FIRM-LEVEL HUMAN CAPITAL AND INNOVATION: EVIDENCE FROM CHINA

A Thesis
Presented to
The Academic Faculty

by

Xiuli Sun

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Economics

Georgia Institute of Technology
August 2015

FIRM-LEVEL HUMAN CAPITAL AND INNOVATION: EVIDENCE FROM CHINA

Approved by:

Professor Haizheng Li, Advisor
School of Economics
Georgia Institute of Technology

Professor Vivek Ghosal
School of Economics
Georgia Institute of Technology

Professor Xiaoming Huo
School of Industrial & Systems Engineering
Georgia Institute of Technology

Professor Patrick S. McCarthy
School of Economics
Georgia Institute of Technology

Professor Shatakshee Dhongde
School of Economics
Georgia Institute of Technology

Date Approved: 18 July 2015

To my parents and siblings

ACKNOWLEDGEMENTS

I owe my gratitude to all those people who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever.

My deepest gratitude is to my advisor, Prof. Haizheng Li. I have been amazingly fortunate to have an advisor who gave me endless trust, encouragement, the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. He provided me with every bit of guidance, assistance, and expertise that I needed. I am grateful to him for holding me to a high research standard and enforcing strict validations for each research result, and thus teaching me how to do research. His patience and support helped me overcome many crisis situations. In the days of job seeking and dissertation writing, his selfless time and care were sometimes all that kept me going. I hope that one day I would become as good an advisor to my students as Prof. Li has been to me.

Prof. Patrick S. McCarthy is one of the best professors that I have had in my life. His lectures largely deepened my understanding in econometrics. This helps me a lot in the empirical analysis of my dissertation. I am also thankful to him for carefully reading through my dissertation, commenting on my work, and further clarifying and discussing those comments with me. His help and support means a lot to me. As an expert on innovation, Prof. Vivek Ghosal's insightful comments and constructive criticisms of my research are always thought-provoking. His familiarity with the innovation literature always helps me keep an eye on related studies and inspire me to think more about my work. In addition, as the director of graduate programs, Prof. Ghosal has been always there to be supportive and give advice. I am grateful to Prof. Shatakshee Dhongde for her encouragement and practical advice. I am thankful to her for commenting on my research papers and helping me understand and enrich my ideas. I am also thankful to Prof. Dhongde for her help and suggestions on job market. Prof. Xiaoming Huo, from school of Industrial and Systems Engineering, provides me insightful comments and they helped me enforce strict validations for each research result. I am grateful to him for providing me an opportunity to examine my work from statistics perspective.

I am also deeply indebted to Prof. Emilson C.D. Silva, Prof. Juan Moreno-Cruz, and Prof. Byung-Cheol Kim. They helped me a lot build a systematic understanding of academic research at the first years of my graduate studies. They were and remain my role models for a researcher, mentor, and teacher. Their enthusiasm and love for research and teaching are contagious.

I am also thankful to other professors in School of Economics (SOE), Prof. David N. Laband, Prof. Usha C. Nair-Reichert, Prof. Tibor Besedes, Prof. Mikhail M. Klimenko, Prof. Levent Kutlu, and Prof. Erik Johnson with whom I have interacted during the course of my graduate studies. Their lectures and talks helped me understand my research area better and laid a solid foundation for my research in economics . I also would like to thank Jyldyz Hughes and Tony Gallego for providing me great help and support during my graduate study in SOE. In addition, I must acknowledge the financial support and other supports from SOE and Ivan Allen College of Arts.

I am surrounded by angels but I call them true friends. Together we share knowledge, dreams, smiles and tears. I greatly value their friendship and I deeply appreciate their belief in me. True friends are hard to find, harder to leave, and impossible to forget.

I would also like to thank my family for the support they provided me through my entire life and in particular, I must acknowledge my grandfather, without whose love and encouragement, I would not have finished this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis examines firm innovation in China from firm-level human capital perspective since resource-based theory and upper echelon theory reveal that the reason why firms vary in performance is that they differ in human capital. Two types of human capital are examined: general human capital measured by number of highly educated workers, and managerial human capital measured by characteristics (education and age) of general manager and management team. Besides human capital indicators, we also take R&D, firm size, market structure, firm age, ownership, city fixed effects, and industry fixed effects into account. Given the fact that innovations are made up of multifarious elements and hard to measure and define, this thesis examines firm innovation from three different aspects, patent applications, product innovation and total factor productivity.

The datasets from two comparable firm-level surveys conducted by the World Bank including data from 1998-2000 and from 2000-2002 respectively. Firms in the first dataset (year 1998-2000) are from large cities such as Beijing, while firms in the second dataset (year 2000-2002) are from provincial middle cities. Notably, these two datasets are comparable but at the same time are from different market environment. This enables us to test model consistency and at the same time examine how different market environment influences firm innovation. The most important results of patent application regression are that general human capital, or skilled human capital, plays an important role in innovation in more developed areas, i.e., large cities, while it has a much smaller or no effect on innovation in less developed areas, i.e., provincial middle cities; R&D has a very significant and positive effect on innovation in less

developed areas while it has a smaller or no effects on innovation in more developed areas. Interestingly, we get the same results when we use total factor productivity to study technological change. A possible explanation is that there's enough R&D in more developed areas while it is not enough in less developed areas. Thus, the bottleneck in more developed area is human capital while in less developed area it is R&D. A policy implication is that to make innovation process more efficient, policies in more developed areas should focus more on promoting human capital aspect rather than R&D and meanwhile policies in less developed areas should focus more on promoting R&D spending. Moreover, we find that the education of managerial personnel (including both general manager and management team) has a positive effect while their age has a negative effect on firm innovation.

In particular, we find that R&D has a very significant and positive effect on product innovation, or new product introduction. This is consistent with three product innovation measures and across two datasets. Moreover, we find that the marginal effect of R&D is much larger in more developed area than in less developed area when we use value based innovation measure. However, when we use count measure of product innovation, we find that the marginal effect of R&D on product innovation is larger in less developed area than in more developed area. A possible explanation is that the unit value of new products introduced is larger in more developed area than in less developed area. At the same time, we find that the effect of management team's education differs in two datasets. The relationship between number of new products introduced and average years of schooling of management team is concave in more developed area while it is monotonically increasing in less developed area. The implication is that in more developed areas, firms with higher human capital might focus more on the *quality* of product innovation rather than the *quantity*.

CHAPTER I

INTRODUCTION

Manufacturing plays an important role in China's rapid economic growth. As a pillar industry and driver of Chinese economic growth, manufacturing has contributed more than 40 percent of China's gross domestic product (GDP) in terms of industrial scale and structure. According to data from National Bureau of Statistics of China (NBSC), during the last 30-odd years since reform and opening began, manufacturing in China has grown much faster than the overall economy. The average growth of manufacturing output was about 20 percent per annum between 2005 and 2013 compared to around 10 percent of the overall growth rate. In 2012, China's manufacturing added value reached US $ 2,080 billion, or about 20 percent of the global total, making China a manufacturing giant. However, now China faces competition in manufacturing not only from developing countries such as Vietnam, but also developed countries. After the 2008 global financial crisis, western countries around the world focused once again on manufacturing and launched manufacturing development strategies. For example, US launched a national strategic plan for advanced manufacturing in 2012. Given the importance of manufacturing in its overall economy and intense competition faced by China, innovation becomes vital for manufacturing in China.

Clearly, the importance of innovation has been widely recognized in China as we can see that R&D spending expands largely. According to OECD, China was the second largest R&D spender in 2012, allocating around 294 billion dollars compared to top-spending the US at around 454 billion dollars that year. In fact, China is forecast to overtake the European Union and the United States in R&D by 2019

(See Figure 1). Does this mean that China has finished the transition from "Made in China" to "Innovated in China"? Probably not. Despite success in some areas, notably high-speed rail, solar energy, supercomputing and space explorations, China is still far from an innovative country. For example, there was no Chinese company listed on 100 most innovative companies by Thomson Reuters in 2013. There was only one Chinese company, Huawei, entering the list in 2014 while around 40 percent listed firms were from the US.[1] It seems that there's an asymmetry between innovation input and output in China. All these facts motivate us to examine firm innovation in China.

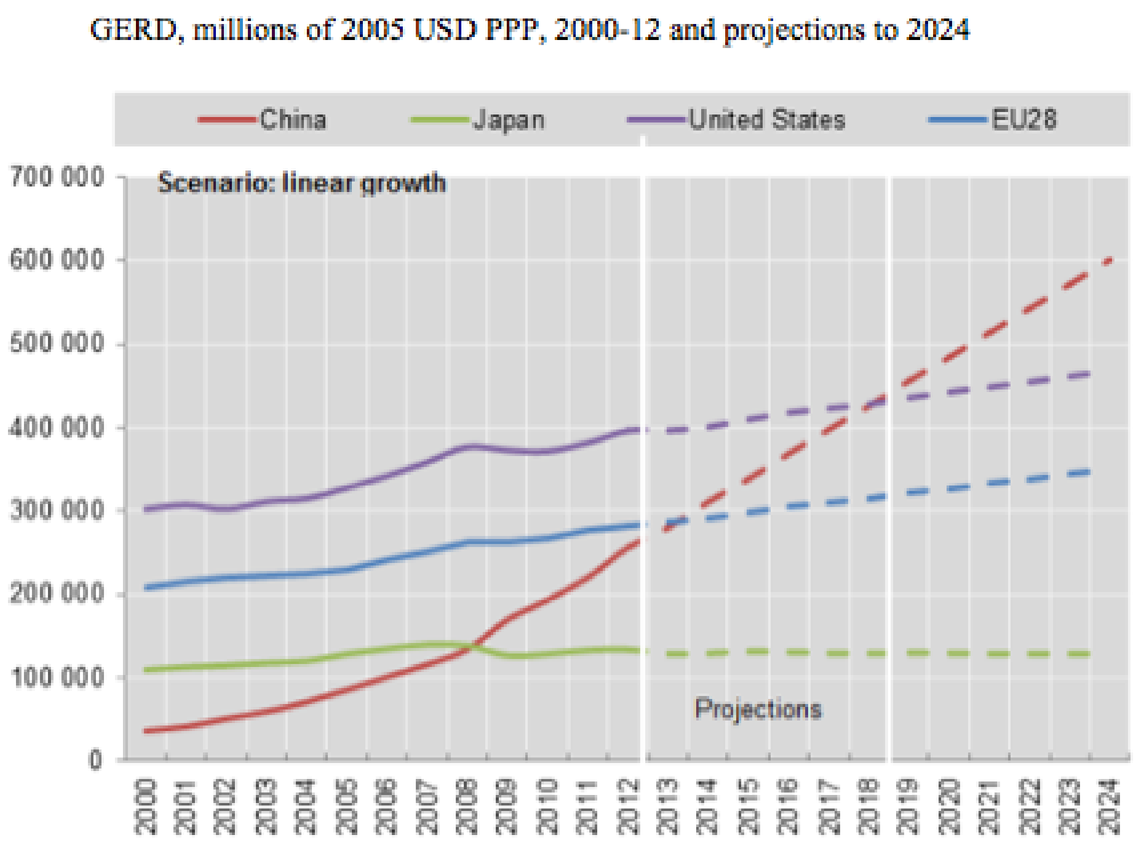

Figure 1: China forecast to outpace the US in R&D spending around 2019. (Source: OECD Science, Technology and Industry Outlook 2014)

[1]For example, I.B.M., Apple, Microsoft, Boeing, Exxon mobile and Intel.

Innovation is significantly influenced by several external factors resulting in specific innovation systems. An innovation system includes not only networks of innovative companies with research organizations, suppliers and customers, but also several institutional factors, such as the way publicly financed research is organized, or the region's system of schooling, training and financial institutions. Production of economically useful new technological knowledge results from collective actions of different actors of the system connected by various linkages ranging from informal to formalized network relationships (Acs et al., 2002). Although the sector in which a firm operates is an important element of the context within which decisions about the commitment of resources to R&D and innovation is made, it is the firm which makes the investment decision and the firm which enjoys the rewards of innovation. According to resource-based theory and upper-echelon theory, firm-level human capital, especially managerial human capital, is very important to firm behavior and thus firm innovation. This is the motivation underlying this thesis.

To understand the exact role of innovation in the economy and innovation itself, the measurement of innovation is critical. However, the unanimity with which its importance is recognized is not matched by adequate development in methods for its measurement. This is due to the fact that the innovative process is difficult to define and measure. Moreover, innovations are made up of multifarious elements, for example, process innovation and product innovation.[2] Generally, there are three ways of measurement: (1) a measure of the inputs into the innovation process, such as R&D expenditure; (2) an intermediate output, such as the number of inventions which have been patented; (3) a direct measure of innovative output. Early innovation studies in 1950s and 1960s, such as Scherer (1982,1984a,1986), mainly relied on R&D

[2]Schumpter (1934) distinguishes five types of innovation: the introduction of a new or a new quality of a good; the introduction of a new method of production; The opening of a new market; the conquest of a new source of supply of raw materials or half-manufactured goods; the carrying out of a new organization of any industry. This thesis will only deal with the technological innovations, i.e., product innovation and process innovation (those classified by Schumpeter in points 1 and 2).

as proxy for innovation. R&D suffer from measuring only the budgeted resources allocated towards trying to produce innovative activity (Acs et al., 2002) and thus cannot represent the whole of innovative activity. It often fails to take into account the innovative activity of small and medium-sized firms. Because of data availability, advances in innovation studies in 1970s made in the use of patent data, an intermediate measure of economic activity, as a proxy for economic output. Although patents are good indicators of new technology creation, they do not measure the economic value of these technology (Hall et al., 2001; Acs et al., 2002). Also, as Pakes and Griliches (1980) pointed out that patents are a flawed measure (of innovative output) particularly since not all new innovations are patented and since patents differ greatly in their economic impact. Also, there is a different propensity to patent between firms and sectors. Compared to proxies of innovation activities such as R&D expenditures or patents, innovation output measures provide a direct indicator of innovation. The advantage of the direct indicators over patents and R&D expenditures is that they document the ultimate end of every innovation process.

When we choose to use innovation output to measure innovation, we still encounter some alternatives: innovation counts, innovation sales, innovation sales share. Innovation counts measure is better than patents in the sense that it includes not only patented innovation but also non-patented innovation. However, like patents, innovation counts measure assumes that all innovations have the same technological and economic relevance, a simplification that obviously does not correspond to reality. In fact, every single innovation has, in fact, a different technological, economic, and commercial value. Furthermore, it frequently happens that the technological and economic values of an innovation fail to coincide. Compared to it, innovation sales are in monetary units and this eliminates the problem of the different commercial value of individual innovations, and the technology flows are, in fact, measured directly in money terms. However, innovation sales also has its limitation. For example, it is influenced by price and market condition. Given the advantage and disadvantage of different measures, we will use all the measures in our study to get a deeper understanding. In addition, the concept of total factor productivity (TFP) change is used synonymously with technological change in the productivity literature (Nishimizu and Page, 1982). By definition, TFP is the proportion of output not explained by the amount of inputs used in production.

The second chapter reviews innovation history during the last decades and related literature. The factors important to innovation in China are policy support, rapid economic growth, FDI, R&D inputs and human capital growth. The third chapter explores firm-level innovation from a human capital point of view using patents as proxy for innovation. In the theoretical model, two firms compete with each other in a three-stage Cournot, innovation stage and production stage. Skilled human capital level can affect innovation success probability directly and via R&D level indirectly. Managerial human capital can affect firm innovation through their choice of projects and R&D level. We find that a firm's innovation is not only determined by its human capital level, firm characteristics, and its market share, but also might be affected by market environment. In the empirical study, we use two firm-level datasets from China, one from metropolitan cities and one from provincial middle cites. Human capital indicators are skilled human capital (number of highly educated workers), general manager's education and experience, and management team's education and age. We find that a firm's skilled human capital and managerial personnel's education have significantly positive effects on innovation while the management team's age has a significantly negative effect on innovation. The effect of R&D on patents is insignificant in large metropolitan cities while it is positive and significant in provincial middle cities.

In the fourth chapter, we study the relationship between firm-level human capital and product innovation. Three measures are used: new product sales proportion, new product sales, and number of new products. We find that for skilled human capital, we find that it also tends to have a positive effect across the three measures, though it is not significant in regression using new product proportion as product innovation measure. When new product sales proportion and new product sales are used as product innovation measures, we find that management team's average years of schooling has a positive effect in determining product innovation, however, its effect in Data 2000 when new product count used as product innovation measure is negative and significant. This indicates that though the results from the three regressions are generally consistent, but they still differ. When compare the results using different innovation measures, we should be very cautious. Another thing is that different from results in last chapter where R&D has a positive effect only in Data 2003, less developed areas, we find that R&D has a very significant and positive effect on product innovation no matter which product innovation measure is used. This indicates that R&D is still important in promoting product innovation and policies supporting investment in R&D is still important. Our results still hold when endogeneity is considered. Notably, in all three models, we all find that general manager's postgraduate degree has a large and significant effect in less developed areas but insignificant at all in more developed areas. The reason might be that in less developed areas, market development is much more incomplete and thus general manger's education matters more. A higher education can enable a general manager to make more insightful decision.

In the fifth chapter, the relationship between firm-level human capital, R&D and total factor productivity (TFP) is examined. Firstly, production function is estimated using methods proposed by Levinsohn and Petrin (2003, LP, thereafter) and Ackerberg, Caves and Frazer (2006, ACF) since inputs are very likely to be correlated with productivity. We then use TFP from ACF method to examine the determinants of TFP. We find that skilled human capital is important in determining TFP even when it is already included in production function. Management team's age has a negative effect on TFP. Management team's average schooling has a positive and significant effect in less developed areas, but we fail to find a significant effect in more developed areas. Notably, R&D has a positive and significant effect across both datasets. When the endogeneity of number of highly educated workers, and general manager's postgraduate degree is considered, our main results still hold.

This dissertation makes four main contributions: (1) it provides both theory and theoretical framework for studying firm innovation in a human capital view, not only skilled human capital but also managerial human capital and R&D human capital. Growth theory only provides a theoretical framework to incorporate skilled human capital into innovation. Though there are some empirical studies on firm innovation human capital, those studies are either on the relationship between firm innovation and managerial human capital or on the relationship between firm innovation and skilled human capital and none of them take care of skilled human capital and managerial human capital at the same time. (2) Using detailed firm-level data, we are able to study the effects of skilled human capital, general manager's education and experience, and management team's education and age. (3) Two datasets from two different levels of cities, metropolitan cities and provincial middle cities enable us to examine the effect of market environment on firm innovation. (4) Several technological change measures are used at the same time and thus we can get a much better understanding of firm innovation in China.

The thesis is organized as follows. Chapter 2 presents the history of innovation in China and innovation literature. Chapter 3 examines the relationship between firm-level human capital and innovation using patents as proxy for innovation. Chapter 4 studies the relationship between firm-level human capital and product innovation. Chapter 5 examines the relationship between firm-level human capital, R&D and TFP. Chater 6 concludes.

CHAPTER II

BACKGROUND AND RELATED LITERATURE

2.1 Innovation in China

Previous studies have shown that government policy, increasing financial and human capital investment in innovation and infrastructures are crucial determinants of national innovative capacity (Furman et al., 2002; Furman and Hayes, 2004). Liu and Buck (2007) further emphasized that learning and imitation strategies are particularly important for developed countries to catch up rapidly and enhance its own innovative capacity.

Firm innovation roots in the development of science and technology in a society. Meanwhile, government policy and social environment play an important role in promoting the development of science and technology in a society. Thus, to trace back the history of firm innovation in China, besides R&D activities conducted by government and firms, we will also focus on government policy and social environment.

Traditionally, it is literature and the arts rather than science and technology that is regarded as important or as carrying social status. But this began to change when a rapidly modernizing West came knocking on the Qing Dynasty's door in the late 1700s. At the same time, long-lasting wars in China from that time until around 1949-1950 began. During wars, the development of science and technology and industrialization is very slow. From 1949 to 1957, China practiced central planning under the direction of the State Planning Commission (SPC). The main function of planning was to set production goals, controlled prices, and allocated resources throughout most of the economy by state-owned enterprises. As a result, by 1978 nearly three-fourths of industrial production was produced by centrally controlled, state-owned enterprise, according to centrally planned output targets (Morrison, 2014). Private enterprises and foreign-invested firms were generally barred but China had a close economic cooperation with Soviet Union. While the Soviets provided a massive technology transfer, China took charge of all science and research facilities and focused them on mammoth projects to create an industrialized economy as fast as possible. Its science establishment was decimated by political campaign and the Cultural Revolution, which lasted from 1966 to 1976. Thus, until 1976, there's very little progress in firm innovation.

Beginning in 1978 the Chinese government changed the economic system gradually towards a market economy, allowing non-state enterprises to produce and compete with state enterprises. When Xiaoping Deng launched reform and opening in 1978, he focused immediately on science and technology as key to China's modernization. Still, at that time, China's economy featured a "up-to-down" type. That is, at national level, central government set objectives and allocate resources to basic science and key areas. For example, in 1978 the State Science Commission summoned some 20,000 experts to draft a new blueprint for science to serve as a drive for restarting China's economy. The plan focused on 27 sectors of research and 108 key research projects. Eight large projects were planned in the fields of agriculture, energy, materials, electronic computers, lasers, space science, high-energy physics and genetic engineering. Especially, to promote the development of science and technology, in 1983 a leading group for science and technology was created, which is under the direct guidance of the premier. Throughout the 1980s and early 1990s, science and technology system reforms and new programs went into fast-forward. In the late 1990s, Chinese scientists and the government together began to push for a "national innovation system", they began to strengthen their research institutes, investing in people and infrastructure with the goal of creating 30 globally recognized research centers. In 2005, CPC elevated indigenous innovation to a strategic level. The landmark document "The National Medium-and Long-Term Plan for the Development of Science and Technology (2006-2020)"(MLP) is launched in 2006, served as the grand blueprint of science and technology development. Throughout the policy changes in the recent decades, we can see that Chinese government almost always put science and technology as its one of its primary objectives. Policy support provides the first condition for the development of innovation in China.

