}} \mu^{\bar{m}} \Delta V_j^{\bar{m},infrig}}_{\text{infringement}} + \underbrace{\sum_{j\in\Omega^m} \mu^{\bar{m}} \Delta V_j^{m,infrig}}_{\text{infringement}} \\
& + \underbrace{\sum_{j\in\Omega^m} \frac{1}{n_j} \bar{\sigma}_j \Delta V_j^{m,lawsuit}}_{\text{being sued}} + \underbrace{\sum_{j\in\Omega^m} \frac{n_j - 1}{n_j} \bar{\sigma}_j \Delta V_j^{m,lawsuit,2}}_{\text{being sued}}
\end{aligned}
\tag{2.4}
$$

where the first line on the right hand side is current gross profit flow of its not-infringed portfolio, infringed portfolio, and infringement portfolio. The second line is the expected payoff from innovation,imitation and non-utilization of the knowledge capital. Among them, $|\Omega^I| + |\Omega^{\bar{m}}| + |\Omega^m|$ is the aggregate number of active products, which indicate the amount of accumulated knowledge capital. The first item of the third line is the net payoff from suing an imitator. The left three terms in the line are the expected loss as the three portfolios suffer from creative destruction shocks. In the forth line, the three portfolios all suffer infringement shocks. Finally, the first item of the fifth line shows the loss of the firm as a defendant. The second item shows how this infringer benefits from other imitators being sued by the innovator. The details of the change of value function are referred to Appendix A.2.

To solve the firm's problem, I assume and testify the value function follows a summation form:

$$
V(\Omega^I, \Omega^{\bar{m}}, \Omega^m) = \sum_{j\in\Omega^I} v^I(0) + \sum_{j\in\Omega^{\bar{m}}} v^{\bar{m}}(n_j) + \sum_{j\in\Omega^m} v^m(n_j) \tag{2.5}
$$

Then I use the first order condition to solve the firm's decision:

$$
w\frac{\partial C^I(\lambda)}{\partial \lambda} = v^I(0), w\frac{C^H(h)}{\partial h} = Ev^m, w\frac{\partial C^{sue}(\sigma_n)}{\partial \sigma_n} = v^{\bar{m}}(n-1) - v^{\bar{m}}(n) \tag{2.6}
$$

Finally, the value of product in different statues can be solved:

$$v^I = \frac{\pi(0) + p^I wC^I(\lambda) + P^H(-wC^H(h) + hEV^m) + \mu^{\bar{m}} v^{\bar{m}}(1)}{r + p^I\lambda + \mu^I + \mu^{\bar{m}}} \tag{2.7}$$

$$v^{\bar{m}}(n) = \frac{\pi(n) + p^I(-wC^I(\lambda) + \lambda v^I(0)) + p^H(-wC^H(h) + hEv^m) - wC^{sue}(\sigma_n) + \mu^{\bar{m}} v^{\bar{m}}(n+1)}{r + \sigma_n + \mu^I + \mu^{\bar{m}}} \tag{2.8}$$

$$v^m(n) = \frac{\pi(n) + p^I(-wC^I(\lambda) + \lambda v^I(0)) + p^H(-wC^H(h) + hEv^m) + \mu^{\bar{m}} + \frac{n-1}{n}\sigma_n v^m(n-1))}{r + \sigma_n + \mu^I + \mu^{\bar{m}}} \tag{2.9}$$

Equation(2.5) indicates that the value of a firm is the sum of its portfolios' value. The equation(2.6) shows how decisions are determined by first order conditions(FOC). The first FOC shows that the marginal cost of innovation is the value of a new innovating product. The second FOC shows that the marginal cost of imitation is the expected value of a new imitating product. The third FOC shows that the marginal cost of the suing decision is the growth of value from the number of competitors reducing by one. The value of an innovating product is referred to equation(2.7). It is composed of the innovation cost, the benefit from imitation , and the infringed value . In addition, it is affected by interest rate, innovation rate, destruction rate, and infringement rate. Similarly, Equation(2.8) and (2.9) show the the details of value of infringed innovating products and infringing products, respectively.

Entry

There is a mass of potential entrants. To proxy the competition policy, I set the measure of entrants exogenous. An entrant invests in F^I labor in return of an innovator, or it have free access to entry by imitation. The problem faced by the entrant is:

$$\max\left\{\left(-wc^I(\lambda) + \lambda\Delta V^{RD}\right)\epsilon_1^{ent}, \left(-wc^H(h) + h\Delta V^{Imitation}\right)\epsilon_2^{ent}\right\}$$

Size distribution

The status of the industry is summarized by the measure of firms in different statuses. The firm status is determined by the three portfolios. As every portfolio is possible to experience innovation, imitation and other shocks, the evolution of a firm is quite complex. The same is true for the distribution of firm size, as products of different market structure vary in size.

Assume the status of a firm is defined by the three portfolios, $\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, \Omega^m = \{b1, b2, ...\}$. The inflow of the measure comes from innovation, imitation, winning a lawsuit, destruction shock, infringement shock, and suing shock. Similarly, the outflow of the status also results from these decisions and shocks.

Table 2.2: The inflow of firms with the state $\Omega^I = \{\{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, \Omega^m = \{b_1, b_2, ...\}\}$

source	the status in the last day	rate
Innovation	$\Omega^I = \{a_0 - 1\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, \Omega^m = \{b_1, b_2, ...\}$	$\left((a_0 - 1)P^I + \sum_n a_i P^I(n) + \sum_n b_n P^H(n)\right) \lambda$
Suing	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ..., a_{n-1} - 1, a_n + 1, ...\}, \Omega^m = \{b_1, b_2, ...\}$	$(a_n + 1)\sigma(n)$,$\forall n$
Type 1 shock	$\Omega^I = \{a_0 + 1\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, \Omega^m = \{b_1, b_2, ...\}$	$(a_0 + 1)\mu^I$
(Destruction)	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ..., a_n + 1, ...\}, \Omega^m = \{b_1, b_2, ...\}$	$(a_n + 1)\mu^I$,$\forall n$
	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, \Omega^m = \{b_1, b_2, ..., b_n + 1, ...\}$	$(b_n + 1)\mu^I$,$\forall n$
Type 2 shock	$\Omega^I = \{a_0 + 1\}, \Omega^{\bar{m}} = \{a_1 - 1, a_2, ...\}, \Omega^m = \{b_1, b_2, ...\}$	$(a_0 + 1)\mu^H$
(Infringement)	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ..., a_n + 1, a_{n+1} - 1, ...\}, \Omega^m = \{b_1, b_2, ...\}$	$a_n + 1, \forall n$
	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, \Omega^m = \{b_1, b_2, ..., b_n - 1, ...\}$	$(b_n - 1)\mu^I$,$\forall n$
Suing shock	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, \Omega^m = \{b_1, b_2, ..., b_n - 1, b_n + 1, ...\}$	$(b_n + 1)\bar{\sigma}, \forall n$
Imitation	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, \Omega^m = \{b_1, b_2, ..., b_n - 1, ...\}$	$\left(a_0 P^H + \sum_i a_i P^H(n) + \sum_i b_i)\bar{P}^H(n)\right) hD(n)$

Table 2.2 lists the source state and the rate of it to become the current state. In the first line, it shows how the firm with one less innovating product innovates to become the current status: As all knowledge capital can be used for innovation, and the probability of the usage depends on the product status, the rate is a sum of every specific rate of every piece of knowledge capital to innovate. In the second line, suing the infringer in a market with status n reduces the number of products in market of status n, but increases the number of product in markets with $n - 1$ competitors. In the third line, a destruction shock reduces the number of not infringed innovating product. In the forth line, a destruction shock ruins an innovating product of market structure status n. In the fifth line, destruction shock terminates an imitating product of market structure n. In the sixth line, one of the $a_0 + 1$ product is infringed, then the number of product of status zero adds 1 to a_1. In the seventh line, one innovating product with market structure status n is infringed, and then the number of products with status $n + 1$ pluses 1 to a_n. In the eighth line, the innovating product imitated by the firm suffers new infringement. In the ninth line, the infringing product of (n+1) loses the lawsuit and the number of product with status n adds 1. In the tenth line, all knowledge capital is also possible to be used for imitation, then the number of products with status n increases by one. Note that $D(n)$ is the possibility of the new infringed product being status n.

Table 2.3: The outflow of a state $\Omega^I = \{\{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, b\}$

source	the status in the next day	rate
Innovation	$\Omega^I = \{a_0 + 1\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, b$	$\left(a_0 P^I + \sum_n a_i P^I(n) + \sum_n b_n P^H(n)\right) \lambda$
Imitation	$\Omega^I = \{a_0 - 1\}, \Omega^{\bar{m}} = \{a_1 + 1, a_2, ...\}, b,$	$a_0 P(0)^H h$
	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ..., a_n - 1, a_{n+1} + 1, ...\}, b$	$a_n P^H(n)\lambda$
Suing	$\Omega^I = \{a_0 + 1\}, \Omega^{\bar{m}} = \{a_1 - 1, a_2, ...\}, b$	$\sigma(1)(a_1)$
	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ..., a_n + 1, a_{n+1} - 1, ...\}, b$	$\sigma(n)a_n$
Type 1 shock	$\Omega^I = \{a_0 - 1\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, b$	$a_0\mu^I$
(Destruction)	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ..., a_i - 1, ...\}, b$	$a_i\mu^I, \forall i$
	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ..., a_i, ...\}, b - 1$	$b\mu^I$
Type 2 shock	$\Omega^I = \{a_0 - 1\}, \Omega^{\bar{m}} = \{a_1 + 1, a_2, ...\}, b$	$a_0\mu^H$
(Infringement)	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ..., a_i - 1, a_{i+1} + 1, ...\}, b$	$a_i\mu^H, \forall i$
Type 3 shock	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, b$	$\frac{n}{1+n} b\bar{\sigma}$
(suing)	$\Omega^I = \{a_0\}, \Omega^{\bar{m}} = \{a_1, a_2, ...\}, b - 1$	$\frac{1}{1+n} b\bar{\sigma}$

Table 2.3 lists the states of firms which follow the given state. By the inflow and outflow,

I can solve the variation in measure for every status of firms. Especially, when the change of measure is zero, the measure of firms is in steady state.

The stationary distribution. Since I have shown the inflow and outflow of a specific state, I find a trick to solve the stationary distribution of firms. However, since the inflow and outflow is different from the classical example, there is no closed-form solution. A numerical method is offered below to solve the distribution of firms. In addition, I have some analyses about decisions and shock on the distribution. As I define the neighbour as firms have one more product compared to the given firm. The way to find the measure of the innovating neighbor of given status is provided by Proposition 1. The method to find the imitating neighbor is quite similar, and is omitted.

Proposition 1. *The stationary distribution of firms with 1 more innovating product satisfies:*

status	inflow	outflow
x_0	$(\sum_i a_i + \sum_i b_i)p\lambda M + (a_1+1)\sigma_0 x_1$	$(a_0+1)(\mu^I+\mu^H)x_0$
x_1	$(a_0+1)\mu^H x_0 + (a_2+1)\sigma_1 x_2$	$(a_1+1)(\mu^I+\mu^H+\sigma_0)x_1$
x_n	$(a_{n-1}+1)\mu^H x_{n-1} + (a_{n+1}+1)\sigma_1 x_{n+1}$	$(a_n+1)(\mu^I+\mu^H+\sigma_n)x_n$

Here, I use x_n to denote the measure of firms has one more product in market with status n than the original firm. For example, x_0 is the measure of firms with status $[a_0+1, (a_1, a_2, a_3, ...), (b_1, b_2, b_3, ...)]$, the counterpart of x_1 is $[a_0, (a_1+1, a_2, a_3, ...), (b_1, b_2, b_3, ...)]$

Based on the transition dynamics, we have a matrix equation to solve the distribution.

$$\begin{pmatrix} (a_0+1)(\mu^I+\mu^H) & -(a_1+1)\sigma_0 & 0 & 0 & 0 \\ -(a_0+1)\mu^H & (a_1+1)(\mu^I+\mu^H+\sigma_0) & -(a_2+1)\sigma_1 & 0 & 0 \\ 0 & -(a_1+1)\mu^H & (a_2+1)(\mu^I+\mu^H+\sigma_1) & -(a_3+1)\sigma_2 & 0 \\ 0 & ... & ... & ... & 0 \\ 0 & ... & -(a_{n-1}+1)\mu^H & (a_n+1)(\mu^I+\mu^H+\sigma_{n-1}) & -(a_{n+1}+1)\sigma_n \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} (\sum_i a_i + \sum_i b_i)p\lambda M \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.10}$$

To simplify the analysis, I set $\sigma_n = \sigma$. From the equation above, I show

- $\lambda \uparrow$, then the measure of the firm with one more innovating product $\uparrow$. If μ^I and μ^H do not change, the degree of competition, when controlling the number of products, does not change.
- $\mu^I \uparrow$, then the measure of the firm with one more innovating product reduces proportionally. But the possibility of the added product in any status is not changed.
- $\mu^H \uparrow$, then the new added product has a larger possibility to be infringed, which leads to a more competitive market. It does not change the aggregate measure of firms with one new product.

- $\sigma \uparrow$, then the new added product has a less probability being infringed, which leads to a less competitive market. In addition, it does not change the aggregate measure of firms with one new product.

Proposition 2. *In steady state, the following matrix equation solves the stationary distribution of markets of market structure $M(n)$.*

$$\begin{pmatrix} (\mu^I+\mu^m) & -\sigma_1 & 0 & 0 & 0 \\ -\mu^m & (\sigma_1+\mu^m+\mu^I) & -\sigma_2 & 0 & 0 \\ ... & ... & ... & ... & ... \\ ... & 0 & -\mu^m & (\sigma_{n-1}+\mu^m+\mu^I) & -\sigma_n \\ 0 & 0 & 0 & -\mu^m & (\sigma_n+\mu^I) \end{pmatrix} \times M = \begin{pmatrix} \mu^I \\ 0 \\ ... \\ ... \\ 0 \\ 0 \end{pmatrix} \tag{2.11}$$

In addition, it offers the solution for $D(n)$ in Table 2.2

$$D(n) = \frac{M(n)}{1-M(0)} \times (1 - M(n^{max}))$$

The proposition solve the distribution of product markets of various market structure.

2.5.3 The aggregate setting in steady state

At the industrial level, there are several restrictions on the aggregate measure of product lines and other parameters in steady state. First, the inflows of aggregate measure of the three types of product lines should equate the outflows. Second, the aggregate destruction rate is the result of innovation. Third, the aggregate infringement rate is caused by firms' individual innovation decision. Forth, the individual suing shock suffered by infringing firms is due to the related infringed firms' suing decision.

There a balance between innovation and destruction.

$$\eta^I P^I_\eta + \left(M(0)P^I + \sum_{n=1}^{max} M(n)P^I(n) + \sum_{n=1}^{max} nM(n)\bar{P}^I(n) \right) \lambda = \mu^I \tag{2.12}$$

where $M(n)$ denotes the measure of product markets with n imitators; $P^I_\eta, P^I, P^I(n), \bar{P}^I(n)$ denote the possibility of the entrants, not infringed innovating product, innovating product with market status n, imitating product with market status n to adopt innovation as the method to acquire new technology, respectively.

The relationship between infringement decision and the aggregate infringement rate follows

$$\left(\eta^I P^H_\eta + \left(M(0)P^H + \sum_{n=1}^{max} M(n)P^H(n) + \sum_{n=1}^{max} nM(n)\bar{P}^H(n) \right) h \right) (1 - M(n^{max})) = \mu^H \tag{2.13}$$

where the entrant, and incumbents imitate from a random object. However, imitating a product of the most competitive degree does not bring positive profit, thus only $1 - M(n^{max})$ of the total infringement will succeed. The balance between suing and suing shock is

$$\sigma(n) = \bar{\sigma}(n)$$

where $\sigma(n)$ and $\bar{\sigma}(n)$ denote the suing effort and the exogenous termination rate for infringement in the same product market, respectively.