Economic growth provide the source and market for innovation. On the one hand, as economy prosper, firms are more easily to gain profits and thus they can put more in R&D. On the other hand, with fast economic growth, market demand will also increase. Both effects can promote firm innovation. Figure 2 presents the aggregate growth from 1979-2013. For the last decades, the average economic growth rate of China is around 10%. However, there are also two low economic growth period. In late 1980s, the growth rate began to decrease and this might reflect the impacts of China's austerity economic policies in the late 1980s. The low growth rate around 1999 might reflect the 1997 Asian financial crisis (Zheng and Tong, 2014).

Like any other developed country, learning and imitation strategies are particularly important(Kim and Nelson, 2000; Hu et al., 2005). Thus, international technology spillover channels, FDI and international trade, are very important drivers of innovation in developed countries. There are several channels through which international technology can promote local innovation. Firstly, local firms can learn about the designs of new products and technology through interaction with foreign firms. Secondly, technology spillovers from foreign firms to local firms can be realized through human capital mobility or labor market turnover. Thirdly, foreign firm's R&D activities in China can stimulate local firm's innovation (Liu and Buck, 2007). Fourthly, the import of technology also acts as a channel through which domestic

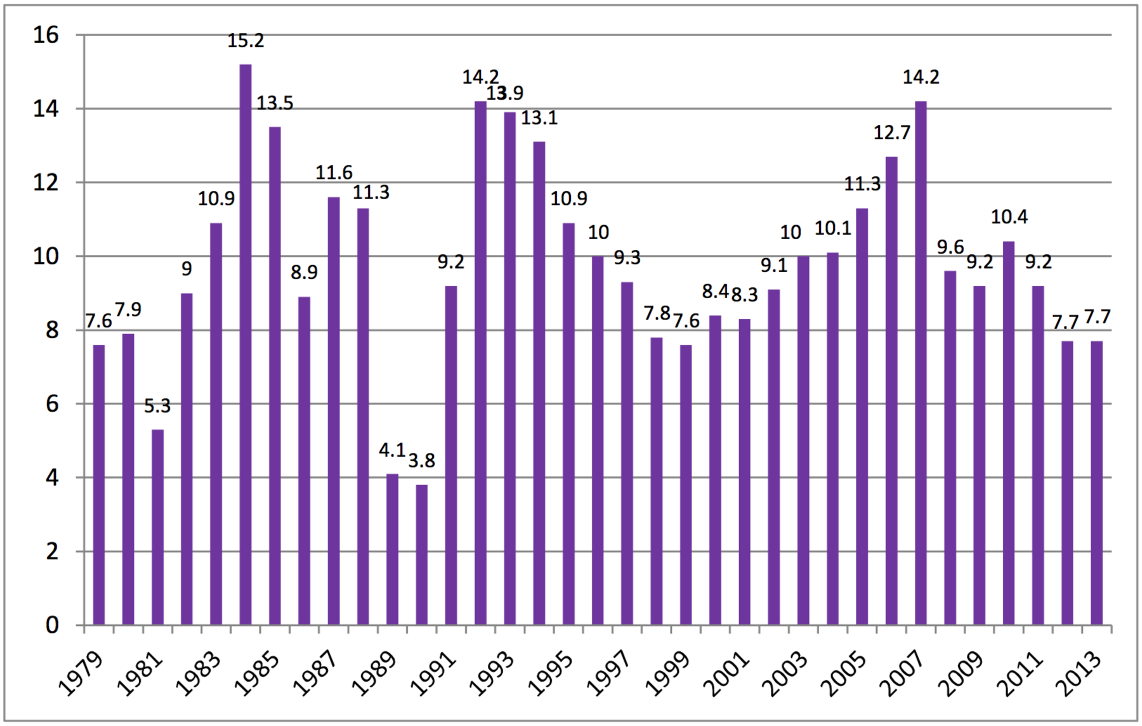

Figure 2: Chinese Real GDP Growth: 1979-2013 (percent) (Source: Morrison (2014))

local firms may "reverse engineer" the products of their foreign rivals (Coe and Helpman, 1995). Finally, exports can also promote domestic firm's innovation through learning by exporting (Salomon and Shaver, 2005). Figure 3 shows that from the end of 1980s, FDI began to increase very fast, from around 4 billion dollars in 1991 to 124 billion dollars in 2013.

We can see that, "indigenous innovation" policies do not advocate closed-door innovation or technological autarchy. Global technology sourcing and the integration of acquired technologies into new technological solutions are explicitly mentioned in the MLP as types of indigenous innovation. At the same time, the plan also sets as target the increase in domestic R&D expenditures relative to expenditure on technology import. Figure 4 shows that China's R&D expenditure and import. We can see that for more than ten years, china's technology import always stay in a high level and domestic R&D even grow much faster. In 1997, R&D is less than half of technology import expenditure., but in 2010, R&D is almost triple of technology

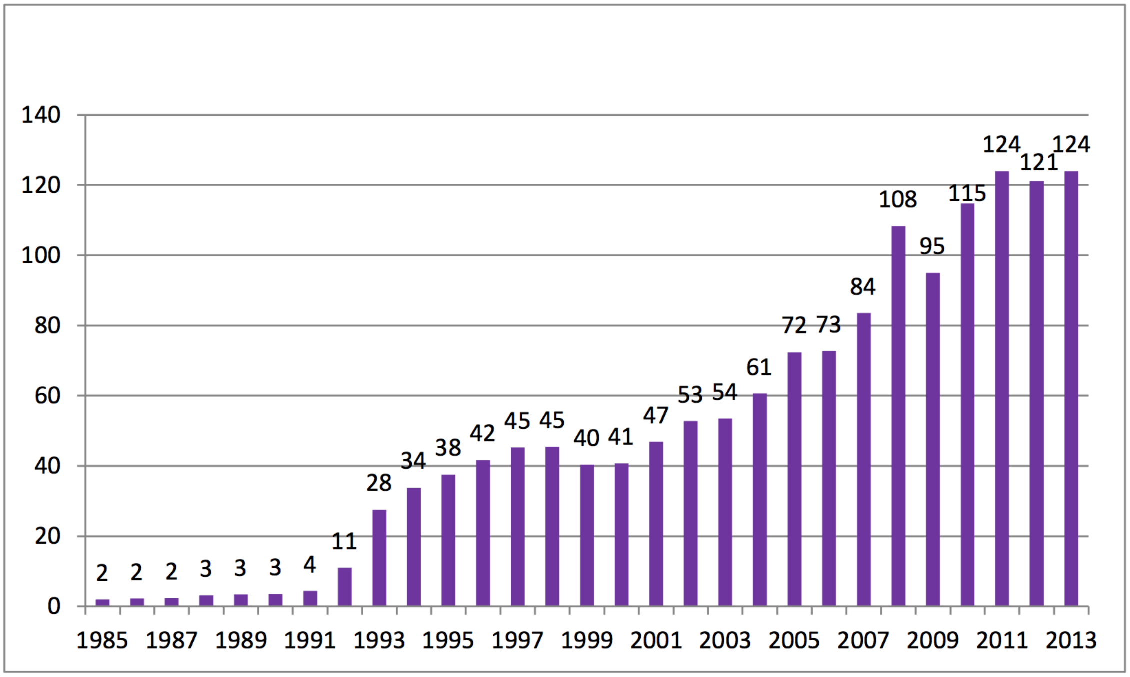

Figure 3: Annual FDI Flows to China: 1985-2013 ($ billions) (Source: Morrison (2014)). Data are from the United Nations.

import expenditure. Besides large R&D expenditure spending, China's human capital also increases a lot and this provides a base for innovation. Figure 5 shows the growth of human capital from 1985 to 2007. Human capital index in 1985 is around 100, then it grew gradually until the beginning of 1990s to around 150. From the middle of 1990s, human capital growth becomes faster and in 2007, it is around 450.

To sum up, we can see that there are several reasons account for the development of innovation in China: government policy support, fast economic growth, FDI, Technology Import, R&D spending growth and human capital growth.

Last but not the least, intellectual property protection is also important to a country's innovation. Kanwar and Evenson (2003) found evidence that intellectual property rights can serve as incentives for spur innovation using cross-country panel data. The first patent law in China was passed in 1984. China's patent office grants three types of patents: invention, utility model and design patents. An invention patent is a new or improved technical solution for a product or process. A utility

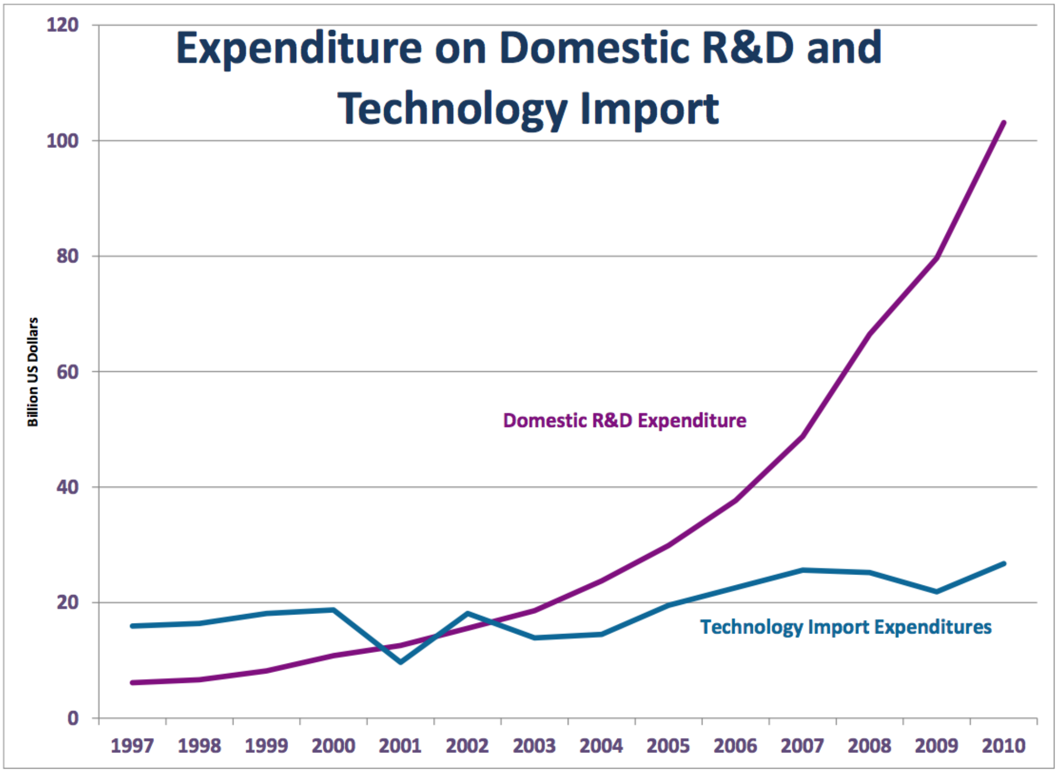

Figure 4: Expenditure on Domestic R&D and Technology Import. (Source: Ernst and Naughton (2012)

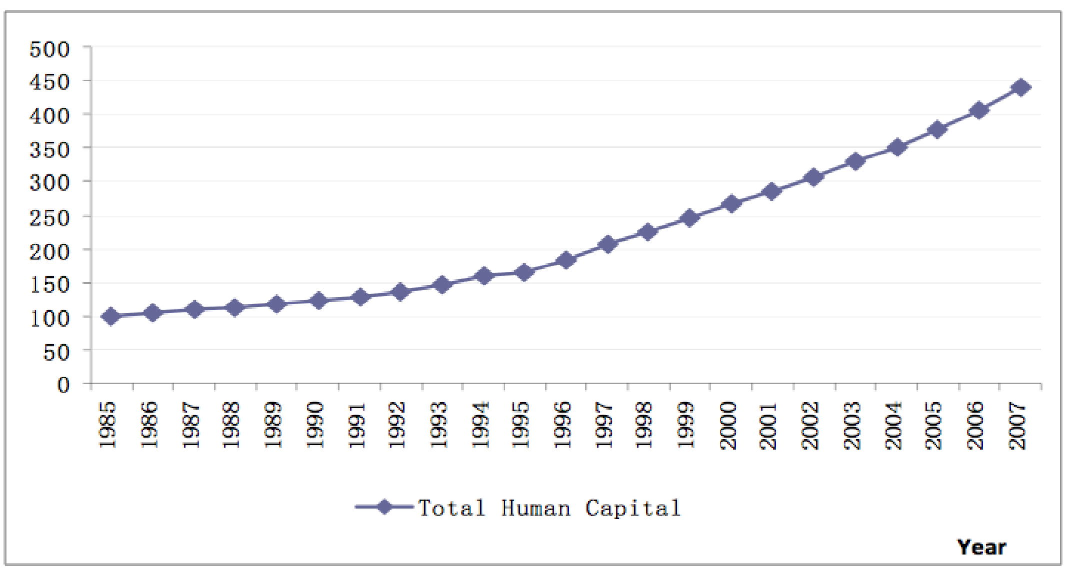

Figure 5: The index of total human capital, 1985-2007. (Source: Li et al. (2009)

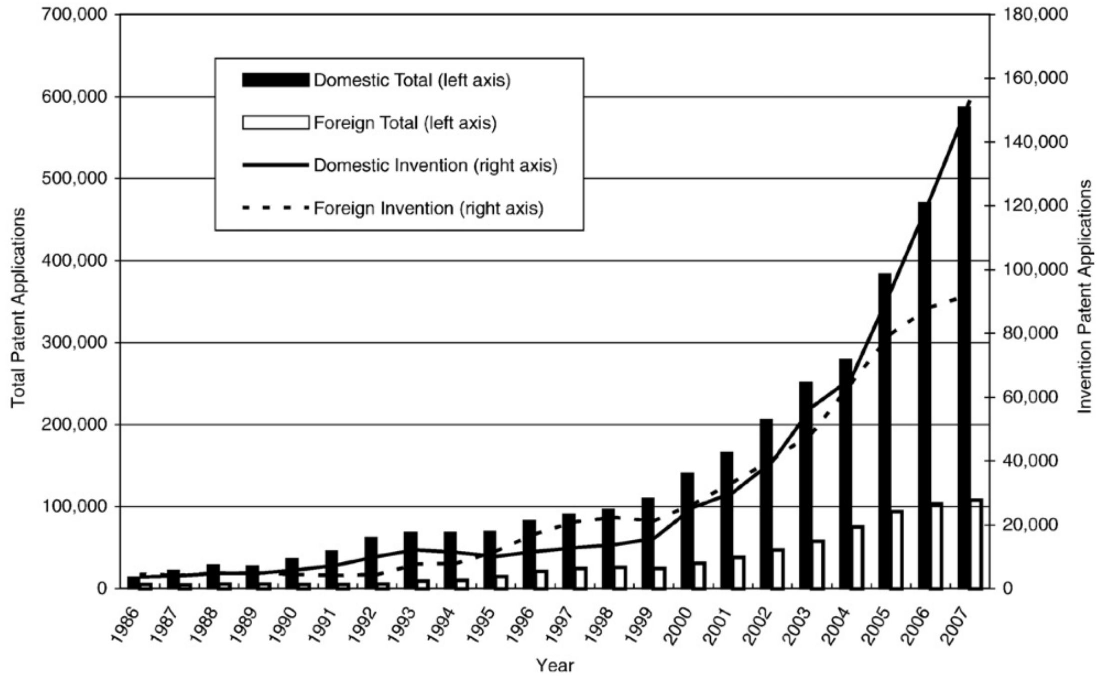

Figure 6: Chinese patent applications, 1986-2007. (Source: Hu and Jefferson (2009)

model is a new technical solution for the shape and/or structure of a product fit for practical use. A design patent is considered more narrow than the other classes and covered a new design of the shape, pattern, or the combination. Applications for invention need to pass a substantive examination for utility, novelty and non-obviousness before granted. The utility model and design patents generally cover more incremental innovations and are only subject to examination for utility. Figure 6 shows patent applications in China during 1986-2007. We can see that from 1986 to 1999, patents grow gradually, but from 2000 patent application began to surge. Moreover though at the end of 1990s there were more foreign inventions, but from 2004 domestic invention began to surpass foreign invention.

2.2 Previous Studies in Innovation

The studies on innovation began to be a long-lasting hot topic since Schumpeter (1942). He asserted that large firms operating in concentrated industries constitute the engine of technological progress. Moreover, he argued that monopoly and oligopoly firms are more able of conducting meaningful R&D because they can use funds earned from profits to finance R&D. In particular, Schumpeter argued that oligopolistic market structures, with their perceived intensity of product and factor cost competition, will achieve more innovation and thus make a greater contribution to social welfare than the severe price competition exhibited by perfectly competitive market structures.Two hypotheses are formulated as two hypotheses (Symeonidis 1996). The first hypothesis postulates a positive relationship between the incentive to innovate and market share or power. Large market share implies greater certainty that a new product will also achieve higher market share and generate profits. Higher profit margins, due to market power, also provide finance for R&D, which is important since capital markets may be reluctant to fund innovative projects. The second hypothesis states that large firm size and innovation are correlated. He believed that a large firm in a concentrated market can innovate more. Since then, a lot of studies began to test Schumpeterian hypotheses. Scherer (1965a, 1965b) used a sample of 448 firms from the 500 largest US industrial firms in 1955 and he found evidence of an inverted-U relationship between R&D employment intensity and sales for the pooled sample and for most sub-samples (the chemical sector was an exception, the relationship being clearly positive). He also found that the number of patents increased less than proportionately with sales, except for a few very large firms. These results were interpreted by Scherer as a clear refutation of the Schumpeterian hypothesis of a positive effect of firm size on innovation. Acs and Audretsch (1988) use direct measure of innovation, find that the total number of innovations is negatively related to concentration and unionization, and positively related to R&D, skilled labor, and the degree to which large firms comprise the industry; these determinants have disparate effects on large and small firms.

Schumpeter, particularly in his early work, also emphasized the important role that committed entrepreneurs capable of overcoming an inert or resisting environment may play in innovation and, largely for this reason, Schumpeter is also acknowledged as an important source of inspiration in the entrepreneurship literature (Landstrom et al., 2012). Nelson and Winter (1982) proposed a theory of firm-level knowledge and believed that the strategies that firms pursue with respect to innovation is important and the outcomes of their actions are shaped by these strategies. Cohen and Levinthal (1990) also focus on the importance of firm-level knowledge, in particular so-called absorptive capacity, which they see as critical for the ability to identify and exploit external sources of knowledge in innovation. These studies provide theoretical base for our study.

Another important factors related to innovation is human capital. Human capital of firm is defined as the knowledge and skills of its professionals that can be used to produce professional services. In Romer model, skilled workers or human capital is explicitly included in the innovation model. A constant problem for human capital study is how to measure it. Pennings et al.(1998) use firm tenure, industry experience, and graduate education to capture firm-level human capital to study technology adoption. Ballot et al.(2001) constructs measures of a firm's human capital stock based on firms' past and present training expenditures to study its effect on productivity. Instead of making any comparison among different measures, we believe that a good measure should meet its own research objective. In our paper, our objective is to study how different types of human capital affect firm's technology decision, to innovate or to imitate. Aghion et al.(2009) suggest that firms with more workers with "high brow" education tend to innovate and firms with more workers with "low brow" education tend to imitate.

Numerous studies in management have sought to identify factors that can stimulate firm innovation. Calantone et al. (2002) examined the relationship among learning orientation and firm innovativeness. They believed that learning orientation was composed of four factors: commitment to learning, shared vision, open-mindedness, and intraorganizational knowledge sharing. Their empirical results supported that learning orientation is critical for innovation using data collected from large US firms. Jung et al. (2008) examined how transformational leadership by top managers (CEOs) affect their companies' innovativeness. Using data collected through multiple sources on 50 Taiwanese electronics and telecommunications companies, their results confirmed a direct and positive effect of CEO transformational leadership on organizational innovation. Moreover, their results revealed moderating effects by the uncertainty and competition of market environment.

CHAPTER III

FIRM-LEVEL HUMAN CAPITAL AND PATENTS

3.1 Introduction

Why do firms differ in innovation? Economists have long sought answers to this question because characteristics of innovative firms can have significant implications, not only for firm success but also the economic growth of a country. Along this avenue, literature on testing Schumpeterian hypotheses (Schumpeter, 1942), which argue that large firm operating in a concentrated market is the main engine of technological progress, has offered numerous insights as to how firm size and market structure affect firm innovation. Although it provides an effective framework for exploring the issue, the large body of work also leaves many questions unanswered. Most notably, the literature, which is almost exclusively on firm size, industry characteristics, market structure, and/or part of human capital information, seldom touches the characteristics of firms besides size and strategic choice. Moreover, leaving strategic choices outside the framework results in ignoring the interaction between a firm and its market environment. In this article, we attempt to fill those gaps first theoretically and then by exploiting the detailed and comprehensive firm-level data in Chinese manufacturing industry.

Traditionally, the most important explanatory factor of innovation is R&D spending because it is believed that R&D is the input in producing innovation.[1] However, in essence, those studies are deficient. First, they simply regard R&D as the most important input of innovation without going deeper into the details of R&D and

[1]This idea is expressed through the knowledge production function by Griliches (1979) and has been followed by many studies (Pakes and Griliches, 1984; Hall, Griliches and Hausman, 1986; Hall and Ziedonis, 2001).

the mechanisms of R&D in affecting innovation. Thus, they fail to take other firm resources, mainly a firm's skilled human capital, into consideration. Second, they ignore the non-R&D innovation, which is usually important to firm innovation.