2.5.4 General equilibrium

The general equilibrium is the allocation of resources to production, innovation, imitation, suing, $\{l^p(n), l^I(n), l^H(n), l^{sue}(n)\}$, the possibility of firms' choice on innovation $\{P^I_\eta, P^I, P^I(n), \bar{P}^I(n)\}$, the possibility of firms' choices on imitation $\{P^H_\eta, P^H, P^H(n), \bar{P}^H(n)\}$, innovation rate λ, imitation rate h, suing effort $\sigma(n)$, under wage w, the Bertrand competition price $p^{Bertrand}$, and the Cournot competition price $\{p^{Cournot}(n)\}$, and aggregate destruction rate $\{\mu^I, \mu^H\}$, given the firm i's state $\Omega^I(i), \Omega^{\bar{m}}(i), \Omega^m(i)$:

1. When new innovation occurs, the active innovating firm i engages in Bertrand competition and choose the optimal price for product j, $\{p_t^{Bertrand}(i, j)\}$.
2. When new imitation occurs, the active firm i engages in Cournot competition and choose the optimal price for product j, $\{p_t^{Cournot}(i, j)\}$.
3. Given the value of different infringed status, an infringed firm i chooses the effort σ to drive the imitator out of the market.
4. Given the value of newly innovating product, a firm chooses the innovation rate λ to extend the product portfolio by innovation.
5. Given the value of expected infringement, a firm chooses the imitation rate h to extend the portfolio by imitation.
6. Facing the value of different choices, the firm decide the probability to innovate or imitate to acquire new technology.
7. Facing the benefit of newly innovating product and expected value of infringement, the potential determines the method to enter the product market.
8. the aggregate destruction rate, imitation rate and suing shock are determined by the aggregation of firms' decisions.
9. The representative household consumes all products with the same amount.
10. Labor market clearing condition should be satisfied.

2.6 Explanation for motivating facts

By solving the model, the motivating facts are supported by the following propositions. Proposition 3 provides the mechanism that firms' expenditure in innovation decreases in competition when post-innovation profit is not as good as other options, like imitation.

In addition, proposition 4 explains why IPR protection reduces the diminished amount of innovation resulted by competition.

By assuming firms have independent taste shocks on choices of acquiring technology, firms' decisions depend on two parts: the expected profit from newly developed product and the value of the current product. When (1). themarket is not competitive, and (2). return of innovation is low, the ratio of payoff from innovation and current value to the payoff from imitation and current value is not bad. when market becomes more competitive, the payoff depends more on the newly innovating or imitating product. But the net profit from innovation is low, so the probability of choosing innovation is lowest when market is most competitive. However, in strict-IPR environment, the return of imitation is low. When market is extremely competitive, innovation is still a good choice. So strict IPR makes the R&D expenditure reduces more slowly.

Proposition 3. *When the benefit of innovation* ($-wc^I(\lambda)\Delta t + \frac{\lambda\Delta t}{1+r\Delta t}v^I + \frac{1}{1+r\Delta t}v^I(n)$) *is smaller than a cutoff, which is the weighted mean of benefit of imitation* ($-wc^H(h)\Delta t + \frac{h\Delta t}{1+r\Delta t}Ev^m + \frac{1}{1+r\Delta t}v^I(n)$) *and outside option* ($o\Delta t + \frac{1}{1+r\Delta t}v^I(n)$)*, then the possibility of choosing innovation as the method to achieve new technology is decreasing in* n*.*

$$\frac{\partial P^I(\Delta t, n)}{\partial n} < 0$$

where $P^I(\Delta t, n) = \frac{\left(-wc^I(\lambda)+\frac{\lambda\Delta t}{1+r\Delta t}v^I+\frac{1}{1+r\Delta t}v^I(n)\right)^\theta}{\left(-wc^I(\lambda)+\frac{\lambda\Delta t}{1+r\Delta t}v^I+\frac{1}{1+r\Delta t}v^I(n)\right)^\theta+\left(-wc^H(h)+\frac{h\Delta t}{1+r\Delta t}Ev^m+\frac{1}{1+r\Delta t}v^I(n)\right)^\theta+\left(o\Delta t+\frac{1}{1+r\Delta t}v^I(n)\right)^\theta}$.

Proof. First, I simplify the form into $P^I(\Delta t, n) = \frac{(a+f(n))^\theta}{(a+f(n))^\theta+(b+f(n))^\theta+(c+f(n))^\theta}$, where $a \equiv -wc^I(\lambda) + \frac{\lambda\Delta t}{1+r\Delta t}v^I + \frac{1}{1+r\Delta t}v^I(n)$, $b \equiv -wc^H(h) + \frac{h\Delta t}{1+r\Delta t}Ev^m + \frac{1}{1+r\Delta t}v^I(n)$, $c \equiv o\Delta t$ and $f(n) \equiv \frac{1}{1+r\Delta t}v^I(n)$.

$$P^I(\Delta t, n) = \frac{1}{1 + \frac{(b+f(n))^\theta}{(a+f(n))^\theta} + \frac{(c+f(n))^\theta}{(a+f(n))^\theta}} \tag{2.14}$$

$$log(\frac{1}{P^I(\Delta t, n)} - 1) = log\left((b+f(n))^\theta + (c+f(n))^\theta\right) - \theta log\,(a+f(n)) \tag{2.15}$$

$$\begin{aligned}
\frac{\partial log(\frac{1}{P^I(\Delta t,n)} - 1)}{\partial f} =& \frac{\theta\,(b+f)^{\theta-1} + \theta\,(c+f)^{\theta-1}}{(b+f)^\theta + (c+f)^\theta} - \theta\frac{1}{a+f} \\
=&\theta\frac{(b+f)^{\theta-1}(a+f) + (c+f)^{\theta-1}(a+f) - \left((b+f)^\theta + (c+f)^\theta\right)}{\left((b+f)^\theta + (c+f)^\theta\right)(a+f)} \\
=&\theta\frac{(b+f)^{\theta-1}(a-b) + (c+f)^{\theta-1}(a-c)}{\left((b+f)^\theta + (c+f)^\theta\right)(a+f)}
\end{aligned} \tag{2.16}$$

In addition, I have

$$\frac{\partial logP^I}{\partial f} = -(1-P^I)\theta\frac{(b+f)^{\theta-1}(a-b)+(c+f)^{\theta-1}(a-c)}{\left((b+f)^{\theta}+(c+f)^{\theta}\right)(a+f)}$$

Since $\frac{\partial f}{\partial n} < 0$, the sign of $\frac{\partial P^I}{\partial n}$ is determined by $\frac{\partial P^I}{\partial f}$.

To guarantee $\frac{\partial P^I}{\partial n} < 0$, I need $\frac{\partial P^I}{\partial f} > 0$. Then I have

$$\begin{aligned} a &< \frac{(b+f)^{\theta}}{(b+f)^{\theta}+(c+f)^{\theta}}b + \frac{(c+f)^{\theta}}{(b+f)^{\theta}+(c+f)^{\theta}}c \\ &= wb + (1-w)c \end{aligned} \tag{2.17}$$

□

The meaning of the last inequality is that bad payoff from innovation leads that firms prefer less innovation when the market is more competitive. The intuition is that the value of innovation shrinks faster than the value of other candidate choices when a market becomes more competitive. That's because when market becomes more competitive, the added value brought by innovation directly and the added value brought by imitation directly determines the firm's choice. However, the added value brought by innovation is so little that firms do not like to innovate.

Lemma 1. *If the benefit of innovation is quite high, then I have*

$$\frac{\partial P^I(\Delta t, n)}{\partial n} > 0$$

Proposition 4. *Following the setting in proposition 1, when the benefit of innovation* $(-wc^I(\lambda)\Delta t + \frac{\lambda\Delta t}{1+r\Delta t}v^I + \frac{1}{1+r\Delta t}v^I(n))$ *is smaller than a cutoff, which is the weighted mean of benefit of imitation* $(-wc^H(h)\Delta t + \frac{h\Delta t}{1+r\Delta t}Ev^m + \frac{1}{1+r\Delta t}v^I(n))$ *and outside option* $(o\Delta t + \frac{1}{1+r\Delta t}v^I(n))$*, then the elasticity of probability of innovation on competition decreases in* γ_1^{sue}*,*

$$\frac{\partial^2 logP^I(\Delta t, n)}{\partial n \partial \gamma_1^{sue}} < 0$$

Proof.

$$\frac{\partial logP^I(\Delta t, n)}{\partial n} = -(1-P^I)\theta\frac{\partial logP^I}{\partial f} = -(1-P^I)\theta\frac{(b+f)^{\theta-1}(a-b)+(c+f)^{\theta-1}(a-c)}{\left((b+f)^{\theta}+(c+f)^{\theta}\right)(a+f)}\frac{\partial f}{\partial n}$$

When $a < b$, if $b\uparrow$,then (1) $1-P^I\downarrow$;(2)the numerator$\downarrow$ (3)$\frac{\partial f}{\partial n}$ is not affected, but the value is negative. Finally, I have

$$\frac{\partial^2 logP^I(\Delta t, n)}{\partial n \partial b} < 0$$

so that

$$\frac{\partial^2 logP^I(\Delta t, n)}{\partial n \partial \gamma_1^{sue}} < 0$$

□

The two propositions show the existence of a distribution of firms' market structure. Further on, the market structure can be observed by profit rate, though we cannot identify the firms of the same product market. The market structure is due to infringement, which is governed by IPR policies.

2.7 Estimation

Now I turn to the estimation of the parameters of the model. I use a two steps method repeatedly to estimate them. In the first step, I choose parameters relating to three functions, market size, fixed cost of R&D, innovation cost by entrant. By the FOCs and the aggregate balance conditions, firm decisions and aggregate rates are solved. In addition, the distribution of product markets of market structure is solved at the same time. In the second step, I simulate thousands of firms to enter, innovate, imitate, sue in the economy. The simulation result is used to generate the distribution of firms. By guessing the initial parameters in the first step, I try to find a group of parameters which makes the simulated distributions of firm size and profit rate fit data in the second step. Finally, The fit is good.

2.7.1 Model Identification

The distribution of firm size and profit rate, and the growth information identify the parameters. If the distribution of the firm size is fat-tailed, it implies large product portfolios held by firms. However, the ratio of innovating firms to the aggregate number is only 12% in 2007. The low innovation ratio indicates that the large portfolio mainly depends on imitation rather than imitation. So I deduce that the innovation efficiency is low while the imitation efficiency is high. In addition, the distribution of firms' profit rates reflects the situation of imitation. That's because successful innovation results in monopoly while the imitation from other firms reduces the monopoly. Overall, more than 70% firms's profit is lower than 20%. In addition, Cournot competition helps us identify the number of competitors based on the profit rate. To support the widely imitation, the suing efficiency must be low compared to the imitation efficiency.

2.7.2 Estimation procedure

There are two steps to estimate the model. In the first step, firms decisions are made and the aggregate rate are calculated by guessing parameters. In the second step, distribution of firm size and competition index are generated by simulating firm dynamics with the known decisions and aggregate rate from step 1. If the model distribution and data distribution do not fit well, I go to step 1 to optimize parameters. Repeat the two steps until the fit is good.

Step 1. Generating firm decisions and the aggregate rate

The input of the economy system is exogenous parameters, which govern the shape of cost functions, technology upgrade size, market size, entrant measure, and choice parameter. With these parameters, firms make decisions dynamically, and the aggregation of the decisions becomes exogenous shocks to the individual firm.

The explanation for parameters are given in Table 2.4. With the given initializing parameters $\{y, \eta, \gamma_1^I, \gamma_2^I, \gamma_1^H, \gamma_2^H, \gamma_1^{sue}, \gamma_2^{sue}, f, f^I, \theta, o\}$, I solve the mapping firm decisions and the mapping aggregate shocks. First, I use the backward iteration of value function as it is a dynamic model. To be noted: equation 2.5 provides a method to decompose the firm dynamics into product dynamics. So I do not solve heterogeneous firm decisions, and only discuss how firm decision made based on its every product status independently. There are $1 + 2n^{max}$ value functions to be estimated. Among them, One is the innovating product without infringement; there are n^{max} statuses for innovating products suffering infringement; and there aren^{max} statuses for imitating products. The backward method started from the last day, firms do not survive to the next day but they earn profit, so they only produce and do not make innovation, imitation or suing decisions for the future. With the definition of the last day, the infinite dynamic model becomes finite. One day earlier, every individual product brings profit, firms make innovation decision, imitation decision, and suing decision only based on the statues of the product. At the same time, the aggregate destruction rate and infringement rate need to be calculated. To solve μ^I, μ^H, M, Ev^m, h, a group of equations is used, including FOC of imitation, equation 2.11, equation 2.12, and equation 2.13. With the solved firm's decisions and the aggregate rates, I solve the value of products in one day earlier. Repeat the step until the value converges. Then the unknown decisions and shocks, $\{\lambda, h, P^I, \bar{P}^I(n), P^I(n), P^I_{ent}, P^H, \bar{P}^H(n), P^H(n), P^H_{ent}, \sigma(n), \mu^I, \mu^H, M(n)\}$, in the steady state of the dynamic problem with given parameters is solved.

Table 2.4: Parameters

parameter	explanation	value
y	market size	25000(1000 Yuan)
η	entrant ratio	0.05
γ_1^I	coefficient of innovation cost function	500
γ_2^I	the power of innovation cost function	3
γ_1^H	coefficient of imitation cost function	20
γ_2^H	the power of imitation cost function	2
γ_1^{sue}	coefficient of suing cost function	50
γ_2^{sue}	the power of suing cost function	4
f	entrant's innovation cost	600(1000 Yuan)
f^I	incumbent's innovation fixed cost	250(1000 Yuan)
θ	parameter for discrete choice error term	0.0001
o	value of other option	250
η	the measure of entrant	0.05

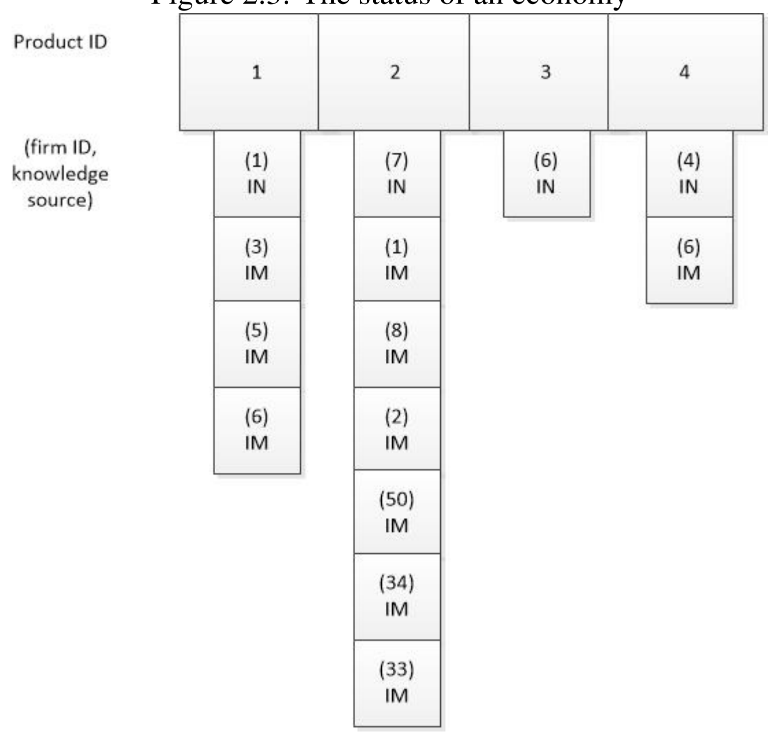
Step 2. Simulation

Now, I have firm decisions $\{\lambda, h, P^I, \bar{P}^I(n), P^I(n), P^I_{ent}, P^H, \bar{P}^H(n), P^H(n), P^H_{ent}, \sigma(n), M(n)\}$ and shocks $\{\mu^I, \mu^H, \bar{\sigma}(n)\}$. Assume the number of entrant is $x = 1000$. On Every date, the number of entrants is 10. I use an id generator to grant firm id to identify entrants.

Status for a firm's product is (k, n). The first state variable k denotes the source of knowledge $k \in \{innovation, imitation\}$, and the second state variable n denotes market structure, $n \in \{0, 1, ..., n_{max}\}$. The number of product market is $p^{max} \equiv \frac{x}{\eta^I}$, where η^I denotes the measure of innovating entrants. An array with $\frac{x}{\eta^I}$ elements is generated to simulate the all product markets. The element in the array is one product, which records ids and knowledge sources of active firms in the product market.

The status of a product market is $ms \equiv (id_0, id_1, id_2, ..., id_n)$. The knowledge source is indicated by the order of ids. In detail, id_1 is the innovator, and other ids are of imitators. The whole market is composed of $\{ms_1, ms_2, ..., ms_{p^{max}}\}$, described by Fig 2.3. The id of a product market is p, $0 < p < p^{max}$.

Figure 2.3: The status of an economy



For every entrant, generate uniform random variable $u^I_{ent} \in (0, 1)$. If $u^I_{ent} < P^I \Delta t$, the entrant chooses innovation, else the entrant chooses imitation.

For incumbent products, I divide them into $2n^{max} + 1$ groups based on knowledge source and market structure. For each group n, I use a uniform random variables $x(n)$ to generate random numbers to every product. The value of the number determines the method of acquiring new technology: (1). If $x < P^I(n)\lambda \ Deltat$, the firm uses the knowledge capital of the product to innovate; (2). If $P^I(n)\lambda\Delta t < x < P^I(n)\lambda\Delta t + P^H(n)h\Delta t$, it imitates.

To simulate the verdict result of Lawsuit, I use a uniform random variable $x^{lawsuit}(n)$ to generate random number for each infringed products of market structure n. If $x^{lawsuit}(n) < \sigma(n)\Delta t$, then one of the infringing firms leaves that product market.

Now that I divide products in the economy into $2(n^{max} + 1)$ groups, the firm ids of new innovating products, new imitation products are collected in two sets. For the new innovative products, I generate the uniform random variable to generate random numbers and map them to the product market space, and then destruct the existing product market. The number of competitors of the destructed market becomes zero. The first id in the market becomes the firm id. For the new imitation products, I generate another uniform random variable and maps them to the product market space, and add the imitating firm id to the existing product market. In addition, the number of competitors in the markets pluses one.

The winning firm in lawsuit destructs the product in order 1 of the same market.

2.7.3 Estimation results

Table 2.5 shows a comparison between data moments and the simulated moments. Overall, the model matches closely to the targeted moments.

The estimates of the model parameters are reported in Table 2.4. From the estimated parameters, I find that innovation in China is quite expensive: investing in 500+600 workers can exchange for the arrival of new technology in one year. Imitation only needs 20 workers. In addition, the protection is also not efficient. It provides evidence that it is easy to imitate but not easy for the infringed firm to stop the infringement. In addition, the low return of innovation is due to R&D fixed cost.

Table 2.5: Model fit

	data	model		data	model
firm size $1st$ percentile	5,060	5,000	firm profit rate $5th$ percentile	0.029	0.092
firm size $5th$ percentile	5,797	5,625	firm profit rate $10th$ percentile	0.045	0.103
firm size $10th$ percentile	7,110	7,777	firm profit rate $25th$ percentile	0.079	0.146
firm size $25th$ percentile	12,024	17,500	firm profit rate $50th$ percentile	0.132	0.226
firm size $50th$ percentile	26,286	62,768	firm profit rate $75th$ percentile	0.209	0.264
firm size $75th$ percentile	64,900	327,160	firm profit rate $90th$ percentile	0.318	0.333
firm size $90th$ percentile	165,065	1,220,300	firm profit rate $95th$ percentile	0.407	0.404
firm size $95th$ percentile	317,888	2,283,531	firm profit rate $99th$ percentile	0.636	0.540
firm size $99th$ percentile	1,340,945	5,392,460			
mean of firm size	121,526	463,155			
standard error of firm size	1,204,254	1,228,823			
innovating firm ratio	12%	45%			
growth rate	18%	79%			

Notes: Firm A is an innovator. Firm B is a counterfeiter. Event 4, 5 or 6 occurs passively when a new technology replace the current technology innovated by innovator A.

2.8 Conclusion

This paper uses China firm-level data and lawsuits data to investigate the relationship between IPR enforcement, competition and, innovation. Empirically, I have two findings: (i). Competition reduces firm individual R&D expenditure; (ii). The elasticity of firm individual R&D expenditure on competition is larger in stricter-IPR provinces. The empirical findings are robust to a number of alternative specifications.

To explain the two findings, the paper develops a Schumpeterian growth model with allowing IPR infringement. In the model, I show how the Schumpeter effect is twisted by the infringement. When the benefit of innovation is lower than imitation, firms prefer imitation more as market becomes more competitive. As strict IPR reduces the infringement profit, which makes innovation as not a bad choice when market is quite competitive.

The model features knowledge spillover through imitation. For example, n imitators in one market also have capability to innovation. They accumulate the knowledge capital through imitation.

As a dynamic general equilibrium model, it allows the quantification of a series of competition policy and IPR enforcement. In the next step, I will simulate how IPR and competition policy impacts on innovation.

Chapter 3

Firm productivity and the Variety of Inputs and Outputs: Evidence from Chinese Trade Data

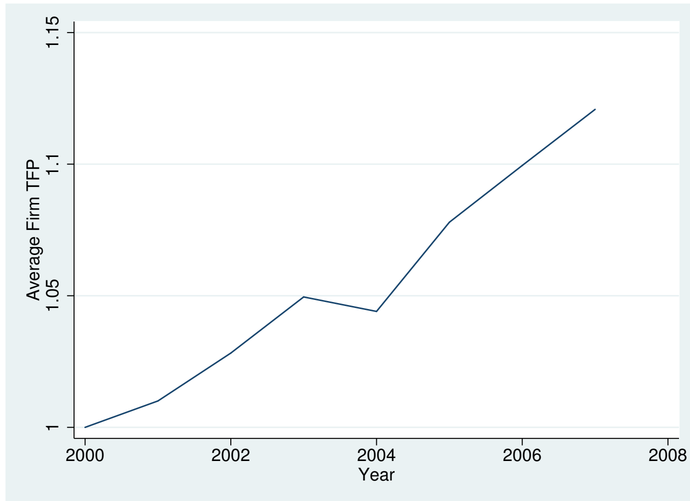
3.1 Introduction

The paper studies how the trade liberalization in China changes the firm productivity. My coauthor Dr. Jianhuan XU, Dr. Ken Onish, Dr. Guang YANG and I develop a theoretical framework to estimate revenue productivity (TFPR) and real productivity (TFPQ) with multi-product firms. We find that the aggregate TFPR increases by 47.1% from 2002-2007 and TFPQ increases by 64%, suggesting that the observed TFPR increase is mainly driven by the rise of real productivity change. The decrease of markup explain the gap between TFPR and TFPQ. Except the traditional firm entry and exit channel, We further decompose the change of productivity of surviving firms into three channels: (1) access to foreign inputs; (2) technology upgrade; (3) resource re-allocation within the firm. I find the three channels leads to 0.1%, 27.2% and 15.3% growth of the aggregate TFPR, respectively. I also find that the aggregate TFPR of private firms and foreign firms improve by 35.3%, and 41.6% while the SOE improves by only 1.65%. In addition, technology upgrade contributes most to private firms. In addition, the growth of small-size firms is due to upgrade rather than the firm entry and exit channel.

Since China join the WTO, the export and import of China grow significantly. Now China is the largest export country and the second largest import country in the world. At the same time, the average firm level productivity of China increases as well. Figure 3.1 plots the average firm level productivity from year 2000-2007.[1] As we can see, the firm average productivity increases by 12 %.

[1]The firm level TFP is computed using the Olley-Pakes (1994). We normalize the average TFP in year 2000 to be 1.

Figure 3.1: The Average of Firm TFP



NOTE: This figure plots the average firm level TFP. The firm productivity is estimated through the Olley-Pakes (1996). The TFP in year 2000 is normalized to 1.

Our paper is to understand the link between international trade and the change of firm level productivity.[2] Three possible channels are considered: (1) Access to the international inputs. Foreign inputs may affect firm productivity through two channels: as in quality ladder models, foreign inputs may have a higher price-adjusted quality, and as in product-variety models they imperfectly substitute domestic inputs. Halpern et al. [2015] find that a quarter of Hungarian productivity growth is attributed to imported inputs. (2) Resource allocations within the firm. International competition may help firms to focus on products they have the largest comparative advantage and improve the resource allocations within a firm. Redding et al. [2006] documents that the unproductive products will be dropped when access to the international market improves. (3) Firms improve their productivity by upgrading their technology or management. Bustos [2011] finds the Argentina firms increase the R&D after the trade liberalization. We quantify the increase of firm productivity in China into these three channels.

Our starting point is the China Customs database, which tracks the firms' import and export information at the product level. We combine this data set with the Chinese manufacturing firm survey data. For each firm, we know its export products and imported products as well as other resources the firm uses (capital and labor). We then build a structural dynamic industry equilibrium model that allow firms to optimally choose their export products and imported inputs. The model is quite flexible to permit rich heterogeneity across products and firms.

We estimate this model by the micro data. In doing so, we face two key empirical challenges. First, the imports are chosen endogenously by the firm. We deal with this identification problem using a structural approach following Halpern et al. [2015]. Our

[2] Old trade theory, as Melitz [2003] focuses on the resource allocations across firms when trade cost declines but fixing the firm productivity as given.

model implies that the effect of imports on the firm production function is only through the number of imported varieties and a time-shifter capturing the relative quality-adjusted price change between foreign goods and domestic goods. Then we can identify the productivity gain from the imported inputs channel. Second, we do not observe the resources allocation within the firm. However, our structural model implies that the resources allocations within the firm is related to the revenue shares of different products in a firm.[3] Through the revenue shares distribution change within a firm, we can identify the productivity gain through the second channel. Our model also introduces unobserved products productivity within the firm, similar as Olley and Pakes [1996]. The setting heterogeneity exists in productivity of goods within the firm. We follow the identification strategy in Olley and Pakes [1996] and uses the transition process of productivity to identify it.

A great benefit of our data is that we can observe the import and export prices at very dis-aggregated product levels. It helps us to separate the mark-up and real productivity changes. We are not the first one to do this job. De Loecker [2011] also estimate the markup and real productivity in a multi-product firm model using an Indian data. They assume that for a fixed product, the technology of single product firm and multiple product firm are the same. There are two drawbacks of this approach. First, they are unable to perform counterfactual analysis since they do not model the multiple products firms' pricing and resource allocation decisions. Second, they use a too strong assumption that all products in a firms share a same productivity level. Our approach can overcome both drawbacks.

Our result shows that the aggregate revenue productivity (TFPR) improves by 47.1% from 2001-2007, which triple the productivity change estimated from Olley-Pakes(1994). We further decompose the firm level productivity into the three channels mentioned before. The most significant contribution at the firm level comes from the firm entry and exit channel, 28.4%. The upgrade of the technology is also important, 12.8%. Reallocation of resources between existing products within a firm contributes to 7.20%. Within the surviving firm, the entry and exit of products contributes to 4.45%. Besides, the contribution of access to the foreign imported goods is quite small, 0.05%.

Comparing the real productivity change and markup change, we find that most TFPR improvement in China comes from the real productivity (TFPQ) change. As the average demand elasticity grows from 2.6 to 3.3, lower markup reduces prices around 17%. Thus, the aggregate TFPQ grows by 64%, which is a sum of TFPR growth rate and price reduction.

Besides the papers mentioned above, our paper relates to several other literature. First, our paper contributes to the empirical literature of exploring the firm productivity gain of the international trade. In a multi-product firms model setup, De Loecker et al. [2016] and Dhyne et al. [2016] find that trade liberalization primarily affected markets by causing firms to find ways to reduce marginal costs as opposed to causing output prices to fall. Our paper suggests how the firms can cut their marginal cost and highlight the channel of domestic inputs improvement and allocation of resources across products.

Second, our paper is related to the literature that focuses on productivity in developing countries. The low productivity of the developing countries usually is attributed to resource mis-allocation **across** firms (Hsieh and Klenow [2009]) or lack of competition **across** firms (Bloom et al. [2007]). When the frictions are reduced (such as trade cost), people usually

[3] Dhyne et al. [2016] uses a different approach and gets a similar condition as ours.

start to think the reallocation across firms (Restuccia and Rogerson [2008]). Our paper contributes to the literature by focusing on how the allocation within the firm will change in response of the decline of the trade cost.

The remainder of the paper is organized as follows. In the next section, we provide a brief overview about the data used in the analysis and document some motivation facts. In section 3, we lay out the structural model. Section 4 discusses our estimation strategy and section 5 shows the main results. Section 6 concludes.

3.2 Data and Motivation Facts

3.2.1 Data

In this paper, we match Chinese manufacture firm survey data and Chinese customs database. The first dataset covers operating information of all Chinese manufacture firms whose annual sales are above 500 million RMB (71 million USD) or SOEs. The Chinese customs data records the export and import price and quantity information of each firm at the product level (HS6). We match the two datasets by matching the names of legal representative, head-quarter's address and telephone number. The efficiency of the match process turns out to be good. Take year 2007 as an example, the number of firms in Chinese manufacture firm survey data is 298992, among which 75930 firms are exporters. In the same year, there are 193567 exporters in the Chinese customs database. We can match 46604 firms in the two datasets.

In the end, we can observe the capital, number of employees, total expenditure of materials at the firm level. We also observe other firm characteristics, such as the set-up year and the ownership. At the product level, we know the prices and quantities of all imported inputs and exports.[4]

We choose three industries in our analysis: chemistry, home appliance and clothes. They are very large industries, and actively involved in export and import. We define a product as a HS5 code. Table B.3 reports the summary statistics in our data. On average a firm exports 5 products and imports 20 products.

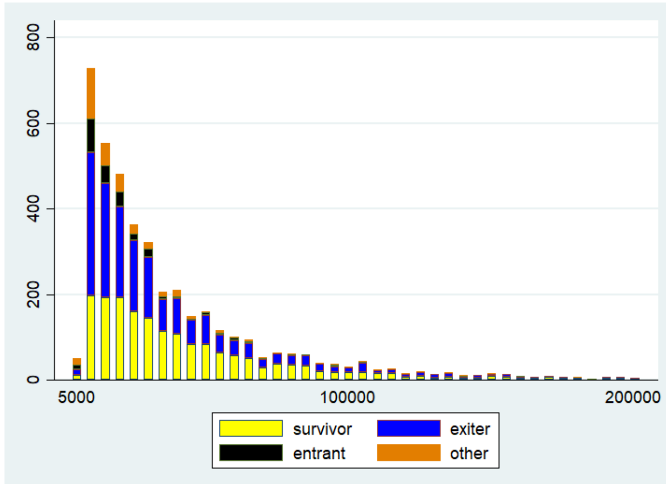
3.2.2 Motivation Facts

Fact 1: the blackbox of survivors' growth

From 2002 to 2007, a large share of firms survive, as Figure 3.2 shows. The ratio of survivor is increasing with the firm size. However, the growth of the survivors is still a blackbox. It is valuable to analyse the growth at the product level, which brings new ideas about how firms switch product to contribute to growth.

[4]We define the domestic output as one product (j=1) and the price of the domestic output is normalized to 1. So for those firms who only sell domestically, we consider these firms sell a single product with unit price.

Figure 3.2: The frequency of firms' revenues (1,000 yuan) in year 2005



Note: If a firm appear in the database both in year 2002 and 2007, then it is defined as a survivor. If the firm appears in year 2002 but disappear in year 2007, then it is defined as an exiter. If the firm only appear in year 2007, then it is defined as an entrant.