In addition, innovation includes not only R&D innovation but also non-R&D innovation. Generally, there are three types of creative activities that do not require R&D. First, Kim and Nelson (2000) found that many imitative activities, including reverse engineering, do not require R&D, and the imitation is mainly dependent on the firm's technical personnel and engineers. Second, firms can make minor modifications or incremental changes to products and processes, relying on engineering human capital. Moreover, Hansen and Serin (1997) noted that the innovation process in low-and medium-technology sectors is more related to adaptation and learning by doing, based on design and process optimization, rather than from R&D. Third, firms can combine existing knowledge in new ways, for example in industrial design and engineering projects (Grimpe and Sofka, 2009). Due to the large share of firms that innovate without performing R&D, we can conclude that studies that only focus on R&D should not be enough to fully explain innovation differences across firms.[2]

This paper explores firm-level innovation from a human capital point of view. In the theoretical model, two firms compete with each other in a three-stage Cournot, innovation stage and production stage. Skilled human capital level can affect innovation success probability directly and via R&D level indirectly. Managerial human capital can affect firm innovation through their choice of projects and R&D level. We find that a firm's innovation is not only determined by its human capital level, firm characteristics, and its market share, but also might be affected by market environment. In the empirical study, we use two firm-level datasets from China, one from metropolitan cities and one from provincial middle cites. Human capital indicators are skilled human capital (number of highly educated workers), general manager's

[2] We find that around 25% of firms with patents have no R&D spending in both of our data sets.

education and experience, and management team's education and age. We find that a firm's skilled human capital and managerial personnel's education have significantly positive effects on innovation while the management team's age has a significantly negative effect on innovation. The effect of R&D on patents is insignificant in large metropolitan cities while it is positive and significant in provincial middle cities.

Several difficulties arise in our study. First, the form of R&D to be included in the empirical study is difficult to be determined, and what's worse, R&D may be endogenous. Lagged R&D expenditure may also affect a firm's number of patents, but R&D is highly correlated in a firm over time.[3] Simply including contemporary R&D and its lag may bring about serious multicollinearity. Also, we notice that R&D is a long-run plan and it may not distribute evenly over years.[4] Thus, we use average R&D over years in our estimation. Moreover, R&D may be endogenous in a firm's knowledge production function. To decrease the endogeneity of R&D, we exclude current R&D when calculating average R&D over time. Second, the endogeneity of skilled human capital and General Manager's education may bias our estimates. We use instrumental estimates to solve this problem. The instruments we used are city average and industry average of the corresponding variables: average skilled human capital over cities, average skilled human capital over industries, average General Manager's education over industries, average General Manager's education over cities. All the above averages exclude firm itself. In addition, for skilled human capital, we also use the number of applicants for the positions and the average number of weeks those positions are vacant as instruments.

This paper makes three main contributions: (1) it provides both theory and theoretical framework for studying firm innovation in a human capital view, not only

[3]In our datasets, we notice that R&D over time does correlate with each other highly. In fact the correlation coefficient is always above 0.90.

[4]For example, a firm may invest a lot of R&D in one year, but in the following two years, its R&D investment may be much less than this amount.

skilled human capital but also managerial human capital and R&D human capital. Growth theory only provides a theoretical framework to incorporate skilled human capital into innovation. Though there are some empirical studies on firm innovation and human capital, those studies are either on the relationship between firm innovation and managerial human capital or on the relationship between firm innovation and skilled human capital and none of them take care of skilled human capital and managerial human capital at the same time. (2) Using detailed firm-level data, we are able to study the effects of skilled human capital, general manager's education and experience, and management team's education and age. (3) Two datasets from two different levels of cities, metropolitan cities and provincial middle cities enable us to examine the effect of market environment on firm innovation.

One of the limitations in our study is that patent is not a perfect measure of innovation. Pakes and Griliches (1980) observed that patents are a flawed measure (of innovative output); particularly since not all new innovations are patented and since patents differ greatly in their economic impact. Thus, the relationship between human capital and patents cannot fully reveal how much human capital contributes to firm productivity via innovation. Moreover, the patent propensity rate for product and process innovation can differ a lot.[5] To better understand how firm human capital affects product innovation and process innovation, it is also interesting to study how human capital affects new product sales. We will pursue the productivity and new product sales in future research.

The chapter is organized as follows. Section 3.2 introduces firm-level human capital into innovation. Section 3.3 presents a theoretical framework where two firms Cournot compete with each other in a two-stage game. In Section 3.4 we introduce the data. Section 3.5 introduces our methodology strategy. In Section 3.6, we present

[5] Arundel and Kabla (1998) found that the propensity rate for product innovations is 35.9% on average and 24.8% for process innovation on average using Europes largest industrial firms survey in 1993.

our main results and interpret the findings. Section 3.7 presents further investigation. Section 3.8 concludes.

3.2 Human capital and Innovation

Why human capital might be essential in firm innovation study in China? First, according to the resource-based view of the firm, performance differences across firms can be attributed to the variance in the firms' resources and capacities. Resources that are valuable, unique, and difficult to imitate can provide the basis for firms' competitive advantages. Among all the resources in a firm, human capital has long been argued as a critical resource (Pfeffer, 1995). Although human resources may be mobile to some degree, because some capabilities are based on firm-specific knowledge, and others may only be valuable when integrated with additional individual capacities and specific firm resources that may not be mobile (Hitt et al., 2001), the idea that a firm's human capital is critical still holds. Moreover, upper echelon theory argued that organizations are just reflections of their top managers (Hambrick and Mason, 1984). Thus, given the importance of firm human capital, studying firm innovation from a human capital view becomes very natural.

Second, on a macro scale, human capital has long been introduced into innovation in endogenous growth theory (Romer, 1986, 1990; Lucas, 1993). Barro (2001) further proposed that higher human capital stock tends to generate higher growth through at least two channels: on the one hand, more human capital facilitates the absorption of superior technologies from leading countries, and for this channel, schoolings at secondary and higher levels should be especially important; on the other hand, human capital tends to be more difficult to adjust than physical capital. The endogenous growth theory takes human capital as one of the most important inputs in innovation from the macro level and this inspires us to notice the importance of human capital in firm innovation. However, we still know relatively little about firm level human capital and innovation given the difference between micro and macro study.[6]

Third, as a developed country, China has a distinct form of innovation compared to developed country. Nahm and Steinfeld (2014) pointed out that these forms pertain not to upstream research and development (R&D) or new-to-the-world invention, but instead to downstream efforts involving both the redefinition of existing technologies and the commercialization of new ones. Thus, given the downstream type of innovation, skilled human capital has a especially important role in innovation in China compared to that in developed countries. For example, Ge and Fujimoto (2004) describe how Chinese motorcycle assemblers, through reverse engineering, effectively modularized the firm-specific, integral designs of Japanese lead firms. In this process, it is engineering personnel's skill rather than formal R&D matters the most. Thus, innovation in those firms exhibits some peculiar features that R&D would not capture, incurring the risk of underestimating their innovation effort. In fact, in some small and medium firms, innovation often occurs without the performance of formal R&D. Meanwhile, a firm might invest more R&D than it should do because of government subsidies or tax incentive. In this case, R&D may overestimate its innovation capability. Thus, traditional firm innovation study based on R&D remains quite limited, causing a significant bias in understanding firm innovation.

In my study, firm-level human capital can be divided into two types: managerial human capital and skilled human capital. Managerial human capital is embodied in CEOs, top management top and all managers. Top executives have the discretion to control R&D expenditure in firms. Also, because R&D expenditure is a long-term investment that is considerably risky with high failure rates, top managers monitor

[6]For example, human capital in a region or a country cannot be adjusted easily in a short run, while in a firm human capital can always be adjusted through hiring and firing workers and job training. Thus, human capital in a firm is more apt to suffer from endogeneity than in the country level. Moreover, in a firm, the role of the CEO or general manager and the whole management team can also be very important for firm innovation because they are related to the firm's innovation strategy and management.

R&D expenditure closely and adjust its level based on their preferences. Moreover, top management teams have the task of formulating and implementing the firm's strategy (Hambrick and Mason, 1984), and as part of their leadership function, CEOs must coordinate and control team behaviors. Research examining the relationship between managers' personal characteristics and organizational outcomes has taken two different approaches. One approach is to directly assess the psychological attributes of the managers and examine their link to outcomes (Miller et al., 1982). Another approach is to assess demographic characteristics (such as age and education), making the assumption that such characteristics are related to cognitive abilities, attitudes, and expertise (Bantel and Jackson, 1989). In this study, we use the demographic characteristics because it is more practical.

On the other hand, skilled human capital is related to all skilled workers in a firm and it can be seen as a general measure of human capital in a firm and thus it is fundamental to firm's behavior and performance. The mechanisms between skilled human capital in a firm and innovation can be in two channels. First, higher skilled human capital means higher ability of learning by doing and thus can improve a firm's innovation ability. The relationship between learning-by-doing and patents has been studied by Lieberman (1987), and it found that patenting in process innovation in the chemical industry was largely an outgrowth of "learning by doing." Second, skilled human capital and a firm's R&D together affect the firm's innovation through R&D innovation. The complementary relationship is modeled by Romer (1990), where innovation is produced by combining R&D and human capital together. [7]

[7]There's also a probability of knowledge spillover channel which means that when theres more skilled human capital in a firm, there will be more internalization of outside R&D spillover or knowledge spillover into the firm. This has been explained in knowledge spillover literature, such as Audretsch and Feldman (1996). We will save this for future study.

3.3 Theoretical framework

This section presents our theoretical framework and it allows us to see how firms choose their optimal innovation project, optimal human capital and R&D investment, and thus we can derive expressions for firm innovation, which we estimate in the empirical part.

Our framework is based on Rosen (1991) and Howitt and Mayer-Foulkes (2005). We examine firm innovation using a duopoly model with risk-neutral firms in a three-stages game.[8] In the first stage, the firms invest in a risky innovation project. In the second stage, they will choose their own human capital level. In the third stage, the firms choose their R&D level.

Since backward induction can give us subgame-perfect equilibrium, we consider first the output market decision in production stage. We consider an industry consisting of two firms with Cournot competition. The firms produce a single homogenous good and each maximize its single-period profit. Assume the expected output market profit be a function of the firms' constant marginal costs of production c_i, $i = 1, 2$.[9] The inverse demand curve the firms face is linear. The single period profits of the ith firm are given by[10]

$$\Pi_i = [A - 2c_i + c_j]^2 \tag{1}$$

where A is subject to $A - 2c_i + c_j > 0 \; for \; i, j \in 1, 2$. Equation (1) implies that firm i's profit is decreasing with its own cost, but increasing with its rival's cost and

[8]Firms engage in innovation because a successful project lowers their production cost in the sequent output market competition. To simplify our analysis, we follow Spence (1984) and Rosen (1991) and use cost-reducing technology to represent innovation because we can always break a product into a Lancasterian bundle of services and model product improvement as a reduction in the cost of producing services.

[9]Constant marginal cost might be restrictive since in reality firms might have increasing or decreasing marginal cost. When there is increasing marginal cost, other things equal, firms will produce less; vice versa. However, a constant marginal cost assumption doesn't influence our main conclusion as long as it has nothing to do with firm innovation behavior.

[10]The profit function form is derived from Cournot.

thus the firm's profit will increase with its own innovation success but decrease with its rival's innovation success.

In innovation stage, the two firms choose their own R&D project, human capital level, and the level of R&D investment sequentially. Moreover, we assume that they conduct their own projects at the same time and are completed before the start of the output game. Two outcomes may arise for each project: either it succeeds or it fails. The set of R&D strategies from which the firms choose and the outcome of innovation are common knowledge.

There are three substages in the innovation stage. First, both firms choose projects $p^i(i = 1, 2)$ from the continuum of projects, α, in the set $(0, 1)$. Higher values of αrepresent projects that have a greater chance of success at any fixed level of investment. If a project α yields a successful innovation for a firm, then the firm's cost is reduced by $\gamma\alpha$, where γ is differentiable in α and $\gamma'(\alpha) < 0$, which means that as α increases from 0 to 1, the cost reduction will decrease. Therefore, if firm i succeeds in innovation, its marginal cost will become $c_i - \gamma(\alpha_i)$ and if it fails its marginal cost of production will still be c_i but the firm will lose its R&D spending and wage for skilled workers. Projects should be made based on the existing technology base, the firm's human capital, and the market. Optimal project should enable the firm to generate the maximum of expected profit.

Second, for each project α, there is an optimal human capital level. Third, based on its human capital level, a firm then decides its optimal R&D level. We know that innovation as a way of knowledge creation is an activity with a basic element of uncertainty (Maskell and Malmberg, 1999). For project α, based on Howitt and Mayer-Foulkes (2005), we also assume that the success probability of a certain project at time t in firm i is given by

$$\mu_i = S_i R_i \tag{2}$$

Where R_i is firm i's R&D expenditure and S_i is firm i's skilled human capital that can promote innovation. Note that both S_i and R_i are standardized into values with range $[0, 1]$. The relationship of S_i and R_i expressed in equation (2) means that we implicitly assume that they are both complementary and substitute to each other. For example, if we have $S_i = 0.2$ and $R_i = 0.5$, we then get a probability of 0.1. If we only have $S_i = 0.1$, we can have more R&D (in fact $R_i = 1$), we then get the same probability. Thus, in this sense, S_i and R_i are substitutes. On the other hand, when we take the derivative with respect to S_i, we will get R_i, this means that the marginal effect of S_i on innovation success probability is increasing with R_i. In this sense, S_i and R_i are complementary to each other. Note that, even if we have $S_i = 1$ and $R_i = 1$, we still cannot ensure a project will be successful with probability 1 because every project further has its own successful probability in nature. That is, the success probability of innovation in firm i is

$$I_i = \alpha_i S_i R_i \tag{3}$$

This means that a firm's innovation depends not only on its R&D investment R_i, but also the firm's skilled human capital and whether or not it chooses the "right' project. A larger α_i means a much easier project, and a smaller α_i means a more difficult project.

Next, we will use backward induction to examine firm's strategic choices at equilibrium. First, we will examine how firms make their R&D decisions. That is, a firm first takes a project and human capital level as given and choose its optimal R&D level, R_i. Second, when it makes its skilled labor decision, it can forecast its optimal R&D level and it will make its human capital decision based on its forecasting of optimal R&D level. Finally, when firm chooses optimal project, α_i, it can forecast the optimal R&D level and optimal human capital level, S_i. and it will make its project choice decision based on his forecasting. Given project α_i and its human capital level,

S_i, firm i will maximize its expected profit

$$\Pi_i = \alpha_i \mu_i \pi_i^S + (1 - \alpha_i \mu_i) \pi_i^F - rR_i - wS_i \tag{4}$$

where π_i^S is the profit firm i will get if its innovation is successful and π_i^F is its profit if its innovation fails. r is the cost rate for R&D and it may include the interest rate, government incentives and subsidies for firm R&D. w is the wage for skilled workers. Equation (4) states that the expected profit is the firm's expected profit after innovation minus its expenditure on innovation, R&D and wages for skilled workers.

The firm's after-innovation profits π_i^S and π_i^F are determined not only by firm i's innovation, but also firm j's $(j \neq i; i, j = 1, 2.)$ innovation because the two firms Cournot compete with each other in the same market. Firm j also may succeed or fail in innovation; thus, we will have

$$\pi_i^S = \alpha_j \mu_j \pi_i^{SS} + (1 - \alpha_j \mu_j) \pi_i^{SF} \tag{5}$$

$$\pi_i^F = \alpha_j \mu_j \pi_i^{FS} + (1 - \alpha_j \mu_j) \pi_i^{FF} \tag{6}$$

where π_i^{SS} is firm i's profit when both firms succeed in their innovation, π_i^{SF} is firm i's profit when firm i succeeds while firm j fails, π_i^{FS} is firm i's profit when firm i fails while firm j succeeds, and π_i^{FF} is firm i's profit when both firms fail. Obviously, the firm's after-innovation profits π_i^S and π_i^F are the weighted average of two situation respectively. We will take π_i^S in equation (5) for example. Specifically, for firm i, if its innovation succeeds, then its profit is π_i^S. This includes two situations: firm j succeeds or firm j fails. The probability of the innovation success of firm j is $\alpha_j \mu_j$ and the probability of the innovation failure of firm j is$(1 - \alpha_j \mu_j)$. For the two situations, firm i's profits are π_i^{SS} and π_i^{SF} respectively. Then π_i^S is the expected profits under two situations.

We then plug equations (5) and (6) into equation (4), and then take the derivative with respect to R_i and thus we get the reaction function of firm i. Following the same procedure, we then get the reaction function of firm j. Combine the two reaction functions together and we then solve for the optimal R&D, R_i, given α_i,

$$R_i^* = \frac{-r + 4\alpha_j\gamma(\alpha_j)S_j(A + c_i - 2c_j + \gamma(\alpha_j))}{4\alpha_i\alpha_j\gamma(\alpha_i)\gamma(\alpha_j)S_iS_j} \tag{7}$$

Equation (7) shows that a firm's R&D spending is determined not only by its own and rival's technology level (cost function) (c_i, c_j), but also both firms' skilled human capital level and their project choices $(S_i, S_j, \alpha_i, \alpha_j)$. The implication is that the rival's information can be used as instruments for the firm's R&D in empirical study. More specifically, we use the rival firm's R&D as an instrument.

When we take derivatives with respect to different variables and parameters respectively, we can analyze a firm's R&D behavior more specifically.[11] Under assumption $\gamma(\alpha_i)+\alpha_i\gamma'(\alpha_i) > 0$, we find that the derivative of R_i^* with respect to α_i is smaller than zero, implying that to ensure a certain level of innovation success probability firm will choose more R&D spending if they choose riskier project given skilled labor, consistent with our intuition. By taking derivative with respective to c_i and c_j respectively, we get that firm will invest more R&D if it has a less advanced technology and if its rival has a more advanced technology. That is, firms in a market with laggard technology have less R&D investment than firms in a market with advanced technology. Similarly, we get that R&D is increasing with its rival's human capital level, and decreasing with its own human capital level. This is very easy to understand since firm tend to invest more R&D when the competition is more intensive and because of the complementary effect between R&D and human capital level, more human capital tends to induce less R&D. Similarly, we get that firms in a market with higher human capital level have more R&D investment than firms in a market with lower human

[11]The details will be available upon request.

capital level.

Also, firms can chooses its optimal human capital level. We can solve that

$$S_i^* = \frac{\sqrt{(-2A + 4c_i - 2c_j + \gamma(\alpha_j))r}}{(2\sqrt{w\alpha_i\gamma(\alpha_i)}} \tag{8}$$

We can see that firm's skilled human capital decreases with local wage, its rival's cost but increase with its own cost. Also, we can easily that that it increase with R&D cost, r. This is the substitution effect between R&D and human capital level. Similarly to R_i^*, when under assumption $\gamma(\alpha_i) + \alpha_i\gamma'(\alpha_i) > 0$, we find that the derivative of S_i^* with respect to α_i is smaller than zero, implying that to ensure some certain level of innovation success probability, given R&D spending, firm will choose higher skilled human capital level if they choose riskier project, consistent with our intuition.

Finally, we can solve for the optimal project. We will not discuss it in detail since our focus is skilled human capital and Rosen (1991) used the same framework to study the choice of project. Thus, if interested, please see Rosen (1991) for more detail. Here we plot the relationship between profit and projects in Figure 7. We can see that profit is a concave function of project, α, and there's an optimal value, α^*, so that firm can get maximum profits.

Plug R_i^*, S_i^* and α_i^* into (3), we finally get the innovation level at equilibrium

$$I_i^* = I_i(A, c_i, c_j, w, r) \tag{9}$$

Therefore, we can see that at equilibrium, firm i's innovation is determined is determined by market demand (A), its own and rival's technology level, and R&D cost rate and wage.

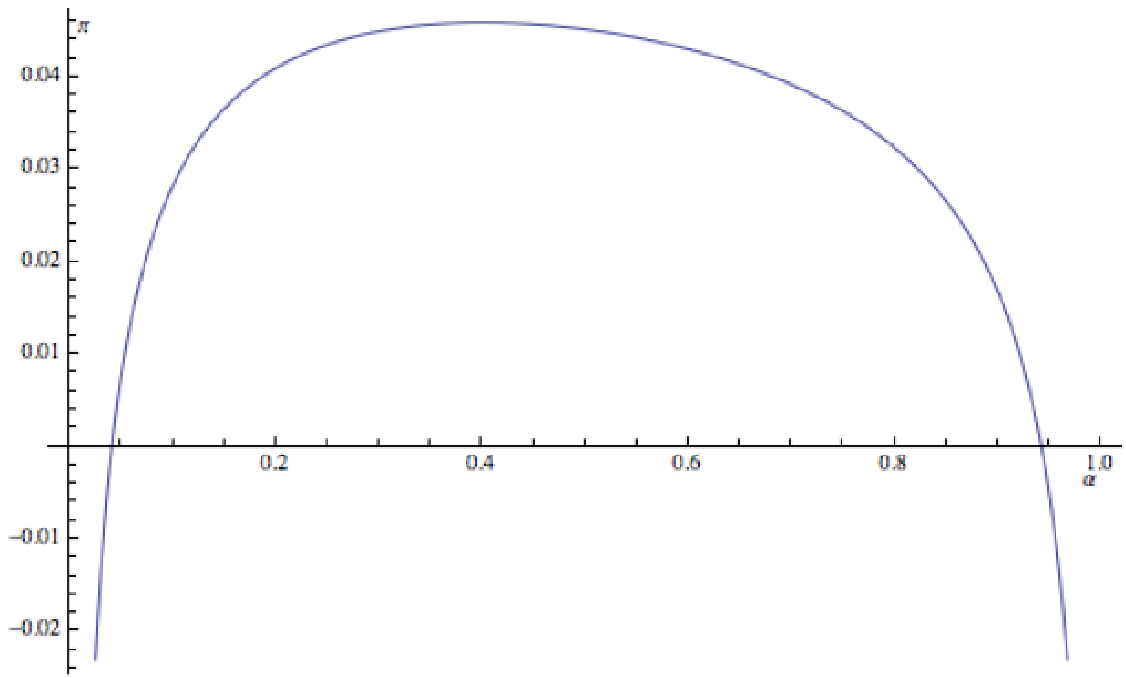

Figure 7: Profit and innovation project

3.4 Empirical strategy

3.4.1 Specification

However, in reality, because of asymmetric information, imperfect decision process, financial constraints, and some other reasons, actual project choice, α_i^r, and actual R&D level, R_i^r, cannot be the optimal levels. That is, $S_i^r \neq S_i^*$ and $R_i^r \neq R_i^*$. Among all the reasons, managerial human capital in a firm is usually a very important factor which affects the difference between actual decision and optimal level. We have

$$|S_i^r - S_i^*| = \varphi^S(M_i) \tag{10}$$

and

$$|R_i^r - R_i^*| = \varphi^R(M_i) \tag{11}$$

where M_i is a firm's managerial human capital, $\varphi^S(M_i)$ is the distance between the actual skilled human capital and optimal skilled human capital, $\varphi^R(M_i)$ is the distance between the actual R&D expenditure and optimal R&D level. Here, we use absolution value of difference to measure distance. Also, we have $\frac{\partial \varphi^S(M_i)}{\partial M_i} < 0$and $\frac{\partial \varphi^R(M_i)}{\partial M_i} < 0$, implying that the higher managerial human capital, the closer the actual decision and the optimal level.