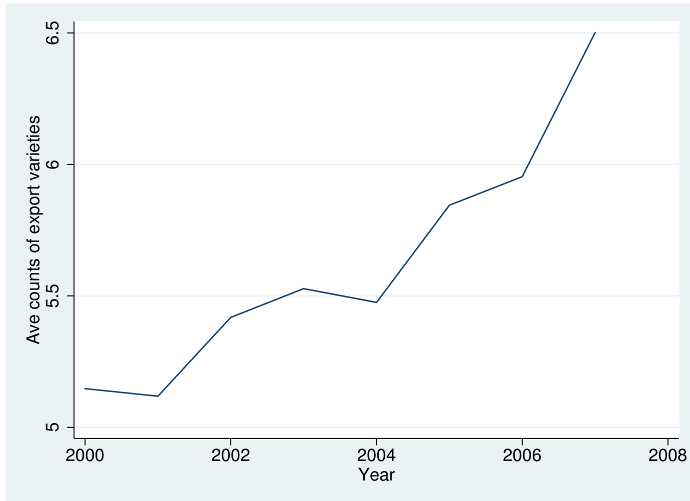
Fact 2: Export Varieties increases, but the revenue shares are more dispersed

Figure 3.3 plots the average counts of export varieties per firm. We define one export variety as an HS5 product. From 2000 to 2007, the average export varieties of a firm increases from 5 to 6.5. The increase of export varieties may be an indicator of firm productivity increase, as pointed out by Goldberg et al. [2009]. In this paper, we will try to see whether the firm productivity increases in China and how does it contribute to the aggregate productivity change.

Then we try to compare the distribution of export revenues from each variety: we compute the Herfindahl index of revenue shares of each export variety within a firm-year observation. The Herfindahl index is computed as follows: fixing a firm-year observation, we compute the export revenue share for each HS5 product. The Herfindahl index is defined as $H = \frac{\sum_{i=1}^{N} s_i^2 - \frac{1}{N}}{1 - \frac{1}{N}}$, where s_i is the revenue share of product i and N is the number of export varieties. Figure B.1 plots the average Herfindahl index of revenue shares.[5] As we can see, the Herfindahl index increases from 0.42 to 0.44 in our data sample. Why the allocations of revenue shares become more dispersed? What is the productivity implication of the change of revenue share dispersion? We will try to explore these questions through the lens of our model.

[5] We restrict the number of export varieties to be greater than 5 to exclude the measurement errors of a few varieties within a firm.

Figure 3.3: Counts of Export Varieties per Firm



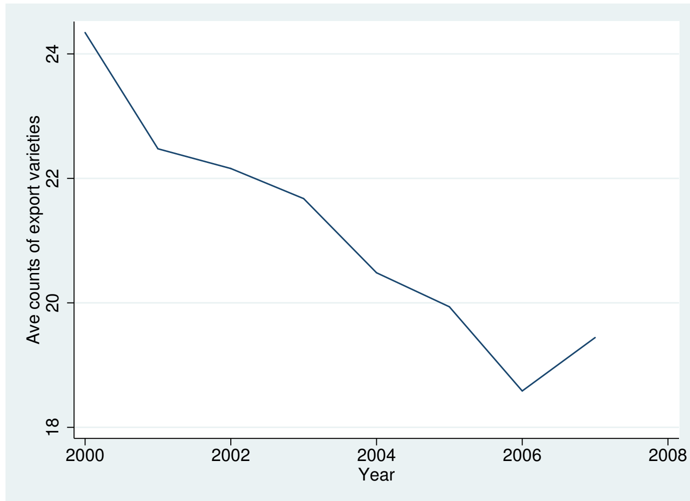
Note: This figure plots the average counts of export varieties per firm. One export variety is an HS5 product.

Fact 3: Import Varieties decreases

Figure 3.4 plots the average count of import varieties per firm. Similarly, we define one import variety as an HS5 product. Our data shows that the firms' import varieties decline from 25 in year 2000 to 18 in year 2006 and then jumps to 19 in year 2007. Overall, a firm imports fewer varieties.[6] Goldberg et al. [2009] documents that the increase of the import varieties increases the export varieties in India and along with the increase in output variety, the export share has also increased. The case of China is different. How will it change the aggregate productivity?

[6]The total number of import varieties increases, although the number of import varieties per firm decreases

Figure 3.4: Counts of Import Varieties per Firm



NOTE: This figure plots the average counts of import varieties per firm. One import variety is an HS5 product.

3.3 The Model

Motivated by the above facts, in this section we build a model of industry equilibrium in which firms use both domestic and imported intermediate goods to produce multiple products.

3.3.1 Model Setup

Production Technology

Firms are indexed by i, time is indexed by t. The feasible set of export products is $\Omega_J = \{1, 2, 3, ..., J\}$. A firm produces multiple products. The firm uses capital K_{ijt}, labor L_{ijt} and intermediate inputs to produce product j. The intermediate inputs are indexed by $n = 1, 2...\bar{N}$, where $\bar{N}$ is the total counts of intermediate input varieties.

For each product j, firm i uses capital, labor and $\bar{N}$ intermediate inputs to produce. The production function is given by

$$Q_{ijt} = \Omega_{ijt}(K_{ijt})^{\alpha_j}(L_{ijt})^{\beta_j}\prod_{n=1}^{\bar{N}}(X_{ijt}^n)^{\gamma_j^n} \tag{3.1}$$

where Ω_{ijt} denotes the firm-product specific productivity, K_{ijt}, L_{ijt} and X_{ijt}^n denotes the capital, labor and intermediate inputs n that firm i allocates to good j. We assume that the Cobb-Douglas weight α_j, β_j and γ_j^n depends on different product j. We denote $\gamma_j = \sum_{n=1}^{\bar{N}}\gamma_j^n$ as the elasticity of the intermediate inputs when producing j. The capital and labor the firm uses is $K_{it}=\sum_{j=1}^{J_{it}} K_{ijt}$ and $L_{it}=\sum_{j=1}^{J_{it}} L_{ijt}$.

Each intermdiate good X_{ijt}^{n} is assembled from a combination of a foreign and a domestic variety

$$X_{ijt}^{n}=\left[\left(B_{nt}X_{ijt}^{n,F}\right)^{\frac{\theta-1}{\theta}}+\left(X_{ijt}^{n,H}\right)^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}. \tag{3.2}$$

where $X_{ijt}^{n,F}$ and $X_{ijt}^{n,H}$ are foreign and domestic variety of intermediate good n. B_{nt} is the relative quality of foreign variety. θ is the elasticity of substitution between foreign variety and domestic variety. We assume different products share the same θ.

Assume that the prices of foreign and domestic varieties are P_{nt}^{F} and P_{nt}^{H} and denote $A_{nt}=\frac{B_{nt}P_{nt}^{H}}{P_{nt}^{F}}$ is the relative price of the domestic input adjusted by the quality of goods. If the intermediate inputs are solely come from domestic variety (such as non-tradable goods), we let $A_{nt}=0$. Following Halpern et al. [2015], we assume that A_{nt} of all traded intermediate inputs is the same across all traded inputs n, $A_{nt}=A_{t}$.

Firms pay costs to access foreign intermediate inputs. Let N_{it} denote the number of imported foreign varieties of firm i in year t. We assume the cost of getting access to foreign inputs is $F\left(N_{it}\right)$ is increasing and convex in N_{it}: firms need to pay more costs to get access to one more foreign variety.

Demand Curve

We assume that for product j, the demand curve facing by firm i is

$$\ln p_{ijt}=-\hat{\sigma}_{jt}\ln Q_{ijt}+D_{jt}+u_{ijt} \tag{3.3}$$

where D_{jt} is the demand shifter of product j, p_{ijt} is the price of product j firm i. $\hat{\sigma}_{jt}$ is the inverse demand elasticity. If $\hat{\sigma}_{jt}$ is a constant, the demand curve has a constant demand elasticity, which would imply a constant markup. There is a recent growing trade literature deviates away from the simple constant demand elasticity assumption using the non-homothetic preference. The demand curve they derived is isomohpic to the eqaution (3.3).

Firm Investment Decision

Firms can invest in innovation to upgrade the cost function of producing a product. A firm rents capital and pays for labor. To simplify the model, the firm do not accumulate capital. The uncertainty exist in future price of factor, like interest rates and wage. The potential profit from a product is determined by the revenue, wage, capital, and input expenditure. If the potential profit is negative, then the firm stops producing the product.

In the following analaysis, we neglect the footnote i and t if it does not cause any confusion. The firm problem is as follows

$$\begin{aligned} V(\{A\}_{j=1}^{J}, \{F\}_{j=1}^{J}, r, w) = & \max_{\{\Delta A_j\}_{j=1}^{J}, N} \pi\left(\{A\}_{j=1}^{J}, \{F\}_{j=1}^{J}, N, r, w\right) - \sum_{i=j}^{J} C^I(\Delta A_j) - F(N) \\ & + \beta E[V(\{b'\}_{j=1}^{J}, \{F'\}_{j=1}^{J}, r', w')] \\ s.t. \quad & A_j' = (1-\delta)A_j + \Delta A_j \\ & F_j' = F_j + \epsilon^F \\ & r' = r + \epsilon^r \\ & w' = w + \epsilon^w \end{aligned} \tag{3.4}$$

where $\pi(.)$ is the periodic profit of the firm, which will be defined later. The firm upgrades productivity by ΔA_j for product j through investing $C^I(\Delta A_j)$ in innovation in j. Next period, the TFP of product j is $(1-\delta)A_j + \Delta A_j$, as TFP discounts at rate δ. The number of feasible product sets is J. Uncertainty exists in future fixed cost of production F, interest rate r and wage w. ϵ^F, ϵ^r and ϵ^w denote the uncertainty for the input fixed cost, interest rate and wage. The firm maximizes the presented value by making dynamic decision ΔA_j and N, and static decisions in (3.5).

The static profit after the realization of product productivity is defined as follows:

$$\pi\left(\{A\}_{j=1}^{J}, \{F\}_{j=1}^{J}, N, r, w\right) = \max_{\{K_j, L_j, X_j, p_j\}_{j=1}^{J}} \sum_{j=1}^{J} \max\left\{p_j Q_j - wL_j - rK_j - \sum_{n=1}^{\bar{N}} P_n X_j^n - F_j, 0\right\} \tag{3.5}$$

Equations (3.1), (3.2) and (3.3)

where the firm determines to produce the product j if the operating profit of j is higher than 0.[7].P_n is the price of intermediate input X_j^n. If some input n is imported, the price can be derived from the aggregator equation (3.2)

$$P_n = [P_n^{H1-\theta} + (P_n^F / B_n)^{1-\theta}]^{\frac{1}{1-\theta}} = P_n^H[1 + A^{\theta-1}]^{\frac{1}{\theta-1}}$$

Otherwise, the price of X_j^n is P_n^H. In the profit equation (3.5), the first term is the sum of revenues from all products; the second and the third term are the labor cost and the intermediate inputs cost. The firm chooses the labor L_j and the input X_j^n. The restrictions include the production function (3.1) and the demand equation (3.3).

3.3.2 Model Solution

Resource Allocation within the Firm

Let R_j denote the revenue of product j. The FOCs of the above optimization (3.5) imply that

[7]is indexed by $j = 1, 2, ...J$, where J is the size of the feasible set of products to produce in t. The operating profit of product j is defined as $p_j Q_j - wL_j - rK_j - \sum_{n=1}^{\bar{N}} P_n X_j^n$

$$\alpha_j\left(1-\hat{\sigma}_{jt}\right)R_{ijt} = \Lambda_{it} r_{it} K_{ijt}$$
$$\beta_j(1-\hat{\sigma}_{jt})R_{ijt} = \Lambda_{it} w_{it} L_{ijt}$$
$$\gamma_j(1-\hat{\sigma}_{jt})R_{ijt} = \Lambda_{it} P_{nt} X^n_{ijt}$$

Thus we define ρ^K_{ijt}, ρ^L_{ijt}, ρ^n_{ijt} as the shares of capital, labor and intermediate input n that are allocated to produce product j

$$\rho^K_{ijt} = \frac{K_{ijt}}{K_{it}} = \frac{\alpha_j\left(1-\hat{\sigma}_{jt}\right)R_{ijt}}{\sum_{j=1}^{J_{it}}\alpha_{j'}\left(1-\hat{\sigma}_{j't}\right)R_{ij't}} \tag{3.6}$$

$$\rho^L_{ijt} = \frac{L_{ijt}}{L_{it}} = \frac{\beta_j\left(1-\hat{\sigma}_{jt}\right)R_{ijt}}{\sum_{j=1}^{J_{it}}\beta_{j'}\left(1-\hat{\sigma}_{j't}\right)R_{ij't}} \tag{3.7}$$

$$\rho^n_{ijt} = \frac{P_{nt}X^n_{ijt}}{P_{nt}X^n_{it}} = \frac{\gamma^n_j\left(1-\hat{\sigma}_{jt}\right)R_{ijt}}{\sum_{j=1}^{J_{it}}\gamma^n_{j'}\left(1-\hat{\sigma}_{j't}\right)R_{ij't}} \tag{3.8}$$

Similarly, we can define the share of intermediate expenditures of product j as $\rho^M_{ijt} = \frac{M_{ijt}}{M_{it}} = \frac{\sum_{n=1}^{\bar{N}} P_{nt}X^n_{ijt}}{\sum_{n=1}^{\bar{N}} P_{nt}X^n_{it}}$, we can get

$$\rho^M_{ijt} = \frac{\gamma_j\left(1-\hat{\sigma}_{jt}\right)R_{ijt}}{\sum_{j=1}^{J_{it}}\gamma_{j'}\left(1-\hat{\sigma}_{j't}\right)R_{ij't}} \tag{3.9}$$

Import Inputs Choice

Then we solve how the firm chooses the intermediate inputs to use. Within a composite intermediate input n, the optimal expenditure share on the foreign good in the total spending for variety n in product j is

$$S_t = \frac{P^F_{nt}X^{n,F}_{ijt}}{P_{nt}X^n_{ijt}} = \frac{A_t^{\theta-1}}{1+A_t^{\theta-1}}$$

which is same across all goods n, products j and firm i.

Consider a firm i and let M_{it} denote the total expenditure on intermediate inputs, and $M^{n,F}_{it}$ as the expenditure of imported inputs n, we have

$$\frac{M^F_{ijt}}{M_{it}} = \frac{M_{ijt}}{M_{it}}\frac{M^F_{ijt}}{M_{ijt}} = \rho^M_{ijt}\sum_{n=1}^{\bar{N}}\frac{\gamma^n_j}{\gamma_j}S_t 1\{\text{import} \quad \text{n}\} = \rho^M_{ijt}S_t G_j(N_{it}) \tag{3.10}$$

where $G(N_{it}) \equiv \sum_{n=1}^{\bar{N}}\frac{\gamma^n_j}{\gamma_j}1\{\text{import} \quad \text{n}\}$. Equation (3.10) says that the share of imported intermediate input of product j against the intermediate input for firm i is determined by 2 parts: the input share for product j, and the import share for product j. Further on, the share of the import input and the corresponding domestic input against the input share for product j is $G_j(N_{it})$. Among these input, the share of import is S.

Production Function

Let $\rho_{jt} = \sum_{n=1}^{\bar{N}} \frac{\gamma_j^n}{\gamma_j} \log P_{nt}^H - \sum_{n=1}^{\bar{N}} \frac{\gamma_j^n}{\gamma_j} \log \frac{\gamma_j^n}{\gamma_j}$, and the production function (3.1) can be rewritten as

$$\begin{aligned} q_{ijt} &= \alpha_j k_{it} + \beta_j l_{it} + \gamma_j (m_{it} - \rho_{jt}) + \gamma_j a_t G_j (N_{it}) + \\ & \quad \alpha_j \log \rho_{ijt}^K + \beta_j \log \rho_{ijt}^L + \gamma_j \log \rho_{ijt}^M + \omega_{ijt} \end{aligned} \tag{3.11}$$

where q, k and l denote the logs of corresponding variables. m is the log values of the interemdiate expenditures. $a_t = \frac{1}{\theta-1} \log(1 + A_t^{\theta-1})$ is a time-shifter measuring the relative technology change of foreign inputs. The first line of the equation (3.12) contains variables at the firm level: firm capital, firm labor, firm level number of imported goods. This part is very similar as the Halpern et al. [2015], with an important difference that all elasticities and function G depend on product j. This difference allows us to quantify the firm productivity change at the product level. The second line of equation (3.12) captures the unobserved variables: allocations of the resources within the firm, and productivity. Comparing with De Loecker et al. [2016], our model builds the link between multiple inputs and multiple output products, while they neglect this channel. Fixing the expenditure of intermediate inputs, when we increase the number of imported varieties, the output quantity will increase. And from equation (3.12), we can see this effect is captured by $G_j (N_{it})$. We will discuss the quantitative implication of neglecting this channel in the next section.

3.4 Estimation

In this section, we introduce our estimation strategies. First, we estimate the demand elasticity $\sigma_{jt} = \frac{1}{\tilde{\sigma}_j}$ from equation (3.3). Second, we estimate the $G_j (N_{it})$ from equation (3.10); Finally, we estimate the production function (3.12).

3.4.1 Estimating the demand elasticity

We first estimate the demand curve (3.3). For simplicity, we assume that σ_{jt} follows

$$\sigma_{jt} = \begin{cases} \sigma_j & \text{if } t \leq 2003 \\ \sigma_j' & \text{if } t > 2003 \end{cases}$$

where the year 2003 is the time when China joined the WTO. To estimate the demand curve, the classical endogeneity problem will arise since the price change may reflect marginal cost difference as well as preference change. Following Wei et al. (2017), we use the average relative price of the firm in other markets, input material's price deflator, log capital, log labor, log material. import variety counts and export variety counts as the instrument variables. The idea is those variables are related to the marginal cost change rather than the consumers' preference change. The detail refers to Appendix B.1. Table B.1 and Table

B.2 show the estimations results of two stages by 2SLS. Figure B.2 show the distribution of demand elasticities. the overall distribution is close to a normal distribution. The frequency of demand elasticity before 2003 reaches largest when the elasticity is 2.7. After 2003, the frequency reaches largest when elasticity is 3.2. The gap between the two period indicates a more competitive market, namely, lower markup.

3.4.2 Estimating the marginal benefit of increases in input variety

In the second step, we estimate equation (3.10). We first compute r_{ijt} by using $r_{ijt} = \frac{(1-\hat{\sigma}_{jt})R_{ijt}}{\sum_{j'=1}^{J_{it}}(1-\hat{\sigma}_{j't})R_{ij't}}$. Then we then assume a parametric functional form

$$G_j(N) = \begin{cases} \bar{G}_j \left(1 - [1 - \left(\frac{N}{\bar{N}_I}\right)^{\lambda_j}]^{\frac{1}{\lambda_j}}\right) & \text{if } N \leq \bar{N}_I \\ \bar{G}_j \quad \text{if } N > \bar{N}_I \end{cases}$$

Here $\lambda_j \in (0,1)$ and $\bar{G}_j \in (0,1)$. This functional form implies that when number of varieties increases, the marginal benefit will decline. And there is a cutoff value, if the import varieties exceed $\bar{N}_I$, the marginal benefit of expanding varieties declines to 0. $\bar{N}_I$ is the total number of traded varieties in the market. If a firm's number of imported varieties equals to $\bar{N}_I$, $\bar{G}_j$ equals to the share of total imported share in the intermediate inputs.

There are four groups of unknowns to estimate S_t, $\bar{G}_j$, λ_j and γ_j. However, γ_j and $\bar{G}_j$ can not be separately identified because they enter equation (3.10) in the same way. We normalize $\bar{G}_j$ to be 0.8 to match the aggregate total imported share in the intermediate inputs from China's input-output table. We then estimate the nonlinear equation to get λ_j. The estimation results γ_j will be ignored in this step. We will estimate γ_j in the next step.

The estimation of S_t is reported in Table B.5 and the estimation of λ_j is reported in Table B.6. There is a declining trend: S_t drops from 0.696 in year 2000 to 0.520 in year 2006. Since the import share is declining on average, it must suggest that there is a growing technology improvement of domestic material goods, which drives down the relative price of domestic goods adjusted by the quality. The value is between 0.455 and 1. The average λ_j is 0.783.

3.4.3 Estimating the production function

In the third step, we estimate the equation (3.12) following the methodology of Olley and Pakes (1994). The estimation equation becomes

$$\begin{aligned} q_{ijt} &= \omega_{ijt} + \alpha_j k_{it} + \beta_j l_{it} + \gamma_j (m_{it} - \rho_{jt}) + \gamma_j a_t G_j(N_{it}) \\ &\quad + h(r_{ijt}, r_{i-jt}) \end{aligned}$$

where $h(r_{ijt}, r_{i-jt})$ denotes $\alpha_j \log \rho_{ijt}^K + \beta_j \log \rho_{ijt}^L + \gamma_j \log \rho_{ijt}^M$, which are functions of (r_{ijt}, r_{i-jt}).

Following Olley and Pakes [1996], there are two problems we should solve: (i). the simultaneity between input and the unobserved productivity shock, (ii) the selection problem, which can be understood that the correlation of product productivity across time is not independent. To solve the first problem, we use exogenous variables, including capital, number of input varieties, number of export varieties, as instrument variable for the unobserved productivity shock. To overcome the selection problem, we use a condition

$$E[\omega_{ijt} - f(\omega_{ij,t-1})|\omega_{ij,t-1}] = 0 \tag{3.12}$$

The key difference from the single product case is that we need instruments for $\frac{r_{ijt}}{r_{it}}$.

Let's define the known variable for estimation ω_{ijt} as x_{ijt}

$$x_{ijt} = (k_{it}, l_{it}, m_{it}, \{r_{ijt}\}_{j=1}^{J}, a_t, G_j(N_{it}))$$

It is obvious that $\hat{\omega}_{ijt}$ can be written as $\Omega(x_{ijt}) = q_{ijt} - [\omega_{ijt} + \alpha_j k_{it} + \beta_j l_{it} + \gamma_j (m_{it} - \rho_{jt}) + \gamma_j a_t G_j (N_{it}) + h(r_{ijt}, r_{i-jt})]$. By substituting the detail form of $\Omega(x_{ijt})$ and $\Omega(x_{ijt})$ into the moment condition,The conditional moment restriction we utilize for estimation is given by

$$\begin{aligned} g(x_{ijt}; \theta) &= E\left[\Omega(x_{ijt}) - f(\Omega(x_{ij,t-1}))|x_{ijt}, x_{ij,t-1}; \theta\right] \\ &= 0 \end{aligned}$$

Table B.6 reports the estimation of α_j, β_j and γ_j. On average the capital elasticity is not large, around 0.038. The largest and the smallest value of $\alpha_j' s$ are 0.175 and 0.014, respectively. The next column reports the point estimation of β_j, labor elasticity. On average of the labor elasticity is 0.07, the largest and the smallest values of which is 0.079 and 0.017. In addition, the average of γ_j, input elasticity, is 0.17. Overall, the production function shows strong decreasing returns to scale (0.038+0.07+0.17=0.279).

3.4.4 Recovering other parameters

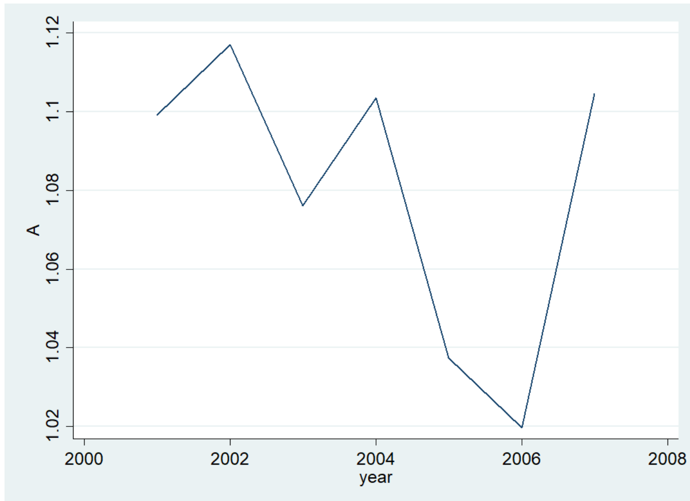
From the estimation of equations (3.10), we get $S_t = \frac{A_t^{\theta-1}}{1+A_t^{\theta-1}}$. In the equation (3.12), we replace a_t by $a_t = \frac{1}{\theta-1}\log(1 + A_t^{\theta-1}) = \frac{1}{\theta-1}\log(1 + \frac{S_{jt}}{1-S_{jt}})$. Then from the parameters in the equation (3.12), we back up the substitution parameters θ.

The A_t is identified from the following equation

$$A_t = \frac{1}{\theta - 1}\log\left(\frac{S_t}{1 - S_t}\right) \tag{3.13}$$

The dynamics of the estimated result is given in Figure 3.5.

Figure 3.5: The dynamics of relative efficiency of import against domestic input (A_t)



Note: A_t measure the relative quality of import input against the domestic input. For example, $A_{2002} = 1.12$, which means that the quality of the import input is 1.12 times of the domestic input.

3.5 Results

3.5.1 Firm Productivity

The revenue productivity of each firm is defined as

$$TFPR_{it} = \sum_{j=1}^{J_{it}} s_{ijt} p_{ijt} \exp(\gamma_j a_t G_j (N_{it})) \Omega_{ijt} \tag{3.14}$$

where s_{ijt} denotes the revenue share of product j against firm i's sales in year t.

There are three parts in the revenue productivity: (1) the weight of products in sales for the firm, which indicates the resource reallocation within the firm;(2) the access to the foreign inputs, $\exp(\gamma_j a_t G_j (N_{it})$; and (3) the firm-product real productivity Ω_{ijt}. Figure B.3, Figure B.4 and Figure B.5 plot the estimation of the log revenue productivity. Table B.7 reports the summary statistics of the estimated log TFPR. The average of the log TFPR is around 5.5 and the standard deviation is 1.4.

We define the aggregate revenue productivity as the weighted average of each firm's TFPR

$$\overline{TFPR}_t = \sum_i \frac{R_{it}}{R_t} TFPR_{it}$$

where $\frac{R_{it}}{R_t}$ is the revenue share of firm i.

We are interested in the change of $\overline{TFPR}$. The first row of table 3.5.1 reports the percentage change of aggregate $\overline{TFPR}$ between year 2002 and 2007. The aggregate TFP increases by 47.1%. We then decompose $\overline{TFPR}$ as follows:

(1) We fix the firms who survive through the year 2002 to 2007 and call those firms as set I. The aggregate productivity change of firms in set I is defined as

$$\ln \overline{TFPR}_{I,2007} - \ln \overline{TFPR}_{I,2002} = \ln \sum_{i \in I} \frac{R_{i,2007}}{R_{2007}} TFPR_{i,2007} - \ln \sum_{i \in I} \frac{R_{i,2002}}{R_{2002}} TFPR_{i,2002} \tag{3.15}$$

(2) For those firms in set I, we first fix the resource allocation within the firm $\exp(\alpha_j \log \rho_{ijt}^K + \beta_j \log \rho_{ijt}^L + \gamma_j \log \rho_{ijt}^M)$ and the firm-product real productivity Ω_{ijt}, using the value in year 2002. At the same time, we also fix the product set J_{it} as well. We only allow the channel of accessing to foreign inputs. The counter-factual productivity of each firm by allowing only the import channel is defined as

$$\begin{aligned} TFPR_{i,2007}^{IMP} = & \sum_{j=1}^{J_{i,2002}} p_{ij,2007} \exp(\gamma_j a_{2007} G_j (N_{i,2007})) \\ & \times \exp(\alpha_j \log \rho_{ij,2002}^K + \beta_j \log \rho_{ij,2002}^L + \gamma_j \log \rho_{ij,2002}^M) \Omega_{ij,2002} \end{aligned}$$

We can get the aggregate productivity change from equation (3.15), replacing $TFPR_{i,2007}$ with $TFPR_{i,2007}^{IMP}$. The aggregate productivity change from this step is contributed to access to foreign inputs.

(3) We then allow the Ω_{ijt} to change and define the productivity as

$$\begin{aligned} TFPR_{i,2007}^{\Omega} = & \sum_{j=1}^{J_{i,2002}} p_{ij,2007} \exp(\gamma_j a_{2007} G_j (N_{i,2007})) \\ & \times \exp(\alpha_j \log \rho_{ij,2002}^K + \beta_j \log \rho_{ij,2002}^L + \gamma_j \log \rho_{ij,2002}^M) \Omega_{ij,2007} \end{aligned}$$

The aggregate productivity change from this step is contributed to the technology upgrade.

(4) Next, we allow the resource allocation within the firm to change, but we still fix the product set J_{it}. We contribute the productivity change to the intensive margin change of allocating resources to existing products.

$$\begin{aligned} TFPR_{i,2007}^{Intensive} = & \sum_{j=1}^{J_{i,2002}} p_{ij,2007} \exp(\gamma_j a_{2007} G_j (N_{i,2007})) \\ & \times \exp(\alpha_j \log \rho_{ij,2007}^K + \beta_j \log \rho_{ij,2007}^L + \gamma_j \log \rho_{ij,2007}^M) \Omega_{ij,2007} \end{aligned}$$

(5) We now allow the entry and exit of products. The productivity change is called the extensive margin change of allocating resources to existing products.

$$\begin{aligned} TFPR_{i,2007}^{Extensive} = & \sum_{j=1}^{J_{i,2007}} p_{ij,2007} \exp(\gamma_j a_{2007} G_j (N_{i,2007})) \\ & \times \exp(\alpha_j \log \rho_{ij,2007}^K + \beta_j \log \rho_{ij,2007}^L + \gamma_j \log \rho_{ij,2007}^M) \Omega_{ij,2007} \end{aligned}$$

(6) Finally, we allow the entry and exit of firms. That is we do not restrict firms in the set I.

In the above analysis, we do not fix the prices. We define the real productivity (TFPQ) of each firm in a similar way as equation (3.14), but we change the price p_{ijt} into the average price of product j. Then the productivity change does not include firm level markup change any more. The difference between real productivity and the revenue productivity is defined as the change due to the markup. We can decompose the real productivity change in a similar way.

Table 3.5.1 reports the decomposition results. The first column reports the decomposition results of the aggregate TFPR. The aggregate TFPR increases by 47.1% from year 2002 to 2007. The second to the fifth row reports the decomposition of step 1 to step 5. We can see that the access to foreign inputs contributes very little to the aggregate productivity change, 0.05%. The most significant contribution at the firm level comes from firm entry and exit channel (28.5%) which is consistent with the literature. The new channels in this paper are upgrade channel, intensive channel and extensive channel. The entry and exit of products (extensive margin) and reallocation of resources (intensive margin) among existing products contribute 4.45% and 7.2% respectively. The upgrade of the technology is also importance. It increases aggregate productivity by 12.8%. In total, the productivity increase of the surviving firms improves the aggregate TFPR by 22%. The between-firm effect is also provided, -5.96%. It indicates that the size of low-TFP firms grow, which might be the evidence for misallocation.

Table 3.1: Decomposition of growth of firm-level TFPR for all firms

Item	TFPR growth
import effect	0.05%
Upgrade	12.8%
Reallocation	7.20%
Surviving firm switch	4.45%
Firm entry/exit	28.5%
Between firm	-5.96%
Aggregate	47.1%

3.5.2 Ownerships and TFPR growth decomposition

As heterogeneity in the management, information, and preference exist in the ownerships, we decompose the TFPR growth by the type of ownerships, including SOE, private firms and foreign firms. Among them, foreign firms achieves the largest growth, 41.6%. the second growth is achieved by private firms, 35.3%, while SOE increases by only 1.65%.