Finally, we get the actual innovation level is determined by

$$I_i^r = I_i(A, R_i, R_j, S_i, S_j, c_i, c_j, M_i, M_j, w, r) \tag{12}$$

In equation (12), we also include R_i, R_j, S_i, S_j to control for the influential factors other than what we consider in our model. In sum, our theoretical framework indicates that a firm's innovation is determined by a combination of the firm's skilled human capital, managerial human capital, firm R&D, and market demand. Thus, in our empirical study, we not only need to include a firm's skilled human capital, managerial human capital, firm's R&D, firm characteristics and market share in our estimation, but also market environment. We use firm characteristics to control for the firm's cost and market share for the demand faced by the firm. Moreover, we use two datasets to control for the effects of market environment (or the other firm) on firm innovation.

From equation (12), we get that skilled human capital in a firm and its managerial human capital together with firm R&D all are vital for a firm's innovation. Thus, they should be included in studying a firm's innovation. In addition, firm characteristics, i.e., firm size, firm age and ownership structure, market structure, industry fixed effect, and city fixed effect are also controlled. Also, R&D is added into specification to control for factors affecting R&D other than the variables we already controlled. In reality, innovation is usually very hard to measure and a common practice is to use the number of patent application to measure innovation. It is assumed that

$$pat_i = \rho_i I_i^r \tag{13}$$

where ρ_i is the patenting propensity ratio of a firm, an it is usually determined by the characteristics of innovation, firm size, government policy, and some other factors inside the firm.

Combined equations (12) and (13), the knowledge production function, or patent production function, in our study is specified as[12]

$$\log(pat_i) = \beta_0 + \beta_1 HC_i + \beta_2 \log(RD_i) + \beta_3 SZ_i + \beta_4 MKTSHR_i + \beta_5 W_i + u_i \quad (14)$$

where pat_i is the number of patents applied for in China, HC_i is human capital indicators, RD_i is R&D expenditure, SZ_i is firm size, $MKTSHR_i$ is market share, W_i is some control variables, such as industry and city fixed effect, and u_i is a disturbance term, assumed to be distributed independently but not necessarily identically across firms, for firm $i = 1, 2, ..., n$.

Though patent number is not a perfect measure for innovation output as we mentioned before, it still constitutes a relevant measure of the technological effectiveness of R&D activity (Griliches, 1990). We use the number of patents applied for in China as our dependent variable though there are data for both number of patents applied for and actually granted in our dataset.[13] There are two reasons for us to use number of patents applied for. One is to decrease the external effect of patent granting process. The other is that in our database, the two variables differ very little and give us similar results.[14] By using patents applied for we implicitly assume that firms apply patent honestly, that is, firms only apply patents when they feel their innovation can

[12]Another way to get empirical specification is to adopt patent production function directly from literature (Kortum and Lerner, 1999, 2000; Hall and Ziedonis, 2001). However, our theoretical framework features two advantages that make them more appropriate for our study. First, it illustrates how firm-level human capital influences firm innovation. Second, it shows how firms affect each other through strategic choices and thus we can see how market environment can affect firm innovation. This provides us theoretical explanation for why we get similar but different results across two datasets in our empirical study.

[13]In addition, we have patented data applied for and actually granted in US, but there's too little useful value and thus we only use patent data in China.

[14]The results are available upon request.

meet the criterion for a patent. This is not an unreasonable presumption since patent application has its cost and firms will not waste its resources.

We use skilled human capital, GM's tenure and education, and the average age and average years of schooling of management team as our human capital indicators. GM's tenure is the years he holds his position. We use number of highly educated workers or skilled workers to measure a firm's skilled human capital. We use GM's graduate degree dummy to account for his education. Education of management is the average years of schooling of the management team.

However, an identification problem, ignored by almost all Schumpeterian studies, arises as we include a firm's skilled human capital level in our estimation because factors affecting a firm's workforce adjustment are very likely to be related to factors affecting the firm's innovation. For example, a firm that wants to be active in innovation tends to hire more highly educated workers. Thoenig and Verdier (2003) mentioned that by employing a larger share of skilled labor, firms can reduce informational leakages and spillovers, which can be freely acquired by outside competitors, and thereby lessen the threat of imitation and technological leapfrogging because of tacit knowledge and non-codified know-how embedded in skilled workers. Moreover, successful innovation may also increase the proportion of skilled workers in the whole workforce (Krueger, 1993) because more advanced technology needs to be complementary to be productive. The endogeneity of skill adjustments in response to technological changes within a firm is also mentioned by Fleisher et al.(2011).

Following Fleisher et al.(2011), for skilled human capital, we use the number of applicants for the positions and the average number of weeks those positions are vacant as instruments. Given labor market supply level, usually a firm with more killed human capital will attract more applicants and it is also much easier for them to get a proper candidate and thus less vacant weeks for those positions than firms with less human capital. On the other side, they are not likely to be correlated with firm's innovation behavior except through human capital. The underlying reason is that number of applicants and number of vacant weeks are largely determined by applicants' behavior, which is independent of firm's innovation behavior, other than through human capital channel. The way firm's skilled human capital affects applicants' behavior is that a firm with higher skilled human capital usually pays higher than firms with less human capital. Thus, they can be instruments for skilled human capital. In addition, we use average skilled human capital over cities excluding firm itself and average skilled human capital over industries excluding firm itself as instruments. The two instruments are correlated with a firm's skilled human capital since a firm's skilled human capital is to some extent determined by its industry characteristics and the city where it locates. Moreover, since we already exclude firm itself from the average values, the correlation between instruments, the average values, and the error term in equation (14) should be very small. Similarly, General Manager's postgraduate degree dummy also might be endogenous and we use city average excluding firm itself and industry average excluding firm itself as instruments.

Another important variable in the patent production function is R&D spending by the firm. However, how to include R&D in the patents estimation equation is still a question. Much of the early work focused on how the lag structure of R&D affects patents (Pakes and Griliches, 1980; Hausman, Hall and Griliches, 1984). They largely concluded that the lag structure is very poorly identified because of the high within-firm correlation of R&D expenditure over time. Moreover, when many lags are included in the model, the estimate of the sum of the coefficients is roughly the same as the estimated coefficient of contemporaneous R&D when no lags are included. Following their conclusion, some literature use only contemporaneous level of R&D in their specification (Hall and Ziedonis, 2001). However, by doing so, two problems might arise. First, R&D expenditure is a long-term investment (Barker and Mueller, 2002). Thus, only including contemporaneous R&D cannot capture a firm's real innovation efforts.[15] In this point of view, an average R&D over years rather than R&D of a certain year is a better innovation input measure for the firm.[16] Second, contemporaneous R&D is very likely to be endogenous. That is, there is a possible correlation between unobserved innovation productivity shocks and R&D level. Thus, we exclude current R&D from the averages to lessen endogeneity.

Firm size is measured by the log of the net value of total assets in the survey year rather than the log of total sales to lessen the correlation between firm size and other variables. Intuitively, firms with more resources will tend to innovate more because it has the ability to innovate. Generally, we expect a positive effect of firm size and when human capital is considered. We use two approaches to study the effect of market environment on innovation. First, we include market share of each firm in our model to account for a firm's market position. Second, we use two datasets, one from metropolitan cities and the other from provincial middle cities, to examine how firms in different markets, a more advanced one and a less advanced one, innovate.

3.4.2 Regression models of count data

The number of patents applied for by a firm is a count variable, so we need to use models of count data. In the following, we will introduce Poisson model, Poisson QMLE and negative binomial model in the framework of linear exponential family (LEF). This presentation follows Cameron and Trivedi (2013). A density $f_{LEF}(y|\mu)$ is a member of a linear exponential family if

$$f_{LEF}(y|\mu) = \exp\{a(\mu) + b(y) + c(\mu)y\} \tag{15}$$

where the function $b(.)$ is a normalizing constant, and $\mu = E[y]$, and the function

[15] Specifically, if a firm decides to develop a new product, its R&D investment structure over time may be a combination of a large initial R&D spending in the first year and some additional R&D investment in the following years. In this case, the R&D level in a certain year cannot represent the firm's innovation endeavor.

[16] For example, in our data, theres a firm with R&D in 1998 of RMB 944.660 million and its R&D in year 1999 and year 2000 are RMB 249.075 million and RMB 191 million, respectively.

$a(.)$ and $c(.)$ are such that

$$E[y] = -[c'(\mu)]^{-1}a'(\mu) \tag{16}$$

where $a'(\mu) = \frac{\partial a(\mu)}{\partial \mu}$ and $c'(\mu) = \frac{\partial c(\mu)}{\partial \mu}$, and

$$V(y) = [c'\mu]^{-1} \tag{17}$$

Different functional forms for $a(.)$ and $c(.)$ lead to different LEF models. From equations (16) and (17), we can see that for LEF family, the variance is proportional to the expectation. Special cases of the LEF include Poisson and binomial (with number of trials fixed), and exponential. For example, the Poisson density can be written as $\exp\{-\mu + y ln\mu - ln y!\}$, which is an LEF model with $a(\mu) = -\mu$, $c(\mu) = ln\mu$ and $b(y) = -lny!$.

A regression model is formed by specifying the density to be $f_{LEF}(y_i|\mu_i)$ where $\mu_i = \mu(X_i, \beta)$, for some specified mean function $\mu(.)$. The MLE based on an LEF, $\hat{\beta}_{LEF}$ maximizes

$$L_{LEF} = \sum_{i=1}^{n} \{a(\mu_i) + b(y_i) + c(\mu_i)y_i\} \tag{18}$$

The first-order conditions can be written as

$$\sum_{i=1}^{n} \frac{1}{v_i}(y_i - \mu_i)\frac{\partial \mu_i}{\partial \beta} = 0 \tag{19}$$

where $v_i = [c'(\mu_i)]^{-1}$ is the specified variance function that is a function of μ_i and hence β.

Under the standard assumption that the density is correctly specified, then we have

$$\sqrt{n}(\hat{\beta}_{LEF} - \beta_0) \xrightarrow{d} N(0, A^{-1}) \tag{20}$$

where $A = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{v_i} \frac{\partial \mu_i}{\partial \beta} \frac{\partial \mu_i}{\partial \beta'} |_{\beta_0}$.

When the density is unknown or misspecified, the estimator is called Quasi-maximum likelihood estimator (QMLE).[17] Gourieroux, Montfort and Trognon (GMT, 1984) show that when the mean is correctly specified, but other features of the distribution such as the variance and density are potentially misspecified, $\hat{\beta}_{LEF} \xrightarrow{p} \beta_0$ so the MLE is still consistent for β_0. Also,

$$\sqrt{n}(\hat{\beta}_{LEF} - \beta_0) \xrightarrow{d} N(0, A^{-1}BA^{-1}) \tag{21}$$

where A is defined above and $B = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \frac{\omega_i}{v_i^2} \frac{\partial \mu_i}{\partial \beta} \frac{\partial \mu_i}{\partial \beta'} |_{\beta_0}$. Note that v_i is the working variance, the variance in the specified LEF density for y_i, whereas ω_i is the variance for the true dgp. Given specification of a true variance function, so $\omega_i = \omega(.)$, one can potentially obtain a more efficient estimator. The negative binomial model with mean μ and variance $\mu + \alpha\mu^2$ is one of the examples. We can see that NB model generalizes the Poisson QMLE model by allowing for an additional source of variance.[18]

3.5 Data

In this paper, we use data from two surveys. The first is "The Study of Competitiveness, Technology & Firm Linkage" conducted by the World Bank in China in 2002. The second is "Investment climate survey" conducted also by the World Bank in 2003. Though with different names, these two surveys are very similar.[19] The first dataset was carried out in 2001-2002, covered firms in five big cities, Beijing,

[17] In application, because it is only ordinary Poisson with robust error, we still call it Poisson in our results analysis.

[18] The other difference between QMLE and NB is that the regression coefficients are fitted differently because different weights are used when estimated and these weights are inversely proportional to the variance.

[19] Both of them collected information on innovation and technology, firm productivity, finance, labor, and the obstacles to doing business, etc. Both are filled up by the senior manager of the main production facility of the firm and the accountant and/or personnel manager of the firm.

Chengdu, Guangzhou, Shanghai, and Tianjin.[20] Most quantitative questions covered the period 1998-2000; most qualitative questions covered only the time of the survey, 2000 (We call the first dataset as Data 2000, thereafter). The second dataset was conducted in 2003 and covered firms in 18 cities, smaller than the cities surveyed in 2000.[21] Most quantitative questions covered the period 2000-2002; most qualitative questions covered only year 2002 (We call the second dataset as Data 2002, thereafter). Both samples consist of both manufacturing and service firms.[22]

The data are randomly selected from all firms in their respective cities and industries. The resulting size range is extreme, with the reported number of production workers ranging from 3 to 83542 in Data 2000 and from 1 to 70169 in Data 2002. In order to reduce the heterogeneity among firms, we restrict our data only in manufacturing industry and also confine our research to the subsample with at least 50 total workers, at least 10 highly educated workers and 10 less educated workers and RMB 3000,000 sales. As a result, there are 624 firms in Data 2000 and 913 firms in Data 2002. We also collect Consumer Price Index (CPI) for each city from the statistic yearbook of the city's corresponding province. We then use them to transform all price-related variables into real value. The same datasets are used through this dissertation. Thus, we will only discuss new variables in the data part of in later chapters but full statistics tables will be presented for convenience.

[20]The sample includes 1548 observations and 1206 variables.

[21]The 18 cities are: Benxi, Changchun, Changsha, Chongqing, Dalian, Guiyang, Haerbin, Hangzhou, Jiangmen, Kunming, Lanzhou, Nanchang, Nanning, Shenzhen, Wenzhou, Wuhan, Xi'an and Zhengzhou. This sample includes 2400 establishments and 1073 variables.

[22]The industries cover electronic components, autos and auto parts, clothing and leather products, electronic and communication equipment, household electrical goods, information technology services, accounting, auditing, and nonbank financial services, business logistics services, advertising and marketing services, and communication services. In Data 2002, chemical products and medicine, biotech products and Chinese medicine, and metallurgical products are also included.

In the survey, there is information on average education level for each occupation.[23] Following Fleisher et al.(2011), we classify the employees into two categories: highly educated and less educated workers. By averaging the workers' schooling codes for each occupation over sample, we designate each occupation level as either highly educated or less educated based on the average schooling of workers in the occupation.[24] Consistent with Fleisher et al.(2011), our highly educated group mainly consists of "engineering and technical personnel" and "managerial personnel (including sales)". We then use number of highly educated workers (L_s) as our skilled human capital measure in a firm. Note that the survey data only provide us the information on number of employees for different occupation for the years 2000 and 1998. We impute employment for different occupations for 1999 in Data 2000 and year 2000 in Data 2002.[25] Finally, we round them to get L_s 2000.

We use self-reported market share to account for market structure. The self-reported market share is rarely used in literature, but it is much better than calculated market share used in previous literature (e.g., Blundell et al., 1999) because to calculate, one needs first to define the market which is usually not an easy task. By using self-reported market share, we don't need to define the market, and moreover the "market" definition used here is related to the firm the most, and thus we can get the real market effect on the firm.

We use number of patents applied for in China as our dependent variable. There is also information about number of patents applied for in the US and number of

[23]In both surveys, workers are classified into: basic production workers, auxiliary production workers, engineering and technical personnel, managerial personnel, service personnel and other employees. Moreover, there is no explicit explanation for "other employees.

[24]We average years of schooling both over the whole sample and over industries, and both indicate the same classification.

[25]We use weighted average employment of 2000 and 1998, using ratio of total employment of 1999 to 2000 and 1999 to 1998 as weights to impute employment for 1999 in Data 2000. Similarly, for Data 2002, we impute the employment for different occupation for year 2000 using weighted average employment of 2002 and 2001 and use the ratio of total employment of 2000 to 2002 and 2001 to 2002 as weights.

patents actually granted in the US; however, useful values are too few.[26] Here, we have two measures of patent, number of patents applied for by firms and number of patent actually granted in China. However, the two measures are very similar.[27] Moreover, when we try to use the two measures as our dependent variables we get similar results. Thus, in this paper, we only present empirical analysis using patent applied for in China as our dependent variable.

We then present a statistics summary for the full sample in Table 1 (Data 2000) and Table 2 (Data 2002). We can see that firms in Data 2000 do better than firms in Data 2002 even if there is a time trend in Data 2002 with average 0.84 patents in year 2000 in Data 2000 and average 0.74 patents in year 2002 in Data 2002. Also, we can see that for both datasets, number of patents increases over time. Generally, firms are bigger in Data 2002 and they have more highly educated workers and more total workers, with around 180 highly educated workers and 950 total employment in Data 2000 and around 162 highly educated workers and 750 total workers in Data 2002. We can see that sales of firms in metropolis (Data 2000) are more than firms in provincial big cities (Data 2002), though the different is not big, all around 0.3 Billion RMB. However, there's very large difference in R&D between two datasets, with around 15-19 Million RMB in Data 2000 and around 2-4 Million RMB in Data 2002. Another important difference between two datasets is that firms in Data 2000 have a higher market share (16.13%) than in Data 2002 (9.01%). In addition, there's little difference in General Manager's education and experience and the firm's age.

Table 3 presents the differences of human capital indicators between firms with

[26] In Data 2000, number of nonzero values for number of patents applied for in US are 2,4,3 for year 2000, 1999 and 1998, respectively; number of nonzero values for number of patent actually granted in US are 4,5,4 for year 2000, 1999 and 1998, respectively. In Data 2002, number of nonzero values for number of patents applied for in US are 2,1,1 for year 2002, 2001 and 2000, respectively; number of nonzero values for number of patent actually granted in US are 2,1,1 for year 2002, 2001 and 2000, respectively.

[27] There are around 8.5%-11% firms apply for patents in 2000 Data while around 7%-9% firms apply for patents in 2002 Data. Meanwhile, there are around 7%-10% firms are actually granted patents in 2000 Data while 7%-9% firms in 2002 Data actually granted a patent.

Table 1: Descriptive Statistics (Data 2000)

	Year	Observation #	Mean	Std. Dev.	Min	Max
Panel A: Patents and Human Capital Variables						
Number of patents applied by firm in China	2000	624	0.84	4.28	0	50
	1999	624	0.67	3.43	0	38
	1998	624	0.39	1.89	0	20
Number of patents applied by firm in China with patents>0	2000	74	7.11	10.56	1	50
	1999	74	5.66	8.46	1	38
	1998	66	3.73	4.64	1	20
Number of highly educated workers in firm (Hundred)	2000	623	1.84	3.25	0.1	41.33
	1999	619	1.76	2.83	0.1	24.96
	1998	623	1.81	2.79	0.1	27.31
Years of schooling of General Manager (GM)	2000	622	14.03	2.30	5	18
Years of GM holding the position	2000	623	5.69	4.44	0	30
GM's postgraduate dummy (=1, postgraduate)	2000	622	0.16	0.37	0	1
Management team's average age	2000	614	36.29	6.63	18	54
Management team's average schooling	2000	615	11.88	1.50	8	18
Number of applicants for skilled position(Hundred)	2000	418	0.37	1.30	0	15.95
Number of weeks to fill last job for skilled positions	2000	431	3.78	7.73	0.5	87.5
Panel B: R&D and Firm Characteristics						
R&D expenditure by firm (Million RMB)	2000	603	19.00	237.06	0	5673.04
	1999	611	15.20	194.83	0	4618.87
	1998	610	15.31	183.43	0	4238.68
Value of total sales (Million RMB)	2000	624	334.31	1828.58	3	31600
	1999	624	318.91	1933.77	3.01	32200
	1998	614	253.86	1613.49	3.04	28900
Total number of employees (Hundred)	2000	624	9.45	15.05	0.5	170.98
	1999	623	9.12	14.59	0.5	184.66
	1998	621	9.46	15.12	0.5	180.59
Net value of total assets (Million RMB)	2000	622	102.79	430.07	0.013	7554.332
Firm's market share	2000	583	16.13	20.53	0.1	98
Firm age	2000	624	17.81	17.37	0	92
Shareholding firms dummy	2000	624	0.16	0.37	0	1
State-owed firms dummy	2000	624	0.24	0.43	0	1
Foreign invested firms dummy	2000	624	0.39	0.49	0	1
Other firms dummy	2000	624	0.21	0.41	0	1

patents and without patents in both datasets. We can see that in both datasets, skilled human capital, or number of highly educated workers is higher in firms with patents than firms without and the difference is statistically significant in both datasets. Also, we found that firms with general manager with postgraduate are more likely to have patents and the difference is statistically significant. Similarly, we find that firms with management team with higher average schooling are more likely to have patents and

Table 2: Descriptive Statistics (Data 2002)

	Year	Observation #	Mean	Std. Dev.	Min	Max
Panel A: Patents and Human Capital Variables						
Number of patents applied by firm in China	2002	910	0.71	3.68	0	77
	2001	910	0.50	2.29	0	31
	2000	910	0.44	2.12	0	28
Number of patents applied by firm in China with patents>0	2002	114	5.68	8.97	1	77
	2001	94	4.85	5.48	1	31
	2000	90	4.48	5.28	1	28
Number of highly educated workers in firm (Hundred)	2002	904	1.57	2.86	0.1	42.81
	2001	902	1.58	3.07	0.1	53.83
	2000	899	1.59	3.19	0.1	60.86
Years of schooling of General Manager (GM)	2002	903	14.15	2.23	5	18
Years of GM holding the position	2002	901	5.86	4.47	1	23
GM's postgraduate dummy (=1, postgraduate)	2002	903	0.17	0.37	0	1
Management team's average age	2002	879	36.50	5.31	20	51
Management team's average schooling	2002	883	12.13	1.50	8	18
Panel B: R&D and Firm Characteristics						
R&D expenditure by firm (Million RMB)	2002	904	4.14	26.45	0	534.97
	2001	892	3.79	32.73	0	782.41
	2000	891	2.53	17.94	0	451.20
Value of total sales (Million RMB)	2002	909	271.05	1246.51	3.11	29700
	2001	906	224.28	938.37	3.09	21300
	2000	901	198.78	763.11	3.01	15800
Total number of employees (Hundred)	2002	909	7.36	13.21	0.5	155
	2001	908	7.45	13.46	0.5	199.06
	2000	904	7.38	13.65	0.5	220.44
Net value of total assets (Million RMB)	2002	905	96.52	411.77	0.11	8207.21
Firm's market share	2002	884	8.96	16.38	1	99.46
Firm age	2002	910	15.96	14.34	2	52
Shareholding firms dummy	2002	910	0.29	0.45	0	1
State-owned firms dummy	2002	910	0.26	0.44	0	1
Foreign invested firms dummy	2002	910	0.22	0.41	0	1
Other firms dummy	2002	910	0.24	0.42	0	1

the difference is also statistically significant. Thus, a preliminary result from data is that human capital in a firm might play an important role in firm's innovation.

3.6 Results

Controlling for city and industry effects, Table 4 reports the results from regressing number of patents applied for on human capital and a firm's other characteristics

Table 3: Human Capital Indicators Comparison between Firms with Patents and without Patents

	Data 2000			Data 2002		
	Firms with patents	Firms without patents	Difference	Firms with patents	Firms without patents	Difference
Number of highly educated workers in firm (Hundred)	4.56	1.47	3.08***(0.38)	3.82	1.30	2.52***(0.32)
Years of GM holding the position	6.46	5.58	0.88(0.55)	5.68	5.90	-0.22 (0.45)
GM's postgraduate dummy (=1,postgraduate)	0.28	0.14	0.14***(0.045)	0.32	0.15	0.18***(0.037)
Management team's average age	35.41	36.41	1.00(0.83)	35.22	36.69	1.47***(0.53)
Management team's average schooling	12.64	11.78	0.85***(0.18)	12.58	12.03	0.55***(0.16)

using two different regression specifications: OLS and Negative Binomial.[28] For all the specifications, the number of patents applied for in China is used as the dependent variable. Columns (1) and (2) in part I in Table 4 (the left half) are estimated using only cross-sectional data in year 2000 in Data 2000 while in columns (3) and (4) in part II in Table 4 (the left half), Data 2002 is used and only cross-sectional data in year 2002 in Data 2002 are used because our human capital indicators are only available for the survey year. All specifications include city dummy variables and industry dummy variables. OLS estimator is the simplest to use and requires the least requirements to be consistent, but it ignores the count nature of the data. Negative Binomial model fits data better, and thus our analysis will be based on it, this is consistent with overdispersion with large number of zero counts in the data (Dhongde, 2014).[29]

To better understand our results, we calculate marginal effects for Negative Binomial in Column (3) and (6) in Table 4 for Data 2000 and Data 2002 respectively. Marginal effects are attained at mean. The first important result we can see is that the number of highly educated workers has a positive and significant coefficient across both datasets, suggesting a positive effect of skilled human capital on innovation. That is, when a firm has more skilled human capital, it will tend to have more innovation. Specifically, we get a marginal effect of 0.0768 using year 2000 data and a marginal effect of 0.0183 using year 2002 data and both are significant at 1% level, indicating that other things equal, when highly educated workers increase 100 people, for firms with mean values the number of patents will increase 0.0768 in Data 2000 and 0.0183 in Data 2002, respectively. We can see that the effect of skilled human capital is quite significant both statistically and economically, and it is robust across

[28]In our datasets, there are about 90% observations with zero patents. Therefore, zero-inflated models are also a possible choice. In fact, Zero-inflated Negative Binomial (ZINB) regression even fits Data 2000 a little bit better than Negative Binomial regression. The criterion we use is the sum of the absolute differences between predicted value and the observed value. However, ZINB fails to converge since Maximum Likelihood Estimation (MLE) encounter flat region when maximize. We also tried Poisson model, but it is dominated by NB regression because of overdispersion.

[29]Our criterion is the sum of the absolute differences between predicted value and observed value.

both datasets. This is consistent with previous studies. Audretsch and Acs (1991) also concluded that skilled labor has a positive effect on innovation using industry level data.[30] Their results indicated that other things equal, when number of skilled labor increases, innovation would also increase. Very interestingly, we notice that the effect of skilled human capital has a larger effect in Data 2000 than in Data 2002 even when there's a time trend in Data 2002, implying in an advanced market environment might have a larger positive effect of skilled human capital on innovation.

General manager's experience is positive and significant both statistically and economically in Data 2000. In NB model, we find the marginal effect is around 0.0451, which means that for an average general manager in an average firm, when general manager holds the position for one additional year, the number of patent application will increase 0.0451. This is consistent with Lin et al.(2011) that showed that general manager's tenure has a positive effect on R&D expenditure. The reason why we get a positive effect of GM's tenure might be that a GM with longer tenure can be more experienced with the firm and the market structure and the technology opportunity in this industry, and he can thus have a good judgment regarding a firm's innovative capacity and market demand. This is especially true for firm innovation that is full of uncertainty. However, there are also some studies that found that general managers tend to make fewer changes in strategy as their tenure increases. Hambrick and Fukutomi (1991) claimed that this lack of change occurs because when tenure increases, GM became conservative and more strongly committed to implementing their own paradigm for how the organization should be run.[31] The positive effect of

[30]Similarly, they defined skilled labor as the percentage of employment consisting of professional and kindred workers, plus managers and administrators, plus craftsmen and kindred workers.

[31]Moreover, there are some researchers who found that CEOs tend to make fewer changes in strategy as their tenure increases. One reason is that with each increasing year of tenure, CEOs have less interest in pursuing strategies of innovation through higher R&D expenditure, preferring instead to emphasize stability and efficiency (Barker and Mueller, 2002). The other reason is that longer-tenured CEOs may lose touch with their organizations' environments and therefore may not make the changes and investments desired to keep the firm evolving over time.

Table 4: Basic Estimation Results

	Year 2000(Data 2000)			Year 2002(Data 2002)		
	OLS	NB	Marginal Effect	OLS	NB	Marginal Effect
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)	(5)	(6)
Number of highly educated workers	0.425***	0.292***	0.0768***	0.536*	0.182***	0.0183***
(Hundred)	(0.162)	(0.0766)	(0.0240)	(0.319)	(0.0548)	(0.00614)
General Manager's tenure (years)	0.201**	0.171***	0.0451***	0.0161	0.0191	0.00192
	(0.0856)	(0.0404)	(0.0108)	(0.0225)	(0.0285)	(0.00287)
General Manager's postgraduate degree	1.004*	0.353	0.105	0.633	0.902***	0.127**
dummy (=1 if has a postgraduate degree)	(0.609)	(0.365)	(0.121)	(0.440)	(0.258)	(0.0528)
Management team's average age	-0.0552	-0.0853***	-0.0225**	-0.00273	-0.0674***	-0.00677***
	(0.0392)	(0.0290)	(0.00930)	(0.0230)	(0.0249)	(0.00262)
Management team's average schooling	0.0370	0.354***	0.0933**	0.0981	0.116	0.0116
	(0.0749)	(0.136)	(0.0376)	(0.0841)	(0.0975)	(0.00964)
Panel B: R&D and Firm Characteristics						
Log (average R&D in previous two years)	-0.00874	-0.0211	-0.00556	0.0294*	0.112***	0.0112***
	(0.0317)	(0.0282)	(0.00794)	(0.0162)	(0.0197)	(0.00243)
Market Share	0.0156	0.0242***	0.00638***	0.0124	0.0311***	0.00313***
	(0.0102)	(0.00754)	(0.00226)	(0.00775)	(0.00521)	(0.000619)
Firm Size (log (net value of total assets))	0.0355	0.0872	0.0230	-0.122	0.0262	0.00263
	(0.141)	(0.137)	(0.0364)	(0.186)	(0.0928)	(0.00934)
Firm Age (year)	-0.00278	0.00505	0.00133	-0.0149	-0.00739	-0.000743
	(0.00840)	(0.0113)	(0.00301)	(0.00952)	(0.0106)	(0.00106)
Shareholding firms dummy	0.0483	-0.307	-0.0733	-0.0966	-0.134	-0.0131
	(0.539)	(0.528)	(0.114)	(0.224)	(0.338)	(0.0323)
State-owned firms dummy	0.0246	0.0464	0.0124	-0.451	-0.872**	-0.0732**
	(0.447)	(0.466)	(0.127)	(0.314)	(0.363)	(0.0287)
Foreign invested firms dummy	-0.00906	-0.406	-0.102	0.426	-0.940**	-0.0753***
	(0.513)	(0.472)	(0.116)	(0.448)	(0.387)	(0.0266)
Constant	0.134	-6.289**		-0.356	-16.17***	
	(2.808)	(2.576)		(1.296)	(1.860)	
City dummies	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
lnalpha		2.350***			1.942***	
		(0.178)			(0.148)	
Adjusted R^2	0.147			0.184		
Number of observations	562	562	562	824	824	824

Standard errors in parentheses: * p <0.10, ** p <0.05, *** p <0.01. Marginal effect for dummy variable is from 0 to 1.

GM tenure in our study indicates that the effect of good judgment is larger than the effect of conservative leadership.

Instead of including GM's college degree in the estimation as in other literature (e.g., Lin et al., 2011), we include GM's postgraduate degree to indicate GM's education because there are more than 70% of firms' GM with a college degree and thus under this situation, the study of a postgraduate degree will be more meaningful. Table 3 shows that GM graduate is insignificant in NB model in Data 2000 while it is significant both in Poisson model and NB model in Data 2002. In NB model, the coefficient of postgraduate degree is 0.127 indicating that for an average GM in an average firm, when a firm's GM with postgraduate degree, its innovation will increase 0.127 compared to firms having GM only with college degree. This means that compared to college education, postgraduate education of GM is more important to innovation. Notably, compared to other factors, this effect is much larger.

Moreover, Table 4 presents that management team's average age has a negative and significant coefficient in both datasets while their average schooling tends to have a positive coefficient among all the models though it is only significant in column (3). Thus, we can get that management team's average age has negative effect on innovation while their average education has a positive effect on innovation. Specifically, in Table 4 we get that the marginal effect of management team's average age in column (6), -0.00677, means that other things equal (at mean), for average management team in an average firm when management team's average age increases one year, the firm's number of patents will decrease 0.00677. Our results are consistent with our intuition and previous management studies. Older executives tend to be more conservative (Hambrick and mason, 1984) and empirical studies have found that older top managers tend to be risk averse (Barker and Mueller, 2002) and follow lower-growth strategies (Child, 1974). One reason is that older executives have less of the physical and mental stamina needed to implement organizational changes (Child, 1974). Another reason is that older managers may have greater difficulty grasping new ideas and learning new behaviors (Hambrick and Mason, 1984) because some cognitive abilities seem to diminish with age, including learning ability, reasoning, and memory. Finally, younger managers are likely to have received their education more recently than older managers, so their technical knowledge should be superior (Bantel and Jackson, 1989).

Meanwhile, Table 4 shows that management team's average schooling has a positive effect and is significant in negative binomial model in part I, implying that the higher education of management team, the more innovation a firm can have. The importance of the top manager's education has been studied in a number of studies. Attained education level is always assumed to be correlated with cognitive ability, and higher levels of education should be associated with higher ability to generate (and implement) creative solutions to complex problems. Hitt and Tyler (1991) found that more educated executives have greater cognitive complexity and such cognitive complexity provides greater ability to absorb new ideas and therefore increases the tendency toward accepting innovations.

R&D in Part II in Table 4 has a positive effect, consistent with previous studies (e.g., Pakes and Griliches, 1980; Scherer, 1983; Brouwer and Kleinknecht, 1999). Its marginal effect is 0.0112, implying that for an average firm when R&D increase 1000 RMB, other things equal, the number of patents will increase 0.0112, which is comparable to Hall and Ziedonis (2001) which reported the coefficient of the log form of R&D is 0.196 using Poisson model.[32] However, we failed to find a significant effect of R&D in Part I in Table 3. Like us, Lieberman (1987) also failed to detect a strong R&D effect, and they argued that their failure might stem from the poor quality of the available R&D data. Different from him, we claim that the relationship

[32]The relationship between patents and R&D has also been studied at the firm level and they reveal a strong link between total corporate R&D expenditure and patents (e.g., Bound et al., 1984; Pakes and Griliches, 1984).

between R&D and innovation, when measured by patents, may be affected by market environment. In theoretical part, we have showed that market environment with advanced technology and more human capital stimulates firm R&D. Also, in the datasets R&D in Data 2000 is around 5 times R&D in Data 2002. Thus, we conclude that it might be that all firms have more than enough R&D over what is needed so the variation in R&D is not the reason why our dependent variable, number of patents, varies.

Market share has a positive effect across all models and is significant in Poisson in Part I and significant in Poisson and Negative Binomial in Part II, strongly supporting Schumpeterian hypotheses. This is not hard to understand. With bigger market share, firms can have more profit, and thus firms can have more resources to put into R&D. This is important because possible failures in financial markets may force firms to rely on their own supra-normal profits to finance the search for innovation (Bhattacharya and Ritter, 1983). Also, with bigger market share, firms can appropriate more profits from more sales using innovation.[33]

Firm size tends to have a positive effect and is significant in Poisson model in Part I in Table 4, and this is consistent with Schumpeterian hypothesis and the literature (e.g., Scherer, 1983); that is, larger firms tend to have more innovation. Holding R&D expenditures constant, large firms are more likely to apply for patents than small firms. The reason may be that larger firms are more likely to have the specialized staff and legal departments that facilitate filing and enforcement of a patent claim (Lieberman, 1987). The coefficient is 0.346 in Poisson in Part I, comparable to Brouwer and Kleinknecht (1999) where the coefficient for firm size is 0.38-0.62 for different datasets, and industry dummies and R&D collaboration are controlled.

[33]See Blundell, Griffith and Reenen (1999) for more reasons.

3.7 Further Investigation

In estimating our equation, we face a possible econometric problem concerning the potential correlation between the independent variables, skilled human capital and General Manager's postgraduate dummy, and unobservable or unmeasurable firm-specific characteristics, such as the quality of human capital. The ordinary Poisson and NB estimates would then be subject to omitted-variable misspecification and bias. One of the traditional ways to correct the bias is to use panel data. With panel data, we can de-mean the variables and thus all time-invariant firm-specific characteristics would be removed. If none of the unobservable or unmeasurable firm-specific characteristics change over time, we will get unbiased estimates. However, for our data, a three-year panel data, most variation of the data is cross-sectional. Applying the de-mean method will then wipe out useful interfirm variation. Thus, in our study, we use cross-sectional data that make the best use of information on firm characteristics.

Though we can always use GMM to deal with endogeneity in nonlinear model, our main method, control function approach, can be more efficient. Let y_1 denote the response variable, y_2 the endogenous explanatory variable, and $\mathbf{z}$ the $1 \times L$ vector of exogenous variables (which includes unity as its first element). Consider the model

$$E(y_1|\mathbf{z}, y_2, r_1) = \exp(\mathbf{z}_1\delta_1 + \alpha_1 y_2 + r_1) \tag{22}$$

where $\mathbf{z}_1$ is a $1 \times L_1$ strict subvector of $\mathbf{z}$ that also includes a constant and r_1 is the error term. Suppose first that y_2 has a standard linear reduced form with an additive and independent error

$$\begin{aligned} y_2 &= \mathbf{z}\pi_2 + v_2 & (23) \\ D(r_1, v_2|\mathbf{z}) &= D(r_1, v_2) & (24) \end{aligned}$$

so that (r_1, v_2) is independent of $\mathbf{z}$. Then

$$E(y_1|\mathbf{z}, y_2) = E(y_1|\mathbf{z}, v_2) = E(\exp(r_1)|v_2)\exp(\mathbf{z}_1\delta_1 + \alpha_1 y_2) \tag{25}$$

If (r_1, v_2) are jointly normal, then $E(\exp(r_1)|v_2) = \exp(\theta_1 v_2)$, where we set the intercept to zero, assuming $\mathbf{z}$ includes an intercept. This assumption can hold more generally, too. Then

$$E(y_1|\mathbf{z}, y_2) = E(y_1|\mathbf{z}, v_2) = \exp(\mathbf{z}_1\delta_1 + \alpha_1 y_2 + \theta_1 v_2). \tag{26}$$

This expectation immediately suggests a two-step estimation procedure. The first step is to estimate the reduced form for y_2 and obtain the residuals. Second, include $\hat{v}_2$, along with $\mathbf{z}_1$ and y_2, in Poisson QMLE or Negative Binomial.

Though in the linear model, control function estimates are identical to 2SLS estimates, in the exponential model, we can obtain a more efficient estimator via control function method. Moreover, we can still take the count data feature and overidentification feature in the second stage of control function by using Poisson QMLE and Negative Binomial model.

Table 5 shows the results of IV estimation and skilled human capital and GM's postgraduate degree dummy are treated as endogenous. The instruments for skilled human capital in Data 2000 are the number of applicants for the positions, the average number of weeks those positions are vacant, average skilled human capital over cities excluding firm itself, average skilled human capital over industries. The instruments for skilled human capital in Data 2002 are the same but we have no information on the number of applicants and average vacant weeks. The instruments for GM's postgraduate are average postgraduate over cities excluding firm itself, and average postgraduate over industries excluding firm itself for both datasets. In Table 5, we show both results from 2SLS and control function using Negative Binomial. As we analyzed above, we will rely on control function since it is more efficient.

In Table 5, we can see that compared to the results in Table 4 where no endogeneity is treated, skilled human capital or the number of highly educated workers, are still very significant and the magnitude also changes very little. The effect of postgraduate degree in Data 2000 now have a much larger effect, more than twice, and now become

Table 5: IV Estimation Results

	Year 2000(Data 2000)		Year 2002(Data 2002)	
	2SLS	NB CF	2SLS	NB CF
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)
Number of highly educated workers	0.500**	0.250***	0.582**	0.188***
(Hundred)	(0.201)	(0.0746)	(0.284)	(0.0454)
General Manager's tenure (years)	0.170***	0.156***	0.0160	0.0192
	(0.0473)	(0.0480)	(0.0222)	(0.0289)
General Manager's postgraduate degree	1.714**	0.722**	0.552	0.841***
dummy (=1 if has a postgraduate degree)	(0.839)	(0.362)	(0.395)	(0.276)
Management team's average age	-0.0424	-0.0503*	-0.00432	-0.0666***
	(0.0501)	(0.0277)	(0.0239)	(0.0248)
Management team's average schooling	0.0334	0.216	0.0983	0.141
	(0.114)	(0.138)	(0.0764)	(0.0958)
Panel B: R&D and Firm Characteristics				
Log (average R&D in previous two years)	-0.0116	-0.0427	0.0264*	0.109***
	(0.0321)	(0.0274)	(0.0150)	(0.0195)
Market Share	0.0238*	0.0223***	0.0126*	0.0320***
	(0.0129)	(0.00750)	(0.00756)	(0.00539)
Firm Size (log (net value of total assets))	-0.102	0.103	-0.158	0.00232
	(0.134)	(0.138)	(0.152)	(0.0894)
Firm Age (year)	-0.0151	-0.0147	-0.0143	-0.00760
	(0.0135)	(0.0140)	(0.00975)	(0.0106)
Shareholding firms dummy	-0.132	0.00527	-0.0962	-0.106
	(0.710)	(0.513)	(0.213)	(0.336)
State-owned firms dummy	-0.473	-0.191	-0.471	-0.795**
	(0.428)	(0.586)	(0.306)	(0.359)
Foreign invested firms dummy	0.0120	-0.553	0.437	-0.940**
	(0.661)	(0.465)	(0.433)	(0.384)
Constant	1.381	-4.688*	-0.0339	-16.34***
	(3.681)	(2.664)	(1.443)	(2.011)
City dummies	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Residual1		-1.062**		-0.206**
		(0.471)		(0.104)
Residual2		-2.733		0.610
		(2.818)		(0.989)
Over-identification Test: Chi2	7.487		1.481	
P-value	(0.112)		(0.477)	
lnalpha		2.011***		1.921***
		(0.199)		(0.151)
Number of observations	354	354	826	826
Adjusted R^2	0.201		0.183	

(1) Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
(2) In this model, both skilled human capital (number of highly educated workers) and GM's postgraduate degree are treated as endogenous. We use the corresponding city average skilled human capital (excluding firm self) and industry average human capital (excluding firm self), number of applicants for skilled position, and number of weeks skilled positions vacant as instruments for skilled human capital. But we have no information on applicants for skilled position and number of weeks skilled position vacant in Data 2002. Thus, for Data 2002, we only use city average and industry average excluding firm itself as IV. For GM's postgraduate degree, we use the corresponding city average and industry average excluding firm itself as IV.)