With the product-level TFPR, we can compare the growth by technology upgrade, reallocation, and switch for the surviving firms. Among the three behaviors, upgrade contributes most. Private firm's continuing exporting product experience 20.1% in growth while foreign firms and SOE grows by around 10%. In addition, the reallocation effects for private firms and foreign firms are both larger than the SOE. It indicates that high productivity product within the firm of private type and foreign type experience larger growth in revenue size. For the switch effect, the aggregate TFP of foreign firms, SOE, and privates firms grow by 3.62%, 3.12% and 2.82% in the channel.

For the entry and exit channel at the firm level, foreign firms and privates firms grow by 20.2% and 13.2%, respectively. But SOEs grow by only 0.43% through the channel. For the Between-firm channel, the growth for foreign firms, private firms and SOEs are -0.21%, -6.54%, and -16.4%, respectively. It indicates that, productive firms of the three types all experience market losing in the aggregate level. Further on, the productive firms of SOE loses most.

Table 3.2: Decomposition of growth of firm-level TFP for SOE firms

Item	SOE	Private firms	Foreign firms
Upgrade	10.9%	20.1%	12.6%
Reallocation	3.47%	5.84%	5.36%
Surviving firm switch	3.12%	2.82%	3.62%
Firm entry/exit	0.43%	13.2%	20.21%
import effect	0.08%	-0.06%	0.016%
Between firm	-16.4%	-6.54%	-0.21%
Aggregate	1.65%	35.3%	41.6%

3.6 Conclusion

This paper studies how the trade liberalization in China changes the firm productivity. We develop a framework to estimate revenue productivity (TFPR) and real productivity (TFPQ) with multi-product firms. We find that the aggregate TFPR increases 47.1% from 2002-2007 and TFPQ increases 60%, suggesting that the observed TFPR increase is mainly driven by real productivity change rather than the markup change. We further decompose the change of productivity into three channels: (1) access to foreign inputs;(2) resource re-allocation within the firm; (3) technology upgrade. We find the most significant channel is the last one, which explains half of the aggregate productivity increase. We also find that the both foreign and private firms significantly improve the TFPR while SOEs do not. The most important channel for foreign firms growth is firm entry and exit channel. Moreover, upgrade is the most important channel for privates firms and SOE.

Bibliography

P. Aghion, N. Bloom, R. Blundell, R. Griffith, and P. Howitt. Competition and innovation: An inverted-u relationship. *The quarterly journal of economics*, 120(2):701–728, 2005.

P. Aghion, P. Howitt, and S. Prantl. Patent rights, product market reforms, and innovation. *Journal of Economic Growth*, 20(3):223–262, 2015.

J. S. Ang, Y. Cheng, and C. Wu. 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, 2014.

N. Bloom, J. Van Reenen, et al. Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, 122(4):1351–1408, 2007.

R. Blundell, R. Griffith, and J. Van Reenen. Market share, market value and innovation in a panel of british manufacturing firms. *Review of Economic Studies*, 66(3):529–554, 1999.

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

J. De Loecker. Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, 79(5):1407–1451, 2011.

J. De Loecker, P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik. Prices, markups and trade reform. *Econometrica*, 84(2), 2016.

A. V. Deardorff. Welfare effects of global patent protection. *Economica*, pages 35–51, 1992.

E. Dhyne, A. Petrin, V. Smeets, and F. Warzynski. Multi-product firms, import competition, and the evolution of firm-product technical efficiencies. *Unpublished, University of Aarhus*, 2016.

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

P. Goldberg, A. Khandelwal, N. Pavcnik, and P. Topalova. Trade liberalization and new imported inputs. *The American Economic Review*, 99(2):494–500, 2009.

G. M. Grossman and E. Helpman. *Innovation and growth in the global economy*. MIT press, 1991.

G. M. Grossman and E. L.-C. Lai. International protection of intellectual property. *American Economic Review*, 94(5):1635–1653, 2004.

L. Halpern, M. Koren, and A. Szeidl. Imported inputs and productivity. *The American Economic Review*, 105(12):3660–3703, 2015.

A. R. Hashmi. Competition and innovation: The inverted-u relationship revisited. *Review of Economics and Statistics*, 95(5):1653–1668, 2013.

E. Helpman. Innovation, imitation, and intellectual property rights. *Econometrica*, 61(6): 1247–1280, 1993.

C.-T. Hsieh and P. J. Klenow. Misallocation and manufacturing tfp in china and india. *The quarterly journal of economics*, 124(4):1403–1448, 2009.

T. J. Klette and S. Kortum. Innovating firms and aggregate innovation. *Journal of political economy*, 112(5):986–1018, 2004.

R. Lentz and D. T. Mortensen. An empirical model of growth through product innovation. *Econometrica*, 76(6):1317–1373, 2008.

R. Lentz and D. T. Mortensen. Optimal growth through product innovation. *Review of Economic Dynamics*, 19:4–19, 2016.

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

S. J. Nickell. Competition and corporate performance. *Journal of Political Economy*, 104 (4):724–746, 1996.

G. S. Olley and A. Pakes. The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297, 1996. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2171831.

Y. Qian. 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, 2007.

S. J. Redding, A. B. Bernard, and P. K. Schott. Multi-product firms and product switching. 2006.

D. Restuccia and R. Rogerson. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics*, 11(4):707–720, 2008.

P. M. Romer. Endogenous technological change. *Journal of political Economy*, 98(5, Part 2):S71–S102, 1990.

F. M. Scherer. Research and development resource allocation under rivalry. *Quarterly Journal of Economics*, 81(3):359–394, 1967.

J. Schumpeter. Capitalism, socialism and democracy. 1943.

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

Y. Yao. The chinese growth miracle. In *Handbook of economic growth*, volume 2, pages 943–1031. Elsevier, 2014.

J. Zheng, A. Bigsten, and A. Hu. Can china's growth be sustained? a productivity perspective. *World Development*, 37(4):874–888, 2009.

Appendix A

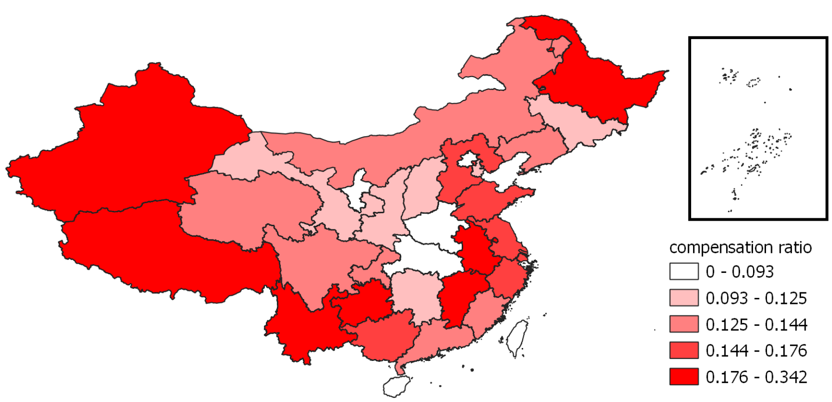
Appendix for Chapter 2

A.1 Appendix to Section 2

A.1.1 Construction of industry-province level IPR enforcement

To achieve infomation on IPR enforcement by industry and location, I finish following steps. First, I retain the cases related to IPR (Details of the related cases are outlined in Appendix A.1.3). There are 260,000 cases belonging to these causes of IPR. Second, I link litigants to firms in ASIEC by name, then the information regarding industry is mapped to litigants. Third, I identify province from the name of the court. By the information of industry and province, I aggregate the case-level observables into industry-province level.

Figure A.1: The degree of IPR enforcement in China



Notes: The calculation of the index of the provincial IPR enforcement follows equation (2.1).

A.1.2 Sample selection

One challenge with using the ASIEC data is that the classification of industry experienced three versions.[1] To be consistent with the industry classification, I retain the sample with version GB2002. Another problem is that the information of R&D expenditure exists only in year 2001, 2005, 2006, 2007 and 2010. Thus I retain the sample where R&D exists. To avoid poor enforcement measure by a small number of lawsuits in each industry-province observation, I aggregate firm-level observables by 2-digit industry and province.

After constructing an industry-province level dataset, I impose a few sample restrictions. From Appendix A.1.4, the aggregated ratio of cases in 13 provinces with the least counts of cases is less than 4% (8426 cases). The size of the sample in these provinces is too small to reflect local government's real attitude to IPR protection. In addition, I restrict my analysis to provinces with more than 2000 related cases, responding to 98% of firms in year 2006 recorded in ASIEC.[2]

A.1.3 Description of lawsuits related to IPR

There are 10 sections of classifications of causes of action of civil cases in China. Among them, second-level classifications in section 5, (1) disputes over the contract of IPR (intellectual property right) and (2) disputes over the ownership or infringement of IPR, are regarded to be related with IPR. Since the variable *cause of action* recorded in China Judgements Database maps to the third-level item, I use keywords from the third-level items in search of lawsuits related to IPR. Table A.2 reports the result.

Table A.1: Summary of lawsuits by third-level classification

Cause of action	Count
Disputes over copyright ownership and infringement	154,118
Disputes over Trademark ownership and infringement	66,575
Disputes over Patent ownership and infringement	27,638
Disputes over copyright contract	8,446
Disputes over contract of franchising	4,438
Disputes over contract of technology	3,396
Others	2,295
Aggregate	266,906

[1] GB1994 is applied to sample of year 1996-2002, GB2002 is applied to sample of year 2003-2012, and GB2011 is applied to sample of year 2013

[2] 6508 firms are excluded from the whole sample of 301,648 in year 2006 from ASIEC

Table A.2: Summary of lawsuits by third-level classification (ASIEC)

Variable	Count	Processing fee	Claim	Compensation	Compensation ration
Cause of action					
Disputes over copyright ownership and infringement	4109	1607	216528	6565	3.78%
		(8148)	(1540491)	(56719)	(12.19%)
Disputes over Trademark ownership and infringement	29446	1024	82766	21195	7.99%
		(3535)	(456658)	(2333615)	(16.85%)
Disputes over Patent ownership and infringement	7891	4156	493225	21560	6.70%
		(15631)	(3002485)	(281184)	(17.02%)
Disputes over copyright contract	262	2497	239010	33430	6.48%
		(5810)	(632877)	(235582)	(18.18%)
Disputes over contract of franchising	218	11192	1051268	533636	38.26%
		(15049)	(1824006)	(1509189)	(45.71%)
Disputes over contract of technology	538	10640	1058347	349098	26.79%
		(17739)	(2316798)	(1407192)	(40.03%)
Others	207	9061	1001825	150523	13.57%
		(23473)	3675252	(863880)	(31.29%)
Aggregate	42671	1856	188517	27612	7.79%
		(8356)	(1442864)	(1953882)	(17.7%)

A.1.4 Description of IPR enforcement index by province or industry

Table A.3: Summary statistics of lawsuits by province

Variable	Ratio of observations		Compensation ratio	
Sample	All	ASIEC	All	ASIEC
Provinces				
Zhejiang	19.48%	15.53%	7.27%	7.62%
Guangdong	16.94%	11.90%	10.14%	11.73%
Beijing	8.00%	4.66%	3.03%	1.91%
Hubei	7.94%	4.22%	2.50%	8.38%
Jiangsu	7.10%	11.72%	11.37%	7.60%
Shandong	6.70%	11.46%	5.27%	7.05%
Shanghai	4.77%	3.79%	5.23%	5.11%
Chongqing	3.91%	0.94%	2.39%	9.68%
Hunan	3.88%	5.54%	5.54%	10.42%
Fujian	3.60%	6.33%	6.49%	6.35%
Sichuan	3.21%	3.10%	4.41%	7.51%
Henan	3.05%	3.49%	2.48%	3.20%
Anhui	3.00%	5.62%	6.36%	9.90%
Tianjin	1.35%	0.97%	4.59%	5.45%
Shaanxi	1.12%	2.53%	7.07%	6.61%
Liaoning	1.05%	0.93%	8.86%	6.66%
Guangxi	0.87%	1.10%	14.36%	10.22%
Hebei	0.86%	1.45%	7.95%	5.68%
Jilin	0.72%	1.30%	8.09%	8.29%
Jiangxi	0.64%	0.98%	5.33%	6.19%
Xinjiang	0.40%	0.37%	16.05%	13.06%
Shanxi	0.34%	0.62%	11.49%	12.52%
Yunnan	0.33%	0.43%	22.02%	11.01%
Heilongjiang	0.18%	0.26%	18.61%	19.86%
Gansu	0.14%	0.19%	9.68%	5.62%
Guizhou	0.13%	0.26%	14.51%	10.57%
Neimenggu	0.12%	0.12%	16.75%	7.23%
Ningxia	0.08%	0.13%	5.48%	6.55%
Hainan	0.04%	0.02%	2.82%	13.50%
Qinghai	0.03%	0.06%	7.56%	6.29%
Xizang	0.00%	0.00%	8.08%	NA
Agg	266,906	46,000		

Table A.4: Summary statistics of lawsuits by industry (ASIEC)

Variable	Count	Observation ratio	Compensation ratio
Industry			
Chemestry	7079	16.59%	7.35%
Education	5132	12.03%	6.62%
Beverage	3779	8.86%	12.06%
Metal Product	3500	8.20%	6.35%
Electronic	3269	7.66%	8.66%
Computer	2535	5.94%	8.41%
Apparel	2070	4.85%	8.37%
Textile	2021	4.74%	5.78%
General Manufacture	1693	3.97%	6.11%
Transport	1599	3.75%	8.19%
Leather	1372	3.22%	11.11%
Non-metal Product	1033	2.42%	5.35%
Furniture	855	2.00%	7.51%
Plastics	809	1.90%	4.75%
Nuclear and other	778	1.82%	7.28%
Medicine	739	1.73%	7.30%
Special equimpment	698	1.64%	9.85%
Food	677	1.59%	6.60%
Food processing	639	1.50%	8.13%
Office	636	1.49%	5.04%
Paper	556	1.30%	7.51%
Print	447	1.05%	4.96%
Other	755	1.77%	

A.1.5 Summary statics for variables in regressions

Table A.5: Summary statics for variables in regressions

Variable	Obs	Mean	SD	Mininum	25th Percentile	Median	75th Percentile	Maximum
average firm-level revenue (y_{pjt})	3,197	2.09×10^5	8.86×10^5	1444	3.75×10^4	6.27×10^4	1.19×10^5	2.44×10^7
average firm-level R&D (RD_{pj})	3,197	1280	1.70×10^4	-36.2	24.1	102	506	8.22×10^5
R&D intensity	3197	0.414%	0.697%	−0.12%	0.0478%	0.150%	0.480%	12%
competition (c_{pjt})	3197	3.03	1.30	0	2.14	3.06	3.92	6.86
compensation ratio (e_{pj})	696	0.094	0.15	0	0.01	0.055	0.11	0.97
winning rate(w_{pj})	733	0.282	0.270	0	0.066	0.236	0.400	1
infringe (s_{pj})	1609	0.187	0.256	0.0198	0.0637	0.115	0.170	1.185
average compensation(F_{pj})	703	50.6	260	0	0	5	20	5482

Note the unit of revenue, R&D and compensation is 1000 RMB

Table A.6: Correlation matrix

	1	2	3	4	5	6	7	8
1.$log(y_{pj})$	1							
2.$log(RD_{pj})$	0.747***	1						
3.RD intensity	0.0135	0.426***	1					
4.c_{pj}	0.03	0.0371	0.0144	1				
5.e_{pj}	0.0319	0	-0.0512*	0.011	1			
6.w_{pj}	-0.0617**	-0.117***	-0.0571*	0	0.561***	1		
7.s_{pj}	-0.360***	-0.161***	-0.0602**	-0.00467	-0.0345	-0.03923	1	
8.$log(F_{pj})$	0.154***	0.117***	0.0634*	0.0211	0.499***	0.277***	-0.194***	1

A.1.6 Alternative reduced form evidence

Folowing Aghion (2015), I compare the effect of competition on innovation across two groups: (1) provinces where plaintiff has a high probability in winning in a lawsuit of IPR; (2) provinces where the probability of winning is low. I estimate the following equations:

$$log(y_{jpt}) = \alpha_0 + \alpha_1 \times c_{jpt} \times G_p + \alpha_2 \times c_{jpt} \times (1 - G_p) \tag{A.1}$$

$$log(y_{jpt}) = \beta_0 + \beta_1 \times c_{jpt} + \beta_2 \times c_{jpt} \times G_p \tag{A.2}$$

where y_{jpt} is the RD expenditure in province p and industry j in year t; c_{jpt} is the competition index; G_p equates 1 if province p belongs to the group with high winning rate; otherwise, G_p is 0.

One concern about the impact of competition on innovation is the reverse causality: more innovation in a industry leads to more advanced firms which will change the market structure. I use competition index in last year as instruments to address this endogeneity problem. Another concern is also a reverse causality : the improvement of R&D leads to a higher demand for IPR protection. To address this endogeneity, I construct a proxy $\hat{G}_i$ for G_i. I take the average winning rate by plaintiff of neighbour provinces to construct this new proxy. This variable is related to local IPR enforcement because of the geography proximity but has little impact on local firms' RD decision.[3] Table A.7 reports the estimation result, which is consistent with the reduced form evidence in section 2.3.3.

Table A.7: Alternative reduced form evidence

Regression Model	OLS	OLS	IV	OLS	OLS	IV
Dependent Variable	Log average R&D expenditure			Log R&D intensity		
$c * G_i$	1.59^b			0.628		
	(0.691)			(0.642)		
$c * \hat{G}_i$		1.72^b	2.29^a		0.370	0.823
		(0.737)	(0.801)		(0.687)	(0.747)
c	-3.63^a	-3.97^a	-8.33^a	-2.740^a	-2.719^a	-6.00^a
	(0.891)	(0.944)		(0.828)	(0.880)	(2.310)
Observations	1531	1531	1435	1531	1531	1435
R-squared	0.69	0.69	0.69	0.60	0.60	0.60

a and b indicate significant at the 1% and 5% level, respectively. In parentheses, I report robust standard errors for coefficients. I include year, province, industry fixed effects in each regression.

A.2 Appendix to Section 5

Solution for Cournot competition

In a market, the unique innovating products meets n counterfeiters producing the same quality, what is the equilibrium price ?

- the demand curve $p = \frac{E}{q}$

[3] $\hat{G}_i = 1$ if the average neighbours' winning rate is higher than the median; 0, otherwise.

- The firm's profit maximization: $\max_{p,q} pq - cq$, s.t.$p = \frac{E}{q+nq*}$, where $q*$ is produced by other firms
- Then we can solve $p = (1 + \frac{1}{n})c$, profit rate $\bar{\pi}(n) = \frac{1}{n+1}$
- It is obvious that the profit rate decreases with n.

The details in value function

- value of innovation

$$\Delta V^{RD} \equiv \underbrace{[V(\Omega^I \cup \{0\}, \Omega^{\bar{m}}, \Omega^m) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)]}_{\text{value from innovation}}\}\epsilon_1$$

- value of imitation

$$\Delta V^{Imitation} \equiv \underbrace{h[\mathbb{E}_j V(\Omega^I, \Omega^{\bar{m}}, \Omega^m \cup \{x\}) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)]}_{\text{value change from imi}}$$

- value of suing

$$\Delta V_j^{sue} \equiv \underbrace{[V(\Omega^I, \Omega^{\bar{m}} - \{n_j\} \cup \{n_j - 1\}, \Omega^m) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)]}_{\text{value change from winning the lawsuit}}$$

- innovative product meets destruction

$$\Delta V_j^{I,destruction} \equiv [V(\Omega^I - \{0_j\}, \Omega^{\bar{m}}, \Omega^m) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)]$$

- imitated product meets destruction

$$\Delta V_j^{\bar{m},dest} \equiv V(\Omega^I, \Omega^{\bar{m}} - \{n_j\}, \Omega^m) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)$$

- counterfeiting product meets destruction

$$\Delta V_j^{m,dest} \equiv V(\Omega^I, \Omega^{\bar{m}}, \Omega^m - \{n_j\}) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)$$

- the innovative product meets infringement

$$\Delta V^{I,infrig} \equiv V(\Omega^I - \{0_j\}, \Omega^{\bar{m}} \cup \{1_j\}, \Omega^m) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)$$

- the infringed product meets infringement again

$$\Delta V_j^{\bar{m},infrig} \equiv V(\Omega^I, \Omega^{\bar{m}} - \{n_j\} + \{n_j + 1\}, \Omega^m) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)$$

- counterfeiter's imitating product meets infringement again

$$\Delta V_j^{m,infrig} \equiv V(\Omega^I, \Omega^{\bar{m}}, \Omega^m - \{n_j\} + \{n_j + 1\}) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)$$

- The counterfeiting product lose the lawsuit

$$\Delta V_j^{m,lawsuit} \equiv V(\Omega^I, \Omega^{\bar{m}}, \Omega^m - \{n_j\}) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)$$

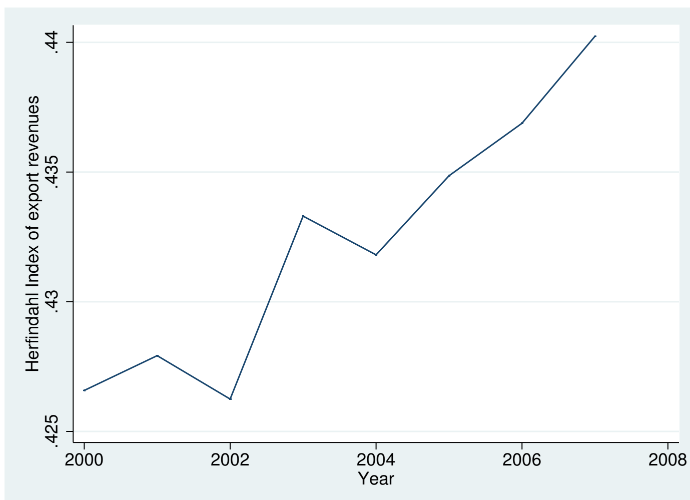
- The other $n_j - 1$ imitator lose the lawsuit

$$\Delta V_j^{m,lawsuit,2} \equiv V(\Omega^I, \Omega^{\bar{m}}, \Omega^m - \{n_j\} + \{n_j - 1\}) - V(\Omega^I, \Omega^{\bar{m}}, \Omega^m)$$

Appendix B

Appendix for Chapter 3

Figure B.1: Average Herfindahl Index of Export Revenues Within the Firm



NOTE: This figure plots the average Herfindahl Index of export revenues within a firm. We restrict the number of export varieties greater than 5.

B.1 Demand side

B.1.1 Estimation problem

$$lnP_{ijy} = \sigma_{jt} lnQ_{ijy} + Q_{jy} + D_{jt} + D_y + \epsilon_{ijy} \tag{B.1}$$

The estimation problem is that ϵ should be irrelevant to lnQ_{ijt}. However the observed (Q^*_{ijt}, P^*_{ijt}) is determined by the demand curve and supply curve together. Thus the observation Q^*_{ijt} is related to ϵ_{ijt}. Thus I should use IV method.

B.1.2 IV-stage 1

$$lnQ_{ijy} = \beta_0 ln\hat{Q}_{ijy} + Q_{jy} + \beta_1 D_{jt} + \beta_2 D_{it} + D_y + \alpha \quad Firm.control + \epsilon^Q_{ijy} \tag{B.2}$$

IV variable :

1. Firm level : $lnK_{iy}, lnL_{iy}, lnM_{iy}, lnN_{iy}, lnJ_{iy}$
2. Firm-product level : $ln\hat{P}_{ijy}$

$$ln\hat{Q}_{ijy} = \frac{\sum_{j' \neq j} lnQ_{ijy}}{N_i - 1} \tag{B.3}$$

Here, t is only different in $(year < 2003, year \geq 2003)$

Table B.1: Estimation in stage 1

	lnq_{ijt}
constant	-2.06
	(0.000)
average price of $(i, -j, t)$	0.202
	(0.000)
lnQ_{jt}	-0.028
	(0.001)
k_{it}	-0.008
	(0.148)
l_{it}	0.013
	(0.005)
m_{it}	0.0893
	(0.000)
J_i	-0.021
	(0.000)
N_i	-0.003
	(0.000)
hs5 × dum2003	Yes
nfid × dum2003	Yes
year dummy	Yes
R^2	0.93
observations	156,075

Then we generate the $lnQ^{IV}_{ijy} = lnQ_{ijy} - \epsilon^Q_{ijy}$

B.1.3 IV-stage 2

$$lnP_{ijy} = \sigma_{jt} lnQ^{IV}_{ijy} + Q_{jy} + D_{jt} + D_y + \epsilon_{ijy} \tag{B.4}$$

Table B.2: Estimation in stage 1

	lnq_{ijt}
constant	0.648
	(0.022)
lnQ^{iv}_{ijt}	-0.205
	(0.000)
hs5 $\times$ lnQ^{iv}_{ijt}	
D2003 $\times$ lnQ^{iv}_{ijt}	0.0715
	(0.005)
hs5 dummy	Yes
year dummy	Yes
R^2	0.827
observations	156,075

B.1.4 Demand elasticity

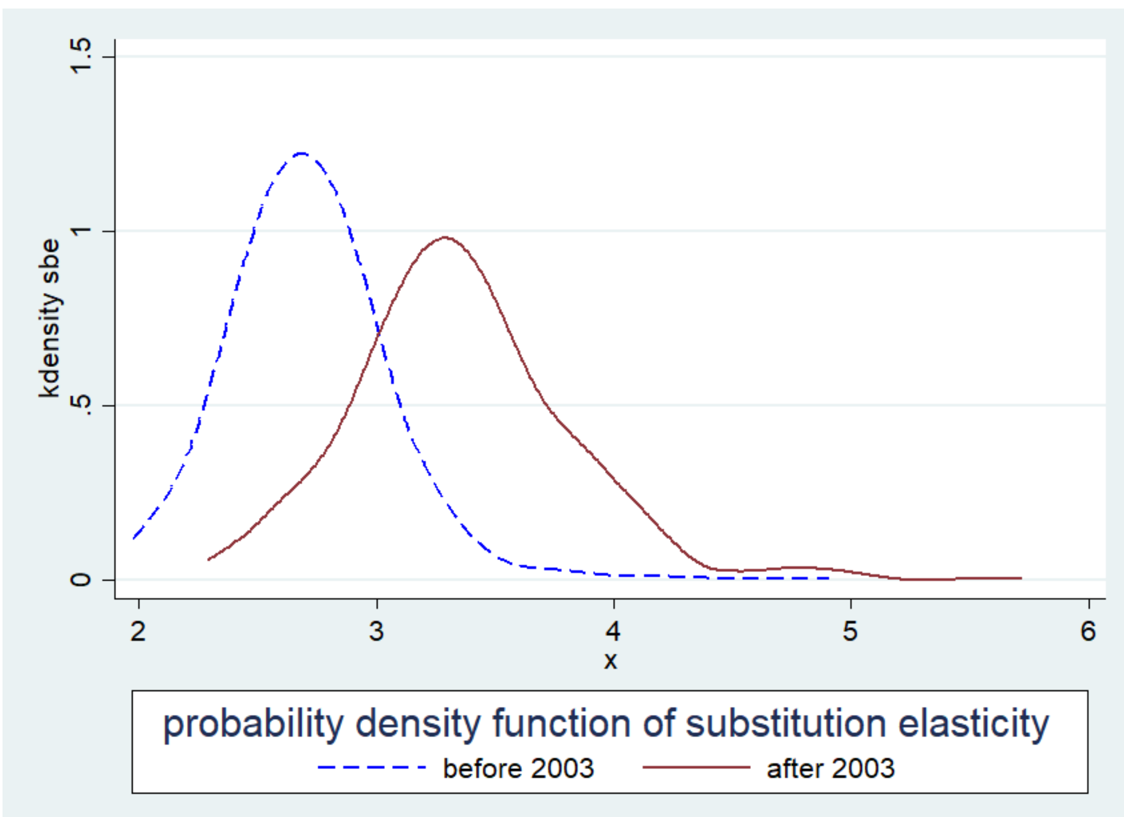


Figure B.2: PDF of demand substitution elasticity in two periods

Table B.3: Summary Statistics

	No. obs	Mean	Std
Capital	373,356	173084.2	1168077.0
Revenue	373,364	143345.5	490274.6
Employee	373,374	486.3	1718.9
Intermediate Inputs	373,364	115742.6	434432.7
Export Counts	286,987	5.7	10.1
Import Counts	195,411	20.7	32.1

NOTE: This table reports the summary statistics of variables. The unit is 1000 RMB.

Table B.4: Demand Elasticity Inverse

	Point est	Sd error est
mean	0.498	0.183
std	0.253	0.141
1%	0.002	0.001
25%	0.317	0.091
50%	0.635	0.186
75%	0.656	0.192
99%	0.807	0.302

NOTE: This table reports point estimation and standard error estimation of demand elasticity inverse of each product-year pair.

Table B.5: The Marginal Benefit of Increasing Input Variety

	S_t
Year	Point est
2000	0.696
2001	0.636
2002	0.642
2003	0.582
2004	0.586
2005	0.536
2006	0.520
2007	0.651

NOTE: This table reports point estimation and standard error estimation of S_t and λ.

B.2 Product-specific parameters report

Table B.6: Report for product-specific parameters

HS4	Obervations	OP step 2	OP step 1			import side	Demand funciton	
		$\alpha*$	$\beta*$	$\gamma*$	scale	λ	σ ($year < 2003$)	σ ($year \geq 2003$)
10000	10989	0.108	0.017	0.241	0.366	0.984	4.500	4.500
6204	19757	0.024	0.071	0.158	0.253	0.713	2.770	3.434
6203	8758	0.038	0.079	0.181	0.299	0.814	2.642	3.285
6110	3952	0.044	0.075	0.173	0.291	0.728	2.430	2.932
6205	3603	0.040	0.073	0.176	0.289	0.819	3.087	3.933
6202	4681	0.018	0.076	0.162	0.255	0.967	2.856	3.565
6109	4163	0.027	0.067	0.164	0.258	0.770	2.602	3.184
6201	3809	0.027	0.071	0.168	0.266	0.902	2.841	3.551
6206	4570	0.034	0.066	0.158	0.258	0.666	2.843	3.550
6104	5646	0.014	0.064	0.163	0.240	0.868	2.371	2.843
6210	1775	0.029	0.075	0.174	0.277	0.790	2.863	3.577
6108	1761	0.032	0.072	0.170	0.274	0.726	2.471	2.991
6211	1942	0.032	0.072	0.172	0.276	0.905	3.523	4.719
6103	2947	0.021	0.066	0.169	0.256	0.767	2.758	3.415
6111	1105	0.042	0.078	0.179	0.299	0.680	2.651	3.252
6208	1865	0.027	0.059	0.185	0.271	0.959	2.406	2.896
6112	593	0.035	0.070	0.176	0.280	0.510	3.932	5.436
6209	896	0.038	0.074	0.173	0.285	0.979	2.634	3.225
6107	677	0.047	0.073	0.163	0.283	0.863	3.022	3.872
6207	603	0.127	0.071	0.175	0.372	0.944	2.688	3.307
6212	379	0.014	0.072	0.192	0.278	0.631	2.173	2.560
6105	194	0.039	0.065	0.171	0.275	0.900	3.223	4.155
6114	199	0.048	0.067	0.171	0.285	0.897	3.108	3.965
5209	41	0.003	0.073	0.169	0.244	0.720	4.500	4.500
5208	24	0.084	0.067	0.161	0.312	0.908	8.860	23.109
6006	49	0.049	0.072	0.169	0.289	0.000	4.875	7.378
6217	132	0.044	0.067	0.165	0.276	0.455	2.904	3.640
6301	47	0.044	0.068	0.168	0.281	0.727	10.183	34.950
4203	100	0.175	0.074	0.166	0.415	0.939	4.908	7.453
mean		0.038	0.070	0.170	0.279	0.783	3.096	4.398
R2		0.561	0.626			0.520	0.827	

Note : the parameters are calculated in hs5-specific product and reported in the simple mean of hs5 product belong in to one hs4 code. In the bottom of the table, I also report the simple mean of the parameters of 65 hs5 products.