Table 6: Basic Results with Smaller Sample

	Year 2000(Data 2000)		Year 2002(Data 2002)	
	OLS	NB	OLS	NB
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)
Number of highly educated workers	0.445**	0.263***	0.556	0.166***
(Hundred)	(0.173)	(0.0752)	(0.372)	(0.0608)
General Manager's tenure (years)	0.372**	0.208***	0.00965	0.0130
	(0.163)	(0.0381)	(0.0442)	(0.0355)
General Manager's postgraduate degree	1.408	0.141	1.049	1.327***
dummy =1 if has a postgraduate degree)	(0.938)	(0.397)	(0.725)	(0.295)
Management team's average age	-0.0565	-0.0845***	-0.0134	-0.0915**
	(0.0655)	(0.0306)	(0.0521)	(0.0360)
Management team's average schooling	-0.127	0.311*	0.213	0.267**
	(0.158)	(0.163)	(0.194)	(0.120)
Panel B: R&D and Firm Characteristics				
Log (average R&D in previous two years)	-0.0558	-0.0590*	0.0551**	0.0932***
	(0.0541)	(0.0328)	(0.0268)	(0.0269)
Market Share	0.0323*	0.0242***	0.0174	0.0352***
	(0.0175)	(0.00934)	(0.0112)	(0.00797)
Firm Size (log (net value of total assets))	0.185	0.0709	-0.148	-0.0383
	(0.293)	(0.172)	(0.373)	(0.155)
Firm Age (year)	0.00534	0.0302**	-0.0194	-0.0163
	(0.0145)	(0.0141)	(0.0176)	(0.0112)
Shareholding firms dummy	0.135	0.658	-0.304	-0.130
	(1.302)	(0.732)	(0.541)	(0.405)
State-owned firms dummy	0.0334	0.410	-0.936	-0.437
	(1.003)	(0.562)	(0.715)	(0.428)
Foreign invested firms dummy	-0.173	0.342	0.556	-0.702
	(1.209)	(0.573)	(0.819)	(0.483)
Constant	-0.694	-6.985***	-0.658	-18.20***
	(5.392)	(2.633)	(3.423)	(3.303)
City dummies	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
lnalpha		2.119***		1.662***
		(0.186)		(0.184)
Adjusted R^2	0.170		0.158	
Number of observations	310	310	388	388

(1) Standard errors in parentheses: * p <0.10, ** p <0.05, *** p <0.01.
(2) Different from Table 4, we now restrict the sample to larger firms and now there are around 20% firms with nonzero patents while in Table 4 there are around 10% firms with nonzero patents.

Table 7: IV Estimation Results with Smaller Sample

	Year 2000(Data 2000)		Year 2002(Data 2002)	
	2SLS	NB CF	2SLS	NB CF
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)
Number of highly educated workers	0.511**	0.207**	0.614*	0.186***
(Hundred)	(0.202)	(0.0823)	(0.323)	(0.0443)
General Manager's tenure (years)	0.274***	0.283***	0.00923	0.0205
	(0.0837)	(0.0649)	(0.0420)	(0.0357)
General Manager's postgraduate degree	2.537**	0.810	0.938	1.567***
dummy (=1 if has a postgraduate degree)	(1.248)	(0.498)	(0.650)	(0.326)
Management team's average age	-0.0279	-0.0578	-0.0165	-0.0896**
	(0.0810)	(0.0386)	(0.0518)	(0.0365)
Management team's average schooling	-0.146	0.0172	0.184	0.305**
	(0.191)	(0.178)	(0.160)	(0.122)
Panel B: R&D and Firm Characteristics				
Log (average R&D in previous two years)	-0.0542	-0.0611*	0.0504*	0.0843***
	(0.0545)	(0.0342)	(0.0259)	(0.0275)
Market Share	0.0404**	0.0149	0.0172	0.0373***
	(0.0196)	(0.00919)	(0.0108)	(0.00834)
Firm Size (log (net value of total assets))	-0.0661	0.155	-0.226	-0.0831
	(0.236)	(0.200)	(0.299)	(0.144)
Firm Age (year)	-0.00834	-0.00182	-0.0189	-0.0148
	(0.0185)	(0.0168)	(0.0173)	(0.0116)
Shareholding firms dummy	0.0235	0.248	-0.313	-0.0704
	(1.689)	(0.711)	(0.478)	(0.415)
State-owned firms dummy	-0.955	0.189	-0.983	-0.393
	(1.093)	(0.700)	(0.660)	(0.421)
Foreign invested firms dummy	0.190	0.103	0.559	-0.806*
	(1.584)	(0.723)	(0.766)	(0.464)
Constant	2.110	-3.849	0.469	-20.54***
	(6.172)	(3.144)	(3.366)	(3.653)
City dummies	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Residual1		-1.245**		-0.330***
		(0.509)		(0.0830)
Residual2		-3.560		-1.062
		(3.409)		(1.253)
Overidentification Test: Chi2	7.583		1.686	
P-value	(0.108)		(0.430)	
lnalpha		1.799***		1.602***
		(0.201)		(0.184)
Number of observations	203	203	390	390
Adjusted R^2	0.201		0.157	

(1) Standard errors in parentheses unless otherwise specified: * p <0.10, ** p <0.05, *** p <0.01

(2) Different from Table 5, we now restrict the sample to larger firms and now there are around 20% firms with nonzero patents while in Table 5 there are around 10% firms with nonzero patents. In this model, both skilled human capital (number of highly educated workers) and GM's postgraduate degree are treated as endogenous. We use the corresponding city average skilled human capital (excluding firm self) and industry average human capital (excluding firm self), number of applicants for skilled position, and number of weeks skilled positions vacant as instruments for skilled human capital. But we have no information on applicants for skilled position and number of weeks skilled position vacant in Data 2002. Thus, for Data 2002, we only use city average and industry average excluding firm itself as IV. For GM's postgraduate degree, we use the corresponding city average and industry average excluding firm itself as IV.

significant at 5% level but it almost has no change in Data 2002 in that it still has a very significant effect and only slight lower effect. The significance of GM's tenure has no change and its effect becomes only slightly lower. The effect of management team's age now has a larger effect and it becomes more significant in Data 2000 but it almost has no change in Data 2002. Management team's average schooling still tends to be positive but neither is significant in both datasets. Thus, we can conclude that our main results still hold.

In addition, the residuals from first stage are significant in both datasets indicating that the existence of endogeneity. We also partially test the validity of the instruments by the over-identification test and do not reject the null that the over-identifying instruments are valid assuming a subset of the instruments is valid and identified the model. The error term in our model is very likely to be heteroskedastic, and thus we use robust standard error. Since both Sargan's and Basmann's tests assume that the errors are i.i.d., then these tests are not valid here.[34] Thus, we use Wooldridge's score test of overidentifying restrictions, which is robust to heteroskedasticity (Wooldridge, 1995).

Moreover, to see the influence of too many zeros in our datasets, we further restrict our sample to firms with at least 200 total employment and 30 Million RMB, and we present the basic results in Table 6 with smaller sample and its IV estimation results in Table 7. Compared to results in Table 4, there's little change in Table 6. For example, the coefficient of skilled human capital is still significant and the magnitude of it becomes a little bit smaller, from 0.292 to 0.263 in Data 2000 and from 0.182 to 0.166 in Data 2002. Similarly, compared to Table 5, the results from Table 7 changes also very little. Thus, we conclude that our results are robust even if there's a lot of zeros in our datasets.

[34]See more in Sargan (1958) and Basmann (1960).

3.8 Conclusion

We first explain the reason why we need human capital in our firm innovation study. First, a firm's human capital is its one of the most important resources, and thus it is how we can differentiate firms. This is from resource-based theory. Second, traditional economic theory assumes that firms are homogenous except for market share and firm size. We emphasize that the characteristics of a firm's decision makers, GM and management team, will largely affect their decision-making and thus the firm's performance and other aspects of the firm. Moreover, upper echelon theory argues that a firm is a reflection of top managers. Human capital plays its role in three ways. First, skilled human capital affects a firm's success probability directly, and it can also affect the firm's success probability indirectly via affecting the firm's R&D level choice. Second, R&D human capital determines innovation directly. Third, managerial personnel can affect a firm's project choice and R&D choice. Better managerial personnel will make decisions closer to optimal levels.

Our major findings are as follows. Skilled human capital has a positive effect in both datasets and the effect in Data 2000 is much larger. Moreover, we find that GM's education and experience have positive and significant effects on innovation. Management team's education has a positive effect on innovation while the team's average age has a negative and significant effect on firm innovation. Notably, R&D has a positive and significant effect on innovation in Data 2002 while it is insignificant in Data 2000.

Implications from our results are that: (1) human capital, skilled human capital and characteristics of managerial personnel, is very important in determining firm's innovation. Without considering them, the study of firm innovation may be biased because of heterogeneity. Moreover, when both variables are included, human capital can account for the impact of other innovation, i.e. all the other "on the job learning" or "learning by doing". (2) Controlling only market share or market fixed effect is not enough for firm innovation study. Comparing firm innovation in different market environments is essential to studying how market environment affects a firm's innovation. Thus, besides knowledge spillover, we find that the strategic choice of a firm plays an important role in how the market environment affects firm's innovation. This is consistent with Grabowski (1968) which stated that firm decisions on R&D are strongly influenced by the behavior of competitors. (3) R&D and skilled human capital level might be endogenous in long run. Without dealing with this problem, we might misinterpret their effects.

Our results are subject to two caveats that warrant further research: one relating to innovation measure; the other relating to innovation strategies. First, as we admitted previously, patent number is not a perfect measure for firm innovation because value of patents varies and a lot of innovation is not patented. Further studies on new product sales and Total Factor Productivity (TFP) might provide us with more insights into firm innovation. Second, in our study, we only focus on firm's R&D and non-R&D innovative activities in house without considering a firm's other innovative strategies, like R&D cooperation and licensing from other firms, and so forth. However, in reality, firms will choose their innovative method among all the possible strategies.

CHAPTER IV

FIRM-LEVEL HUMAN CAPITAL AND PRODUCT INNOVATION

4.1 Introduction

Product innovation or new product introductions are vital to most manufacturing firms' growth and prosperity (Scherer, 1984b). For some technology driven industries, the contribution to profits of new products was over 40 percent (Booz Allen and Hamilton 1982). This is particularly true for China. Firms in China face more complex environmental situations than their counterparts in market economies since the relatively underdeveloped government, legal, and financial institutions in China lead to environmental turbulence as well as dysfunctional competition (Nee, 1992; Xin and Pearce, 1996; Peng and Heath, 1996). Firms tend to adopt a product innovation strategy in a turbulent environment because such an environment triggers unlearning of current routines and offers novel opportunities to take advantage of emerging market needs (Miller, 1987). Moreover, extant research suggests that a product innovation strategy leads to higher performance in volatile environments (Covin and Slevin, 1989). Thus, examination of the determinants of firm-level product innovation opens an important window to examine Chinese economic growth.

It is important to make a distinction between product innovation and process innovation since they have different effects on a firm. A pure process innovation simply changes the way in which a product is made, without changing the product itself (except perhaps the price at which it will be sold). In contrast, a pure product innovation creates a new or improved product for sale without any change in the production process —except that more inputs (labor, machine time and materials) may be required (Swann, 2009). Moreover, Greenhalgh and Rogers (2010) believed that the essential effect of process innovation is one of cost reduction in production while the success development of a new product results in a different configuration of changes in both costs and rewards. Cassiman and Vanormelingen (2013) found that compared to process innovation product innovation bring more firm-specific price-cost margins. Specifically, they indicated that product innovations increase markups on average by 5.1% points by shifting out demand and increasing prices and process innovation increases markups by 3.5% points due to incomplete pass-through of the cost reductions associated with process innovation. Product innovation is vital to a firm since it allows the firm to gain a competitive advantage by differentiating its output and increasing the quality and variety of goods.

In this chapter, based on the studies in economics, management and marketing mentioned above, we try to examine firm innovation from a human capital perspective. Specifically, we will take usual economic factors which have effects on product innovation, such as R&D, firm size, market structure, firm characteristics, industry and city fixed effects, into account. At the same time, we also pay attention to different capacities of a firm, such as the ability to identify and understand the users' needs, and effectiveness in marketing. Rather than assessing the capacities directly, we use demographic characteristics of General Manager (GM) and management team (including sales personnel), such as age and education, in our study. Because of the objective and comparable nature of demographic characteristics, our study is more objective and more generalized. Moreover, based on resource based theory, we regard human resources as the base of a firm, and thus firm's skilled human capital is fundamental to firm activity and thus product innovation. Therefore, skilled human capital also need to be included.

We still use the same datasets as in the previous chapter.One important advantage of our datasets is that we have a consistent definition of new product and this ensures the validity of our results. A product innovation can be recognized as a new product innovation only when it was subsequently sold at a price at least 5% higher or lower than the products the firms sold or increased the sales of the main business line by more than 2%. More importantly, our definition can keep product proliferation out of our datasets. Product proliferation is that firms fill up a product space (or characteristics) with slightly different versions of the same products (Swann, 2009). A new product because of product proliferation is not so much the innovation in any one product. Product proliferation is a strategy by firms to segment markets to get more profits, one way to exert price discrimination, and to deter others from entering the market. Thus, we can see when a new product is introduced only because of product proliferation, it will usually be more difficult for it than a new product with real innovativeness to have a higher price or to increase market share.

Another advantage of our datasets is that we have three product innovation measures: new product sales proportion, new product sales and new product innovation count. These measures enables us to study product innovation from three dimensions, the proportional aspect, value, and count and thus we can have a better understanding of product innovation. If we only have information on the proportion of new product sales in total sales, we may fail to distinguish the meaning of the same proportion for firms with different total sales in the same industry or the same competitive market. For example, a proportion of 0.5 never means the same level of innovativeness for a small bread producer with $1000 total sales and its rival firm with $10,000 total sales. But proportion can be very useful across different industries, for example, if we want to compare the innovativeness of a pen producer and a car producer, new product sales will be less appropriate than proportion. New product innovation count is a good way to compare the innovativeness of firms but it also has its drawbacks. For example, when we use new product count to compare, we implicitly assume that each new product innovation the same value. Moreover, we have information on both self-reported market share and number of competitors faced by firm. Thus, we can use number of competitors rather than market share to capture the effect of market structure on product innovation since market share can be highly correlated with product innovation sales and counts.

Lots of studies only include firms with innovation or R&D their samples and therefore might suffer a sample selection problem. For example, Cohen et al.(2002) only include firms with industrial R&D facilities, and thus their study are essentially heavily biased towards large-scale, technologically-intensive firms, despite the inclusion of around 20 small firms. suffer from a sample selection problem when we try to generalize their findings. Our datasets are from random survey data and in nature include both innovative firms and non-innovative firms and thus our results can be generalized without bias.[1] However, there are also some challenges we need to tackle with our datasets. The first challenge is that our three dependent variables, new product sales proportion, new product sales and new product innovation count, are of three data type, continuous, proportional and count data. Different types of data types needs different estimation techniques.

The second challenge we encountered arises from the fact that there are around 50% observations with zero product innovation and the zeros are very likely to come from two different channels. There are some firms with real zero product innovation who don't do any innovation somehow. But there are some firms with zero product innovation in our datasets are not completely non-innovative firms. They were recorded as with zero product innovation because their product innovations haven't meet the definition of product innovation in this survey. In this way, the distribution of the data has been changed around zero. Taking both the types of dependent variables and the more "zeros" into account, we will use zero-inflated beta model in

[1]More specifically, the sampling methodology for the surveys is stratified random sampling. The strata used are firm size, business sector, and geographic region.

estimating new product innovation proportion, Tobit model in estimating new product sales, and negative binomial model in estimating new product counts. We will explain more details in the methodology part. Finally, similar with previous chapter, we need to deal with the endogeneity of our human capital measure. The reason is that if a firm wants to be active in product innovation it will employ a much more educated GM and management team. Same as in the previous chapter, we will still use the same instruments.

We find that for skilled human capital, we find that it also tends to have a positive effect across the three measures, though it is not significant in regression using new product proportion as product innovation measure. When new product sales proportion and new product sales are used as product innovation measures, we find that management team's average years of schooling has a positive effect in determining product innovation, however, its effect in Data 2000 when new product count used as product innovation measure is negative and significant. This indicates that though the results from the three regressions are generally consistent, but they still differ. When compare the results using different innovation measures, we should be very cautious. Another thing is that different from results in last chapter where R&D has a positive effect only in Data 2003, less developed areas, we find that R&D has a very significant and positive effect on product innovation no matter which product innovation measure is used. This indicates that R&D is still important in promoting product innovation and policies supporting investment in R&D is still important. Our results still hold when endogeneity is considered. Notably, in all three models, we all find that general manager's postgraduate degree has a large and significant effect in less developed areas but insignificant at all in more developed areas. The reason might be that in less developed areas, market development is much more incomplete and thus general manger's education matters more. A higher education can enable a general manager to make more insightful decision.

The paper is organized as follows. Section 4.2 reviews related studies. Section 4.3 introduces the data. Section 4.4 presents empirical strategy. In section 4.5, we present our main results and interpret the findings. Section 4.6 concludes.

4.2 Product Innovation and Related Literature

According the definitions by (Swann, 2009), a pure product innovation creates a new or improved product for sale without any change in the production process-except that more inputs might be required and a pure process innovation simply changes the way in which a product is made, without changing the product itself. In reality, these two types of innovation often coexist, for example, an improved products often require some innovations in the production process. However, this never means that we don't need to make a distinction between them. The primary difference between product innovation and process innovation is that product innovation is more closely and directly related to final products and thus to the market, profits, firm performance and firm growth.

Though firm innovation has always been a hot topic in economics, there's much less literature on product innovation rather than general firm innovation in economics. Early from 1970s, Utterback and Abernathy (1975) argued that characteristics of the innovative process and of a firm's innovation attempts vary systematically with differences in the firm's environment, competition and growth strategies and also with the state of development of process technology and they also empirically examined their arguments. They found that at early stage of process development, product innovation dominated while at the mature stage, process innovation would dominate. Thus, we can get that a market with more demand potential can promote product innovation. Swann (2009) clearly distinguished the differences between the two types of innovation and analyzed them respectively in economics. In his analysis, both product and process innovation allow the innovative firm to capture a larger market share but in different ways. Process innovation takes market share symmetrically from both products with higher quality and lower quality while product innovation takes market share more from the products with much higher quality and much less so from the products with lower quality. Moreover, when firms aim to get a larger market share, product innovation might be more helpful. Thus, from previous studies, we can see that it is important for us to focus only on product innovation rather than general innovation.

Not surprisingly, determinants of general firm innovation in economics still applies in product innovation study. For example, the structural characteristics of industry including market opportunities, technological opportunities and appropriability conditions have been examined in the literature (Souitaris, 2002; Dougherty, 1990; Geroski, 1990; Levin et al., 1985). In particular, the Schumpeterian tradition (Acs and Audretsch, 1988), i.e., in the relationship between firm size and/or market structure and innovation, has always been recognized as an important strand. However, most of the current studies on product innovation is centered in management and marketing literature. Many of these investigations adopt resource-based theory, which emphasize the heterogeneity of firms and the role played by internal attributes in firm strategies. In this perspective, each firm possessed a unique set of resources and capacities, tangible and intangible, which have been acquired and over time and which finally determine a firm's strategy and performance. Of them, firm's human capital is especially important. As early as 1960s, Myers and Marquis (1969) found that identifying and understanding the users' needs are very important to product innovation, and new products were more successful if they were designed to satisfy a perceived need than if they were developed to simply take advantage of a new technology. Moreover, Rubenstein et al. (1976) examined 103 projects in US industrial firms and concluded that internal management factors were primary influences on product success. In fact, Hopkins (1980) and Cooper (1975) concluded the principal causes for failure as ineffective product marketing and poor market research. Zirger and Maidique (1990) examined over 330 new products in the electronics industry and suggested the following key factors affect product outcome are the quality of the R&D organization, the technical performance of the product, the product's value to the customer, the synergy of the new product with the firm's existing competences, and management support during the product development and introduction processes. They concluded that successful product innovation is strongly influenced by the firm's understanding of its customers' needs and its effectiveness in marketing. From above, we can see that most of these studies are case studies and directly assess different capacities of a firm involving in product innovation.

4.3 Data

In our datasets there are two important measures of product innovation: percentage of new product sales in total sales (proportion) and number of introduced new products in existing business line (count).[2] For simplicity, we call them proportion measure and count measure respectively. These two measures have been popularly used in innovation studies (Katila, 2002; Hall et al., 2009). Multiplying the percentage of new product sales to total sales, we then get new product sales (value), which is popular in innovation studies (Atuahence-Gima and Li, 2004; Liu and Buck, 2007). Similarly, we call it value measure. Compared to patents used in last chapter, these three measures having the advantage of including not only patented innovation but also non-patented innovation. Moreover, they can indicate market acceptance of a new product while some patented innovations have no market value. We use these three product innovation measures at the same time to try to get a more comprehensive

[2]Compared to datasets in previous chapter, in Data 2002 we drop 3 observations with new product count 1000, 2000 and 3200 respectively since they are far from the mean without them, 11.04. But our main results are robust to the datasets with these extreme values. In this chapter, we have 624 observations in Data 2000 while we have 910 observations in Data 2002, three less than in previous chapter.

understanding. It should be noticed that the percentage measure and value measure include both new products introduced in existing business line and new products in a new business line. The difference is that by entering a new line, firms make a larger change and face more risk than staying in a existing business line. Note for proportion measure and value measure for each year, that is for year 1998-2000 in Data 2000 and for year 2000-2002 in Data 2002. Though they are three-year panel, we only use data in survey year, i.e., year 2000 in Data 2000 and year 2002 in Data 2002, as we only have information on human capital in survey years. In particular, we should note that for the count measure, it is the number of new products introduced from 1998-2000 and thus it is for three years in Data 2000 while in Data 2002, it is the total number of new products introduced from 1999-2002 and thus it is for four years.

To make product innovation more meaningful, the survey has its own definition of product innovation. A new product in a year should be one which was introduced or produced for the first time at the beginning of the year.[3] For example, a new product is recorded as a new product in the survey in year 2000 should be introduced or produced for the first time after January 1, 2000. Most importantly, it should also need to meet at least one of the following criteria: (1) was subsequently sold at a price at least 5% higher or lower than the products the firms sold on January 1, 2000; (2) increased the sales of the main business line by more than 2%.

Figure 8 and Figure 9 presents the distributions of three product innovation measures over cities in both datasets. In both figures, all histograms indicates that distribution of product innovation over cities is not even. Notably, the distributions of three measures are of different patterns even for the same datasets. Specifically, in Data 2000, when we use count measure, Beijing has the most product innovation while it is Shanghai when proportion measure and value measure are used. The city with the least product innovation is Guangzhou when proportion measure is used, Chengdu

[3]We can see that in this definition, innovation defined as "new to the firm".