Table B.7: Summary Statistics for log form of product-level TFP

year	2002	2007
Observations	16,275	25,723
mean	5.45	5.60
std	1.45	1.42
1%	1.55	1.86
25%	4.52	4.71
50%	5.54	5.69
75%	6.48	6.58
99%	8.43	8.59

Figure B.3: The distribution of log product-level TFP in year 2002 (blue dash) and 2007 (black solid) for product 1 to product 20

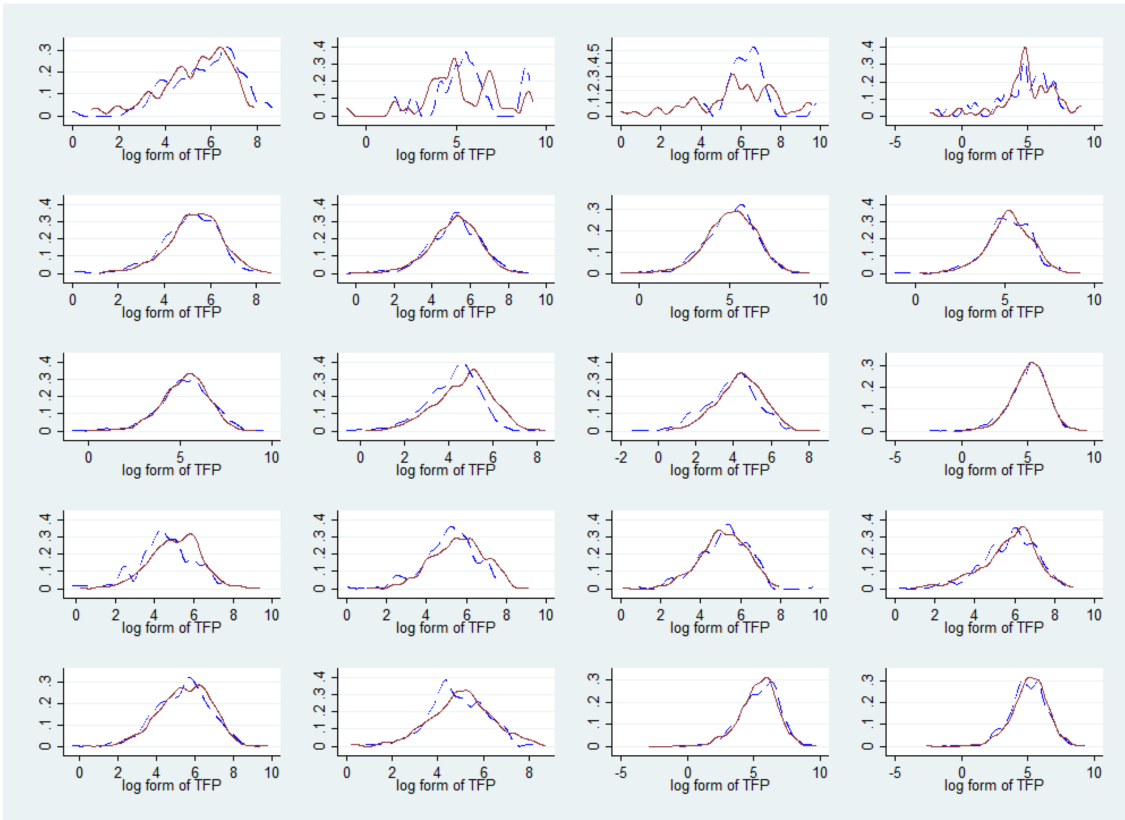


Figure B.4: The distribution of log product-level TFP in year 2002 (blue dash) and 2007 (black solid) for product 21 to product 40

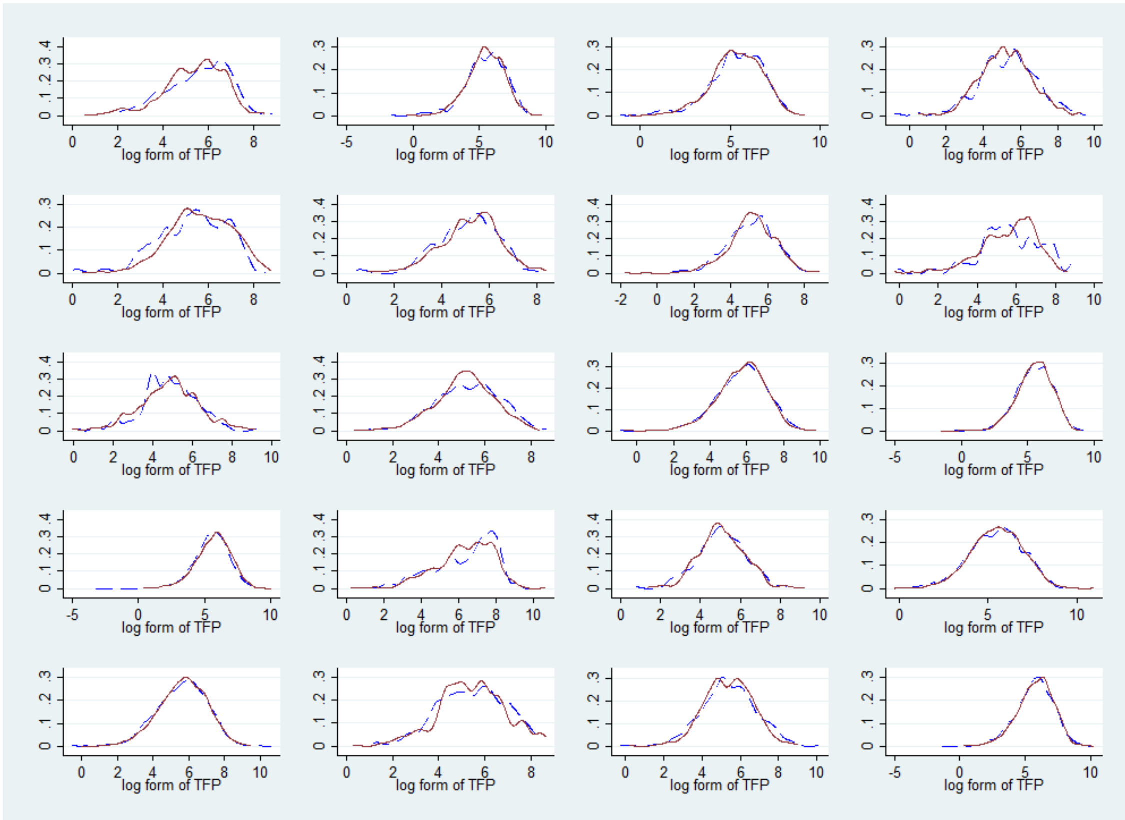
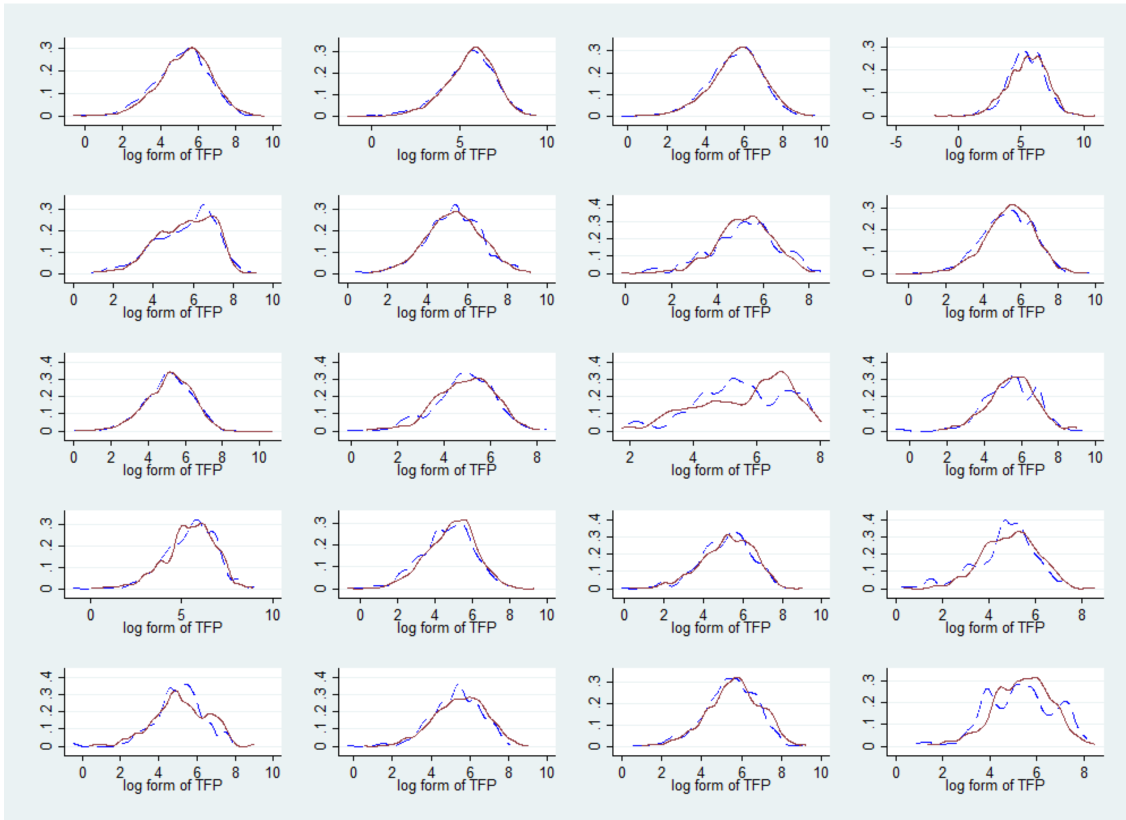


Figure B.5: The distribution of log product-level TFP in year 2002 (blue dash) and 2007 (black solid) for product 41 to product 60



B.3 Details in the decomposition of TFPR

B.3.1 The link between firm-level and product-level TFPR

The relationship between the revenue and TFPR is given as

$$R_{ijt} = A_{ijt}K_{ijt}^{\alpha_j}L_{ijt}^{\beta_j}M_{ijt}^{\gamma_j} \tag{B.5}$$

Rewrite the firm revenue in a function of product-level TFPR

$$\begin{aligned} R_{it} &= \sum_j R_{ijt} \\ &= \sum_j A_{ijt}K_{ijt}^{\alpha_j}L_{ijt}^{\beta_j}M_{ijt}^{\gamma_j} \\ &= (\sum_j A_{ijt}\frac{K_{ijt}^{\alpha_j}L_{ijt}^{\beta_j}M_{ijt}^{\gamma_j}}{K_{it}^{\alpha}L_{it}^{\beta}M_{it}^{\gamma}})K_{it}^{\alpha}L_{it}^{\beta}M_{it}^{\gamma} \end{aligned} \tag{B.6}$$

Then, we find the link

$$A_{it} \equiv \sum_j A_{ijt}s_{ijt} \tag{B.7}$$

where $s_{ijt} = \frac{K_{ijt}^{\alpha_j}L_{ijt}^{\beta_j}M_{ijt}^{\gamma_j}}{K_{it}^{\alpha}L_{it}^{\beta}M_{it}^{\gamma}}$.

B.3.2 Decompose A_{it}

$$TFP_{s,0} = \sum_{i,0} s_{it}A_{it}$$

$$TFP_{s,T} = \sum_{i,T} s_{iT}A_{iT}$$

$y_{i,t} \equiv TFP_{it} - TFP_{s,0}$ Firm level decomposition:

$$\Delta TFP_{s,0} = \underbrace{\sum_{i\in C} s_{i,0}\Delta y_{i,0}}_{Within} + \underbrace{\sum_{i\in C} \Delta y_{i,0}\Delta s_{i,0}}_{Cross} + \underbrace{\sum_{i\in C} y_{i,0}\Delta s_{i,0}}_{Between} + \underbrace{\sum_{i\in E} s_{it}y_{i,t}}_{Entry} - \underbrace{\sum_{i\in X} s_{i0}y_{i0}}_{Exit} \tag{B.8}$$

Decompose Firm-level entry/exit into export market and domestic market.

$$\underbrace{\sum_{i\in E} s_{it}y_{i,t}}_{Entry} - \underbrace{\sum_{i\in X} s_{i0}y_{i0}}_{Exit} = \sum_{i\in E} s_{it}\sum_{j\in EXP} s_{ijt}A_{ijt} + \sum_{i\in E} s_{it}s_{iDt}A_{iDt} - \sum_{i\in E} s_{i0}\sum_{j\in EXP} s_{ij0}A_{ij0} - \sum_{i\in E} s_{i0}s_{iD0}A_{iD0} \tag{B.9}$$

Product level decomposition

$$\begin{aligned}Within + Cross = & \underbrace{\sum_{i \in C} s_{it} \sum_{j \in C_i} s_{ijt} y_{ijt} - \sum_{i \in C} s_{it} \sum_{j \in C_i} s_{ij0} y_{ij0}}_{surviving.Product} \\ & + \underbrace{\sum_{i \in C} s_{it} \sum_{j \in E_i} s_{ijt} y_{ijt}}_{Entry.Product} - \underbrace{\sum_{i \in C} s_{it} \sum_{j \in X_i} s_{ij0} y_{ij0}}_{Exit.Product}\end{aligned} \tag{B.10}$$

Decompose product-level swithch in export market and domestic market

$$\begin{aligned}\sum_{i \in C} s_{it} \sum_{j \in E_i} s_{ijt} y_{ijt} - \sum_{i \in C} s_{it} \sum_{j \in X_i} s_{ij0} y_{ij0} = & \sum_{i \in C} s_{it} \sum_{j \in E_i \wedge EXP} s_{ijt} y_{ijt} + \sum_{i \in C} s_{it} \sum_{j \in E_i \wedge D} s_{ijt} y_{ijt} \\ & - \sum_{i \in C} s_{it} \sum_{j \in X_i \wedge EXP} s_{ij0} y_{ij0} - \sum_{i \in C} s_{it} \sum_{j \in X_i \wedge D} s_{ij0} y_{ij0}\end{aligned} \tag{B.11}$$

Within product decomposition:

$$\begin{aligned}Prod.Surviving = & \sum_{i \in C} s_{iT} [\sum_{j \in C_i} (TFP_{ijT} - TFP_{s0}) s_{ijT} - \sum_{j \in C_i} (TFP_{ij0} - TFP_{s0}) s_{ij0}] \\ = & \underbrace{\sum_{i \in C} s_{iT} [\sum_{j \in C_i} (\Omega_{ijT} e^{\gamma_j a_T G_j(N_i)} - TFP_{s0}) s_{ijT} - \sum_{j \in C_i} (\Omega_{ij0} e^{\gamma_j a_T G_j(N_i)} - TFP_{s0}) s_{ijT}]}_{upgrade.effect} \\ & + \underbrace{\sum_{i \in C} s_{iT} [\sum_{j \in C_i} (\Omega_{ij0} e^{\gamma_j a_T G_j(N_i)} - TFP_{s0}) s_{ijT} - \sum_{j \in C_i} (\Omega_{ij0} e^{\gamma_j a_0 G_j(N_i)} - TFP_{s0}) s_{ijT}]}_{import.effect} \\ & + \underbrace{\sum_{i \in C} s_{iT} [\sum_{j \in C_i} (\Omega_{ij0} e^{\gamma_j a_0 G_j(N_i)} - TFP_{s0}) s_{ijT} - \sum_{j \in C_i} (\Omega_{ij0} e^{\gamma_j a_0 G_j(N_i)} - TFP_{s0}) s_{ij0}]}_{reallocation.effect}\end{aligned} \tag{B.12}$$