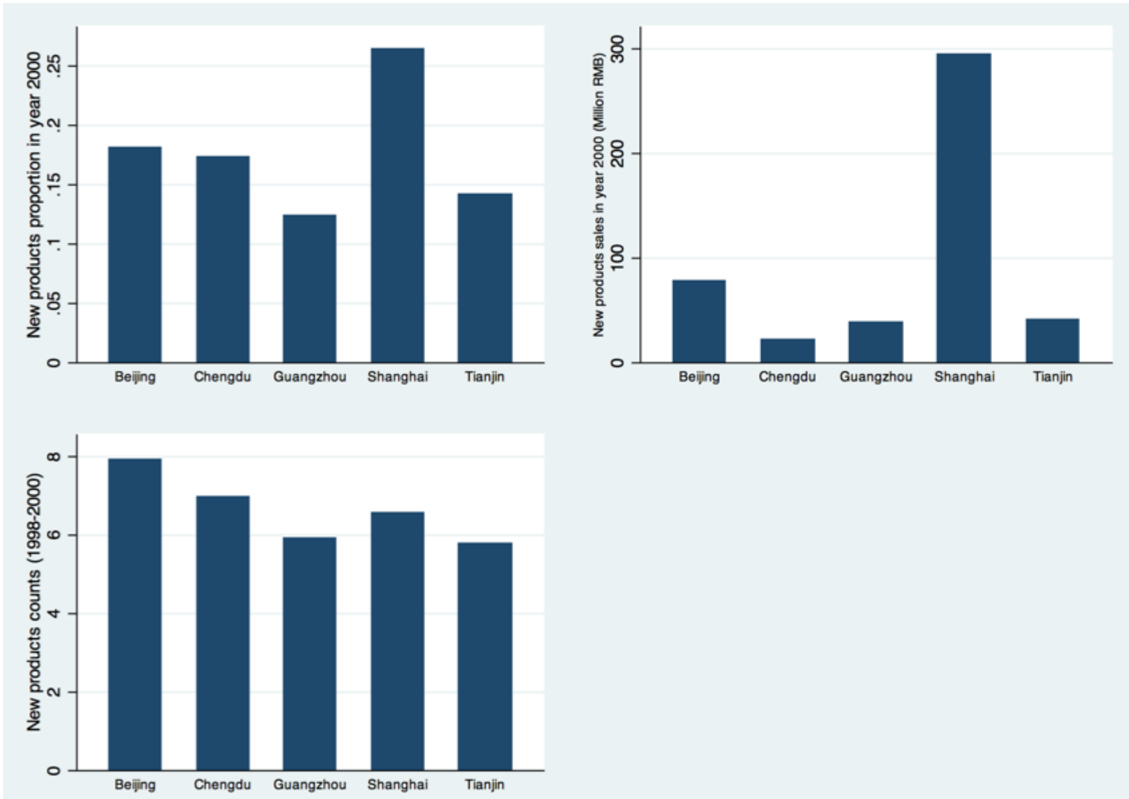

Figure 8: Distributions of Three Product Innovation Measures over Cities (Data 2000)

when value measure is used and Tianjin when count measure is used. Similarly, three measures indicate different distributions over cities in Data 2002. Chongqing has the largest product innovation proportion while Kunming has the smallest product innovation proportion; Shenzhen has the most new product innovation sales while Kunming and Nanning has the least; Hangzhou has the largest product innovation counts while Jiangmen and Kunming has the smallest product innovation by counts. We can see that these three measures have correlated with each other but at the same time they differ a lot. We present the correlation matrix among three measures in Table 8.We can see that across both datasets, the correlation coefficient between proportion measure and value measure is largest, around 0.3, followed by the correlation coefficient between proportion measure and count measure, which is around 0.2. The correlation coefficient between value and count is the smallest, around 0.1. Thus, we can see any conclusions derived by comparing among measures should be cautious.

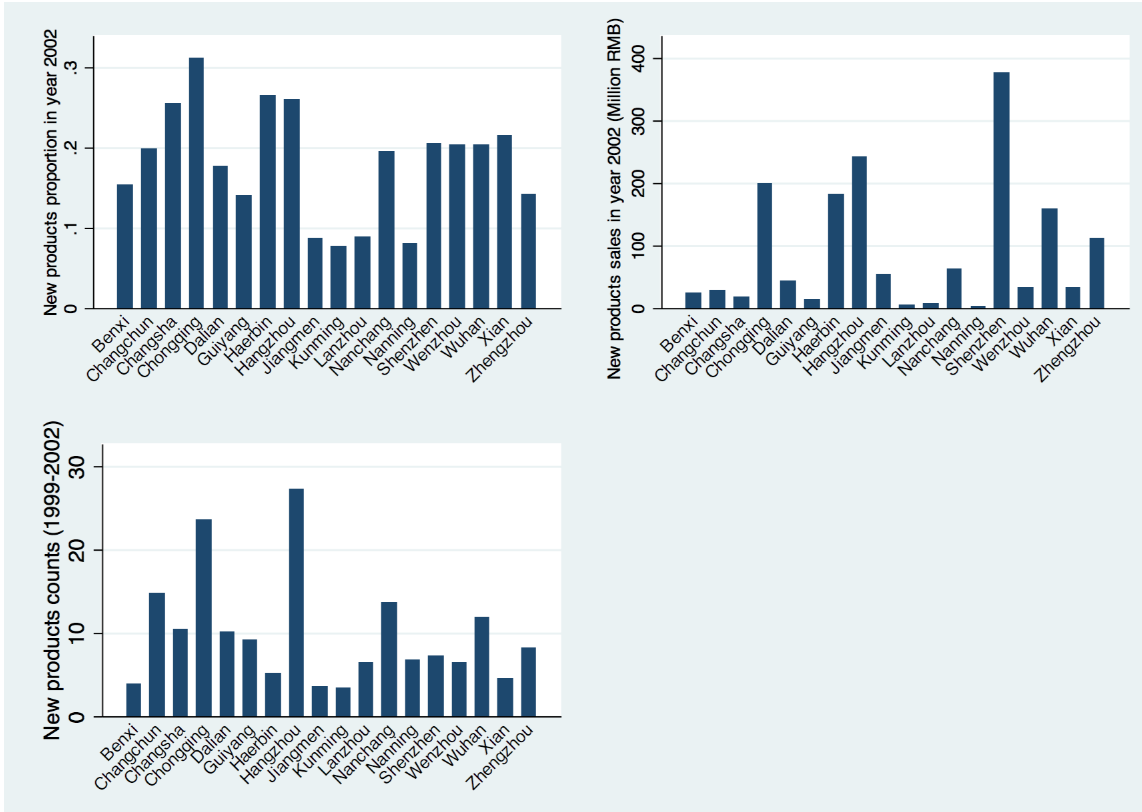

Figure 9: Distributions of Three Product Innovation Measures over Cities (Data 2002)

Table 8: Correlation Matrix among Three Product Innovation Measures (Proportion, Value, and Count)

	Data 2000			Data 2002		
	Proportion	Value	Count	Proportion	Value	Count
Proportion	1.00			1.00		
Value	0.30	1.00		0.35	1.00	
Count	0.25	0.14	1.00	0.20	0.11	1.00

Besides information on new products introduced, we have another important information in this chapter and it is the number of competitors each firm faces. More specifically, it is the number of competitors a firm has within its main business line in domestic market. It is coded as five categories: 1 if the number of competitors is between 1 and 3 (including); 2 if between 4 and 6 (including); 3 if between 7 and 15 (including); 4 if between 16 and 100; 5 if more than 100. On average, there are 2.97

competitors faced by firms in Data 2000 and 3.47 competitors faced by firms in Data 2002. That is, markets in Data 2002 are more competitive than those in Data 2000. This is consistent with self-reported market share, which is 9.01 for firms in Data 2002 while it is 16.13 for firms in Data 2000. Different from previous chapter, we use number of competitors rather than market share to account for market environment as market share, measured using total sales, should be very closely related to total sales which is closely related to the value measure and the proportion measure of product innovation.

We then present statistics summary in Table 8 (Data 2000) and Table 9 (Data 2002). We can see that the average number of products introduced by firms in existing line from 1998-2000 is 6.68 in Data 2000 while it is 11.04 in Data 2002. It seems that there are more product innovation in Data 2002, judging only from these new product counts. However, we still need to notice that 6.68 new products are for *three* years but 11.04 are for *four* years and it is only for *existing* business line rather than total number of introduced new products, including both *existing* and *new* business line. Thus, until now we cannot see that firms in Data 2000 has less product innovation than in Data 2002. There are 321 firms out of total 621 firms (51.69 %) with at least one new product in existing line and 530 firms out of 900 firms (58.89 %) has non-zero new products in existing line. Thus, there are slightly more firms with non-zero new products introduced in the existing business line. When it comes to the percentage of new products in total sales, the two datasets have very similar percentage (13%, 15%, 18% in year 1998, 1999, 2000 respectively) in Data 2000 versus (13%, 16%, 19% in year 2000, 2001, 2002 respectively). Similarly, the percentage for firms with non-zero new products introduced is similar too and is around 30% for both datasets. Though new product sales differ somewhat the years before the survey year, 40.98 and 79.34 Million RMB (around 4.95 and 9.58 Million dollars)in Data 2000 versus 63.97 and 69.91 Million RMB (around 7.73 and 8.45 Million Dollars) in Data 2002, it is

Table 9: Descriptive Statistics for Product Innovation Study (Data 2000)

	Year	Observation #	Mean	Std. Dev.	Min	Max
Panel A: Product Innovation and Human Capital Variables						
Number of introduced new products in existing business line	1998 to 2000	621	6.68	26.73	0	300
Number of introduced new products in existing business line(>0)	1998 to 2000	321	12.92	36.11	1	300
Percentage of new product sales in total sales	2000	607	0.18	0.27	0	1
	1999	610	0.15	0.24	0	1
	1998	606	0.13	0.23	0	1
Percentage of new product sales in total sales(>0)	2000	300	0.36	0.28	0.01	1
	1999	268	0.33	0.27	0.01	1
	1998	244	0.31	0.27	0.01	1
New Products sales (Million RMB)	2000	607	95.76	474.83	0	6952.25
	1999	610	79.34	716.98	0	2894.08
	1998	598	40.98	185.65	0	2149.95
Number of skilled workers (Hundred)	2000	623	1.84	3.25	0.1	41.33
	1999	619	1.76	2.83	0.1	24.96
	1998	623	1.81	2.79	0.1	27.31
Total number of employees (Hundred)	2000	624	9.45	15.05	0.5	170.98
	1999	623	9.12	14.59	0.5	184.66
	1998	621	9.46	15.12	0.5	180.59
Years of schooling of General Manager (GM)	2000	622	14.03	2.30	5	18
Years of GM holding the position	2000	623	5.69	4.44	0	30
GM's postgraduate dummy (=1, postgraduate)	2000	622	0.16	0.37	0	1
Management team's average age	2000	614	36.29	6.63	18	54
Management team's average schooling	2000	615	11.88	1.50	8	18
Panel B: R&D and Firm Characteristics						
R&D expenditure by firm (Million RMB)	2000	603	19.00	237.06	0	5673.04
	1999	611	15.20	194.83	0	4618.87
	1998	610	15.31	183.43	0	4238.68
Value of total sales (Million RMB)	2000	624	334.31	1828.58	3	31600
	1999	624	318.91	1933.77	3.01	32200
	1998	614	253.86	1613.49	3.04	28900
Self-reported Market share	2000	583	16.13	20.53	0.1	98
Self-reported number of competitors	2000	576	2.97	1.30	1	5
Net value of total assets (Million)	2000	622	102.79	430.07	13	7554.33
Firm age	2000	624	17.81	17.37	0	92

quite similar in the survey year across the two datasets, 95.76 Million RMB (around 11.57 Million Dollars) versus 99.56 Million RMB (around 12.03 Million Dollars).[4] Also, judging from the value and proportion measures, new product innovation in both datasets increases with time. In addition, the largest number of new product

[4]We use the exchange rate at that time. The World Bank annual average exchange rate for US dollar to Chinese yuan (1 US to Chinese yuan) is Exchange rate is 8.279, 8.278,8.278,8.277,8.277 for year 1998-2002.

Table 10: Descriptive Statistics for Product Innovation Study (Data 2002)

	Year	Observation #	Mean	Std. Dev.	Min	Max
Panel A: Product Innovation and Human Capital Variables						
Number of introduced new products in existing business line	1999 to 2002	900	11.04	39.32	0	672
Number of introduced new products in existing business line(>0)	1999 to 2002	530	18.75	49.82	1	672
Percentage of new product sales in total sales	2002	887	0.19	0.26	0	1
	2001	890	0.16	0.24	0	1
	2000	885	0.13	0.22	0	1
Percentage of new product sales in total sales(>0)	2002	485	0.35	0.27	0.0001	1
	2001	444	0.31	0.25	0.0001	1
	2000	409	0.28	0.25	0.0028	1
New Products sales (Million RMB)	2002	886	99.56	541.87	0	8609.11
	2001	886	69.91	370.91	0	4602.50
	2000	876	63.97	370.08	0	4739.06
Number of skilled workers (Hundred)	2002	904	1.56	2.86	0.1	42.81
	2001	902	1.57	3.05	0.1	53.83
	2000	899	1.58	3.18	0.1	60.86
Total number of employees (Hundred)	2002	909	7.32	13.15	0.5	155
	2001	908	7.40	13.38	0.5	199.06
	2000	904	7.33	13.58	0.5	220.44
Years of schooling of General Manager (GM)	2002	903	14.15	2.23	5	18
Years of GM holding the position	2002	901	5.87	4.47	1	23
GM's postgraduate dummy (=1, postgraduate)	2002	903	0.17	0.37	0	1
Management team's average age	2002	879	36.51	5.31	20	51
Management team's average schooling	2002	883	12.11	1.53	5	18
Panel B: R&D and Firm Characteristics						
R&D expenditure by firm (Million RMB)	2002	904	4.13	26.45	0	534.97
	2001	892	3.78	32.73	0	782.41
	2000	891	2.52	17.94	0	451.20
Net value of total assets (Million RMB)	2002	909	268.80	1244.89	3.11	29700
	2001	906	222.96	937.67	3.09	21300
	2000	901	197.79	762.64	3.01	15800
Self-reported Market share	2002	884	8.96	16.39	1	99.46
Self-reported number of competitors	2002	893	3.46	1.36	1	5
Asset value (Million)	2002	905	96.30	411.73	0.11	8207.21
Firm age	2002	910	15.94	14.32	2	52

innovation is 300 in Data 2002, much smaller than it in Data 2003, which is is 672. However, the biggest new product sales don't differ so much, 6952.25 Million RMB (around 839.85 Million Dollars) in year 2000 in Data 2000 versus 8609.11 Million RMB (around 1040.12 Million Dollars) in year 2002 in Data 2002.

4.4 Empirical Framework

4.4.1 Estimation Specification

Based on above analysis, we then specify our empirical specification as

$$\log(product_i) = \alpha_0 + \alpha_1 \log(HC_i) + \alpha_2 \log(RD_i) + \alpha_3 X_i + \varepsilon_i \tag{27}$$

where $product_i$ is firm is new product innovation measure, new product sales proportion, new product sales, or number of new product; HC_i is the human capital measures; RD_i is firm's R&D input of last period; X_i is other firm characteristics; ε_i is error term.

For new product innovation, we have information on new product sales proportion (proportion measure), new product sales (value measure) and number of new products (count measure). Each of them measure different dimensions. First, new product sales proportion is the share of new product sales in total sales in a year. The feature that it contains no unit makes it a good measure under the assumption that firms compared are with the same firm size or total sales. However, if the assumption fails the proportion measure will be less meaningful since the same new product proportion means different for firms with different total sales. For example, a large new product proportion might mean very few new product innovation for a firm with small total sales while a small new product proportion for a large firm might means a lot of product innovation. Second, new product sales stands for realized benefits of product innovation for a firm. Thus, we can use it to compare any product innovation given the value is measured correctly and prices don't differ a lot. This feature makes it a very popular innovation measure. However, its limitation is that the prices might differ a lot and the value of product innovation might be measured incorrectly. Third, the number of new product is meaningful in that it provides a comparable measure independent of price and product categories and it is also used a lot in previous studies (e.g., Wu, 2013). Compared to patents, it together with other two measures includes not only patented innovation but also non-patented innovation. But we need to make an assumption that all innovation are of the same value before we use it to compare. Therefore, each measures have its advantages and disadvantages. By using them together, we can gain more understanding about product innovation in a firm.

Like in the previous chapter, our human capital measures are: skilled human capital, or the number of highly educated workers, general manager (GM)'s tenure, GM's postgraduate degree dummy, management team's average schooling and average age. Skilled human capital is used to measure a firm's skilled human capital roughly and it serves as the base of human capital in a firm. It may includes technical personnel's human capital, managerial personnel's human capital and salesmen's human capital and so on. However, it doesn't include basic production worker (usually with high school education or below)'s human capital. The information on GM and management team is to measure how managerial human capital might affect product innovation. The reason why managerial human capital is important to product innovation roots in knowledge search literature and attention-based theory. Knowledge based theory argues that searching broadly enables a firm to see more new technological developments, and learn from them to finally boost its innovation capacities (Katila, 2002). Attention-based theory suggests that the effectiveness of knowledge search is constrained by limited managerial attention, and engaging with too many channels results in a poor allocation of managerial attention (Leiponen and Helfat, 2010). We expect that a firm with more skilled human capital will have a higher knowledge search capacity. GM and management teach with higher education have more knowledge search capacities and also might also can allocate their attention more efficiently. Management team's age is very likely to have an negative effect on knowledge search capacity and attention an the main reason is that older people usually have less energy physically. Moreover, the knowledge structure of older people might be less compatible with current technological development than younger people. However, we are uncertain about the effect of GM's tenure since a GM with more experience can have better judgment and more insight when he makes decisions. On the other hand, a GM holding his position too long might also stick to his own philosophy too much, and he is reluctant to make any changes even when his philosophy is not right. The effect of GM tenure on product innovation is determined by which effect dominates.

R&D is pursued by firms to order to get innovation. It should have an positive effect with product innovation. However, when R&D spending itself becomes one of the firms objective, it might has no effect on innovation. For example, in the landmark document "The National Medium-and Long-Term Plan for the Development of Science and Technology (2006-2010) (MLP), it clearly states that one of its goal is to increase China's gross expenditure on R&D to 2.5 percent of GDP by 2020 from 1.3 percent in 2006. R&D itself cannot produce innovation. There should be complementary human capital. If R&D is higher than needed given its human capital level, R&D will have no effect on innovation. The Scientific Activity Predictor form Patterns of Heuristic Origins (SAPPHO) study also identified the mediating role of human capital in influencing the effect of R&D on innovation. First, R&D teams must be efficient and effective in their development efforts. The second is that there should be an executive champion, a senior member of the firm with power and authority who fought for the product. Product champions facilitate the allocation of resources to the development efforts, and stimulate cooperation and communication between the functional groups, which are also important in product success.

Market structure, especially competition, should be very important for product innovation since it influence firm's innovation incentives and the probability of product innovation success. Arrow (1962) showed that a secure monopolist gains less from perfectly patentable process innovations than would a competitive firm facing the same market demand. Chen and Schwartz (2013) showed that compared to Arrow's result for process innovations, the gain from a product innovation can be larger to a secure monopolist than to a rival who faces competition from independent sellers of the old product. Moreover, monopolist usually have more resources to develop more products and to cover the possible loss from developed and introducing new products. Therefore, we expect that competition should have a negative effect on product innovation. Since market share is obviously endogenous for new product sales, here we use number of competitors for a firm to account for market competition. Thus, we expect that the number of competitors should have a negative effect on product innovation.

In addition, we also control for firm size, firm age, ownership dummies, city fixed effects and industry fixed effects. We want to emphasize the importance of industry effect on product innovation since different industries mean different product life cycle and technology stage. Balachandra and Friar (1997) argued that models that do not take contextual issues into account may lead to erroneous conclusions. Link (1987) found that in the high-tech field, the technology is developed very rapidly, and so new product introductions come quickly.

There are three challenges with respect to estimation. First, our three product innovation measures, the new product sales proportion, new product sales, the number of new products are of three different types, proportion data within 0 to 1, continuous data, and count data. Therefore we need different techniques to deal with them. For proportion data, since its value is within 0 to 1 (0 and 1 are included), the prediction using OLS or count data regression technique such as poisson, is very likely to be out of the range and thus cannot be used. Similarly, for the number of new products, we need to use count data techniques, such as poisson and negative binomial. Second, there are about half zeros in our datasets. Also, according to our new product definition, zero product innovation might come from at least two sources: firms with true zero new product introduced and firms with new product innovation introduced but fail to meet the new product definition. Combined the first and second issues, we will use zero-inflated beta regression for proportion measure, Tobit model for value measure and negative binomial for count data. For proportion data, generalized linear model method (GLM) proposed by Pakes and Wooldridge (1996) is also a good way to model proportion model but it cannot deal with the inflated zero problem. For value measure, Tobit model is used to deal with inflated zero problem. For count data measure, both negative binomial model and zero-inflated negative binomial can deal with count data feature and zero inflated problem. But when we compare predicted values with observed values, we find that negative binomial fits the data better. Since negative binomial model has been introduced in last chapter, in the following, we will only explain Tobit model and zero-inflated beta model in details. Third, skilled human capital and general manger's education are very likely to be endogenous. Firms plan to be active in product innovation introduction will hire general manager with higher education and more highly educated workers. Like in last chapter, for number of highly educated workers, we will use number of applicants for the skilled positions, vacant weeks of skilled position before filled, city average and industry average both excluding firm itself as instruments. Note that for Data 2002, we have no information on applicants and vacant weeks, so we only use city average and industry average excluding firm itself. For general manager's postgraduate degree dummy, we use city and industry average excluding firm itself as instruments. For all the models, we use control function technique.

4.4.2 Estimation Techniques

Tobit model is devised by Tobin (1958) for situations where y^*, the latent variable, is observed for values greater than 0 but is not observed for values of zero or less. The latent variable, y^*, is linear in regressors with additive error that is normally distributed and homoskedastic. It is defined as

$$y^* = \beta_0 + \mathbf{x}\beta + u \tag{28}$$

with the error term

$$u|\mathbf{x} \sim \mathcal{N}(0, \sigma^2). \tag{29}$$

and

$$y = \max(0, y^*) \tag{30}$$

The above equations implies that the observed variable, y, equals y^* when $y^* \geq 0$, but $y = 0$ when $y^* < 0$. Thus, observed 0's on the dependent variable can mean either a "true" 0 or censored data. In our situation, a new product can be counted as a new product if its price is 5% higher price than its current products or increase the total sales by more than 2%. Thus, even if a firm has some new product introduced but cannot meet the above criterion, it is still a product with "0" product innovation. Moreover, firms without any new product introduced, that is, less than "0" product innovation, are also counted as firms with "0" product innovation. But they are two different types of firms with respect to product innovation. Thus, the observed "0", the lower limit, is not necessarily zero, nor is it the same for all firms. Thus, OLS estimator are biased downward and we need to use Tobit model in our analysis.

Maximum-likelihood estimation of the Tobit model is straightforward. Since y^* is normally distributed, y has a continuous distribution over strictly positive values. More specifically, the density of y given $\mathbf{x}$ is the same as the density of y^* given $\mathbf{x}$ for positive values, and we have

$$P(y = 0|\mathbf{x}) = P(y^* < 0|\mathbf{x}) = P(u < -\mathbf{x}\beta|\mathbf{x}) = 1 - \Phi(\mathbf{x}\beta/\sigma) \tag{31}$$

where $\Phi(.)$ is the cumulative density function of the standard normal distribution. We have absorbed the intercept into $\mathbf{x}$ for notational simplicity. We further denote

$\phi(.)$ as the density function of the standard normal distribution, we then have the log-likelihood function

$$\log L(\beta,\sigma) = \sum_{y_i>0}(-\log\sigma + \log\phi(\frac{y_i - \mathbf{x}\beta}{\sigma})) + \sum_{y_i=0}\log(1-\Phi(\frac{\mathbf{x}\beta}{\sigma})) \tag{32}$$

From equation (28), we can see that β_j measure the partial effects of x_j on $E(y^*|\mathbf{x})$, where y^* is the latent variable. In our analysis, if our goal is to understand the underlying *propensity* to innovate, then we will be interested in the marginal effect of $\mathbf{x}$ on y^*. However, we are more interested in understanding the amount of product innovation by innovative firms along, that is, the effect of $\mathbf{x}$ on $y|y>0$. Moreover, we are interested in the effect of $\mathbf{x}$ on the probability of firms having product innovation, $P(y>0|\mathbf{x})$. The three marginal effect expressions are derived using standard results on moments of Tobit, as follows[5]

$$\frac{\partial E(y|y>0,\mathbf{x})}{\partial \mathbf{x}} = \beta(1-\delta(-\frac{\mathbf{x}\beta}{\sigma})) \tag{33}$$

where $\delta(\alpha) = \lambda(\alpha)(\lambda(\alpha)-\alpha), \lambda(\alpha) = \phi(\alpha)/(1-\Phi(\alpha))$, and $\alpha = -(\mathbf{x}\beta/\sigma)$.

$$\frac{\partial P(y>0|\mathbf{x})}{\partial \mathbf{x}} = \frac{\beta}{\sigma}\phi(\frac{\mathbf{x}\beta}{\sigma}) \tag{34}$$

Marginal effect expressed in equation (33) measures how the expected value of product innovation changes for firms with non-zero product innovation (innovative firms) when independent variable, $\mathbf{x}$ changes. Marginal effect expressed in equation (34) measures how the probability of being a firm with product innovation changes as $\mathbf{x}$ changes.

Another important methodological issue arises when considering the modeling of a proportions variable such as our dependent variable, the share of new product sales in total sales. Two features of this variable should be taken into account when estimated

[5]See more in the chapter 17 in Wooldridge (2012).

as the dependent variable. The first feature is that, as a proportion, its value lies in [0,1] (including 0 and 1). One way to handle this is to use logit transformation[6]

$$y = \frac{1}{1 + \exp(-\mathbf{x}\beta)} \tag{35}$$

However, as we can see from equation (35) the transformed dependent variable cannot be 0 and 1, but we have almost half of the observations with 0 proportion of product innovation and thus this method is not appropriate for us.[7] Pakes and Wooldridge (1996) proposed to use generalized linear model (GLM). Their approach combined logit transformation (transform data into (0,1)) and binomial distribution (dealing with the value 0 and 1)together. Thus, if the data are usual proportion data without the second feature mentioned in the following, we can use this method. The second feature is that, as we mentioned in Tobit part, firms with zero product innovation actually comes from at least two sources because of the product innovation definition: firms with product innovation but fail to meet the new product definition and defined as firms with zero product innovation, and firms with real zero new product introduced. We can still use Tobit model if we don't have the first feature. With the first feature, Tobit model, is a conceptually flawed model for proportion model (Cook et al., 2008). Thus, given the two features of our dependent variable, we will use zero-inflated beta model as it is proposed in Cook et al.(2008) to deal with proportional data with a mass at zero. They define the zero-inflated beta probability density as

$$g(y;\theta) = \begin{cases} 0 & \text{if } y < 0 \\ \delta & \text{if } y = 0 \\ (1-\delta) f(\mathbf{X};\theta) & \text{if } 0 < y < 1 \end{cases} \tag{36}$$

[6]This is based on Baum(2008)

[7]Moreover, proportion data has the nature of heteroskedastic nature and its conditional variance must be a function of the conditional mean, a weighted least square method is developed to deal with it based on logit transformation. See more in Baum (2008).

where

$$f(y;p,q) = [\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1}] \text{ for } 0 < y < 1. \tag{37}$$

Thus following Cragg (1971) and Cook et al.(2008), zero-inflated beta regression can be formulated as

$$f(y_i = 0|\mathbf{X}_i) = 1 - C(\alpha'\mathbf{X}_i) \text{ for } y_i = 0 \tag{38}$$

and

$$f(y_i|\mathbf{X}_i) = C(\alpha'\mathbf{X}_i)[\frac{\Gamma(p+q(\mathbf{X}_i))}{\Gamma(p)\Gamma(q(\mathbf{X}_i))} y^{p-1}(1-y_i)^{q(\mathbf{X}_i)-1}] \text{ for } 0 < y_i < 1, \tag{39}$$

where $q(\mathbf{X}_i) = p\exp(-\beta'\mathbf{X}_i)$ and p are parameters for the beta distribution. $C(\alpha'\mathbf{X}_i)$ represents the probability of a firm choosing to have product innovation. α stands for marginal effect on the decision to do product innovation or not, while β stands for the effect on how much to do product innovation for firms with non-zero product innovation. Note that the fact that α and β are allowed to be different indicates that the two effects mentioned above are allowed to be different. This regression is fitted using maximum likelihood estimation.

4.5 Results

We present the results from new product sales proportion regression in Table 11. We use zero-inflated beta model in this regression.[8] Column(1) presents the coefficients from ZIbeta, zero-inflated beta model; column (2) presents the coefficients from zero-inflated part of zero-inflated beta model; column (3) presents the marginal effects of independent variables on observed dependent variable, new product sales proportion,

[8] As we mentioned before, generalized linear model (GLM) proposed by Pakes and Wooldridge (1996) and popular in proportion data regression might be appropriate here. Also, Tobit, which is often used to deal with censored data is might also be appropriate here because our dependent variable with zero value might come from more than one sources. Therefore, in the model selection,I also tried GLM and Tobit. But when we compared actual value and predicted value, we find that zero-inflated beta model fits better.

given uncensored and the marginal effects are calculated at mean; column (4) presents the marginal effects of independent variables on probability of being uncensored, that is, of being a firm with non-zero new products introduced. Columns (5)-(8) are similar but for different dataset, Data 2002. New product sales proportion is used as dependent variable. City and industry dummied are controlled.

We find that in Data 2000, out of all the human capital indicators, we only find a positive effect of the average schooling of management team on the probability of being a new product introducer, or the negative effect of being a firm with zero new product introduced. The marginal effect is 0.0460 in Data 2000, which means that other things equal, for an average firm in the sample, when the average schooling of management team increase 1 year, the probability of a firm being a new product introducer will increase 0.046. In Data 2002, the effect is comparable, an extra year of schooling of management team will increase the probability of being a new product introducer in Data 2002 is 0.0326, only slight lower than in Data 2000. Notably, general manager's postgraduate degree has a positive and significant effect both statistically and economically on the probability of being a new product introducer in Data 2002. More specifically, other things equal, the probability of being a new product producer will increase 0.209 if an average firm with a general manager holing a postgraduate degree than not. We also find a negative effect of the age of management team on the probability of being a product innovator. When management team's age increases one year, the probability of a firm being a product innovator will decrease 0.00848, not a large effect. Also, we fail to find any significant effect of human capital on the amount of sales proportion change given only product innovator.

Notably, we find a positive effect of R&D on new product sales proportion both on the decision of being a product innovator (the probability of non-zero new product introducer) and the amount of new product sales proportion given non-zero product innovator and both effects are comparable across the two datasets. Specifically, other

Table 11: New Product Proportion Regression (Dependent Variable: New Product Sales Proportion)

	Data 2000				Data 2002			
	ZIbeta	ZIbeta(=0)	ME on $E(y\|y>0)$	ME on $P(y>0\|x)$	ZIbeta	ZIbeta(=0)	ME on $E(y\|y>0)$	ME on $P(y>0\|x)$
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of highly educated workers(Hundred)	0.00734	-0.0777	0.00162	0.0194	0.0199	-0.0804	0.00438	0.0197
	(0.0151)	(0.0518)	(0.00333)	(0.0129)	(0.0160)	(0.0517)	(0.00352)	(0.0126)
General Manager's tenure (years)	0.0108	0.0392	0.00238	-0.00978	0.0122	-0.00114	0.00269	0.000279
	(0.0160)	(0.0246)	(0.00355)	(0.00613)	(0.0106)	(0.0202)	(0.00234)	(0.00495)
General Manager's postgraduate degree	0.0458	0.156	0.0101	-0.0390	-0.00233	-0.854***	-0.000513	0.209***
dummy (=1 if has a postgraduate degree)	(0.158)	(0.321)	(0.0349)	(0.0801)	(0.113)	(0.257)	(0.0248)	(0.0625)
Management team's average age	-0.000701	0.00539	-0.000155	-0.00135	-0.0105	0.0346*	-0.00232	-0.00848*
	(0.00990)	(0.0187)	(0.00219)	(0.00468)	(0.0115)	(0.0185)	(0.00253)	(0.00453)
Management team's average schooling	0.0237	-0.184**	0.00524	0.0460**	-0.00334	-0.133**	-0.000735	0.0326**
	(0.0492)	(0.0756)	(0.0109)	(0.0189)	(0.0404)	(0.0618)	(0.00889)	(0.0151)
Panel B: R&D and Firm Characteristics								
Log(average R&D in previous two years)	0.0309***	-0.110***	0.00683***	0.0275***	0.0144*	-0.0857***	0.00316*	0.0210***
	(0.0105)	(0.0172)	(0.00228)	(0.00429)	(0.00805)	(0.0135)	(0.00176)	(0.00330)
Number of competitors	-0.103**	0.0971	-0.0228**	-0.0242	-0.0651*	0.111*	-0.0143*	-0.0272*
	(0.0488)	(0.0786)	(0.0107)	(0.0196)	(0.0365)	(0.0669)	(0.00806)	(0.0164)
Firm size(log(net value of total assets))	0.0123	-0.0288	0.00271	0.00719	-0.0309	0.0770	-0.00682	-0.0189
	(0.0414)	(0.0730)	(0.00915)	(0.0182)	(0.0322)	(0.0667)	(0.00709)	(0.0163)
Firm age (year)	0.00172	-0.00768	0.000381	0.00192	0.00238	0.00229	0.000525	-0.000561
	(0.00461)	(0.00782)	(0.00102)	(0.00195)	(0.00396)	(0.00736)	(0.000871)	(0.00180)
Shareholding firms dummy	0.325	0.266	0.0718	-0.0665	0.127	-0.254	0.0280	0.0623
	(0.205)	(0.328)	(0.0453)	(0.0819)	(0.141)	(0.246)	(0.0309)	(0.0603)
State-owned firms dummy	0.103	0.379	0.0229	-0.0947	0.115	-0.353	0.0253	0.0865
	(0.195)	(0.311)	(0.0432)	(0.0776)	(0.163)	(0.263)	(0.0360)	(0.0643)
Foreign invested firms dummy	0.225	0.376	0.0498	-0.0938	0.0213	0.109	0.00469	-0.0267
	(0.186)	(0.319)	(0.0414)	(0.0796)	(0.172)	(0.267)	(0.0378)	(0.0654)
Constant	-1.126	2.344			-0.00496	-0.0683		
	(0.892)	(1.452)			(0.757)	(1.309)		
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	537	537	537	537	813	813	813	813

Standard errors in parentheses: * p <0.10, ** p <0.05, *** p <0.01. Marginal effect for dummy variable is from 0 to 1.

Column(1) presents the coefficients from ZIbeta, zero-inflated beta model; column (2) presents the coefficients from zero-inflated part of zero-inflated beta model; column (3) presents the marginal effects of independent variables on observed dependent variable, new product sales proportion, given uncensored and the marginal effects are calculated at mean; column (4) presents the marginal effects of independent variables on probability of being uncensored, that is, of being a firm with non-zero new products introduced. Columns (5)-(8) are similar but for different dataset, Data 2002.

things equal, for an average firm with nonzero product innovation, if its R&D increase 1 percent, its number of product innovation will increase 0.00683 in Data 2000 and it will increase 0.00316 in Data 2002. Other things equal, for an average firm, the probability of being a product innovator will increase 0.0275 and 0.0210 respectively in Data 2000 and Data 2002 with 1 percent increase in R&D.

Moreover, we find that number of competitors have a negative effect on product innovation and the effects a little bit across the two datasets. In Data 2000, we only find a negative effect on the amount of product sales proportion for product innovators but there's no significant effect on the probability of being an product innovator. In data 2002, we find both a negative effect on the amount and the probability. More specifically, in Data 2000, we find that other things equal, for an average product innovator with non-zero product innovation, when the number of competitors it faces increase 1, its product innovation proportion will decrease 0.0228. In Data 2002, when number of competitors increases 1, other things equal, the new product sales proportion of an average innovator with non-zero new product introduced will decrease 0.0143; for an average firm, when the number of competitors increases 1, the probability of being a product innovator will decrease 0.0272.

We can conclude from Table 11 that for a firm, management team's education will promote the probability of being a product innovator for both more developed areas and less developed areas. General Manager's postgraduate degree has a large effect on the probability of being a product innovator in less developed areas but no significant effect in more developed areas. The reason might be that in less developed areas, since the market is less developed, a general manager with higher education is more important than in a more developed market. We also find that the management team's age has a negative effect on the probability of being a product innovator in less developed area but no significant effect in more developed area. Moreover, we find that R&D has a positive effect on both the amount of new product sales proportion for product innovators and the probability of being a product innovator. The number of competitors has a negative effect on the amount of new product sales proportion for product innovator in both areas, but it only has decisive effect of being a product innovator in less developed areas.

In Table 12, we present new product sales regression. Column(1) the coefficients of Tobit; column (2) presents the marginal effects of independent variables on observed dependent variable, new product sales proportion, given uncensored and the marginal effects are calculated at mean; column (3) presents the marginal effects of independent variables on probability of being uncensored, that is, of being a firm with non-zero new products introduced. Columns (4)-(6) are similar but for different dataset, Data 2002. Log of new product sales is used as dependent variable. City dummies and industry dummies are controlled.

The first important result is that number of educated workers has a positive effect on both the amount for the uncensored firms but also on the probability of being a non-zero product innovator for both datasets even we have controlled firm size by using log of net value of total assets. The two marginal effects are both comparable across both datasets and this indicates that our result are robust. More specifically, its marginal effect on the amount of new product sales for product innovator is 0.0922 and 0.120 respectively in Data 2000 and in Data 2002, indicating that when number of highly educated workers increases 100, the new product sales will increase 0.0922 percent and 0.12 percent respectively in Data 2000 and data 2002; its marginal effect on the probability of being uncensored or being a product innovator, is 0.0114 and 0.0137 respectively in Data 2000 and Data 2002, indicating that when number of highly educated workers increases 100, the probability of being a product innovator increases 0.0114 and 0.0137 respectively. The effects of number of highly educated workers, or skilled human capital make sense in that skilled human capital is the base of firm innovation. On the one hand, even without R&D, skilled human capital can

Table 12: New Products Sales Regression (Dependent Variable: log(New Product Sales)

	Year 2000(Data 2000)			Year 2002(Data 2002)		
	Tobit	ME on $E(y\|y>0)$	ME on $P(y>0\|x)$	Tobit	ME on $E(y\|y>0)$	ME on $P(y>0\|x)$
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)	(5)	(6)
Number of highly educated workers (Hundred)	0.214**	0.0922**	0.0114**	0.256***	0.120***	0.0137***
	(0.0902)	(0.0389)	(0.00479)	(0.0824)	(0.0386)	(0.00442)
General Manager's tenure (years)	-0.138	-0.0597	-0.00740	0.0338	0.0159	0.00181
	(0.0877)	(0.0378)	(0.00471)	(0.0630)	(0.0296)	(0.00336)
General Manager's postgraduate degree	-0.171	-0.0737	-0.00912	2.203***	1.034***	0.118***
dummy (=1 if has a postgraduate degree)	(1.025)	(0.442)	(0.0548)	(0.634)	(0.296)	(0.0340)
Management team's average age	-0.0318	-0.0137	-0.00170	-0.146**	-0.0687**	-0.00782**
	(0.0648)	(0.0280)	(0.00346)	(0.0612)	(0.0287)	(0.00329)
Management team's average schooling	0.674***	0.291***	0.0360***	0.470**	0.221**	0.0251**
	(0.256)	(0.110)	(0.0138)	(0.204)	(0.0960)	(0.0109)
Panel B: R&D and Firm Characteristics						
Log(average R&D in previous two years)	0.446***	0.193***	0.0239***	0.311***	0.146***	0.0166***
	(0.0630)	(0.0270)	(0.00363)	(0.0456)	(0.0213)	(0.00255)
Number of competitors	-0.342	-0.147	-0.0183	-0.469**	-0.220**	-0.0251**
	(0.278)	(0.120)	(0.0149)	(0.214)	(0.101)	(0.0115)
Firm size(log(net value of total assets))	0.440*	0.190*	0.0235*	0.0935	0.0439	0.00500
	(0.245)	(0.106)	(0.0131)	(0.205)	(0.0962)	(0.0109)
Firm age (year)	0.0244	0.0105	0.00130	-0.0121	-0.00566	-0.000645
	(0.0261)	(0.0113)	(0.00140)	(0.0231)	(0.0108)	(0.00123)
Shareholding firms dummy	-0.614	-0.265	-0.0328	0.889	0.417	0.0475
	(1.173)	(0.505)	(0.0627)	(0.806)	(0.378)	(0.0431)
State-owned firms dummy	-1.371	-0.591	-0.0733	0.752	0.353	0.0402
	(1.074)	(0.462)	(0.0575)	(0.849)	(0.398)	(0.0454)
Foreign invested firms dummy	-0.754	-0.326	-0.0403	-0.558	-0.262	-0.0298
	(1.116)	(0.480)	(0.0598)	(0.881)	(0.414)	(0.0471)
Constant	-10.37**			-0.320		
	(5.037)			(4.349)		
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
City dummies	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	537	537	537	813	813	813

Standard errors in parentheses: * p <0.10, ** p <0.05, *** p <0.01. Marginal effect for dummy variable is from 0 to 1.
Column(1) the coefficients of Tobit; column (2) presents the marginal effects of independent variables on observed dependent variable, new product sales proportion, given uncensored and the marginal effects are calculated at mean; column (3) presents the marginal effects of independent variables on probability of being uncensored, that is, of being a firm with non-zero new products introduced. Columns (4)-(6) are similar but for different dataset, Data 2002.

produce product innovation via "learning from doing", communication with suppliers and consumers and so on. On the other hand, skilled human capital is complementary to R&D activities and be supportive to product innovation.

Moreover, we find that the average years of schooling has a positive effect on both the amount of new product sales for product innovator and the probability of being a product innovator in both datasets and the effects are comparable in the two datasets. More specifically, its marginal effect on the amount of new product sales for product innovator is 0.291 and 0.221 respectively in Data 2000 and Data 2002, indicating that other things equal, for an average product innovator, when management team's average schooling increases 1 year, new product sales will increase 0.291 percent and 0.221 percent, and this is very significant both statistically and economically; its marginal effects on probability of being a product innovator is 0.0360 and 0.0251 respectively in Data 2000 and Data 2002, indicating that other things equal, for an average firm, when its management team's years of schooling increases 1 year, its probability of being a product innovator will increase 0.0360 and 0.0251 respectively in Data 2000 and Data 2002.

Moreover, as in new product sales proportion regression, we only find the significant effects of general manager's postgraduate degree and management team's average age on both the amount of new product sales for product innovator and the probability of being a product innovator in Data 2002. Specifically, the two marginal effects of general manager's postgraduate is 1.034 and 0.118 respectively, indicating that other things equal, when an average firm with a general manager holding a postgraduate degree than not, the product innovator's new product sales will increase 1.034 percent, while the probability of being a product innovator for an average firm will increase 0.118. The two marginal effects os management team's age is -0.0687 and -0.00782 respectively, indicating that other things equal, when the average age of management team in an average firm increases 1 year, the product innovator's new product sales will decrease 0.0687 percent while the probability of being a product innovator for an average firm will decrease 0.00782.

R&D has a significant effect on both the amount of new product sales for product innovator and the probability of being a product innovator across both datasets. The two effects across the two datasets are comparable though they are a little bit larger in Data 2000. More specifically, the marginal effect of R&D on the amount of new product sales for product innovators is 0.193 and 0.146 respectively, indicating that other things equal, for an average product innovator, when R&D increase 1%, its new product sales will increase 0.193 % and 0.146% respectively in Data 2000 and Data 2002; its marginal effect on the probability of being a product innovator is 0.0239 and 0.0166 respectively, indicating that other things equal, for an average firm, when its R&D increase 1%, its probability of being a product innovator will increase 0.0239 and 0.0166 respectively in Data 2000 and Data 2002.

Number of competitors only has significant effects in Data 2002. Other things equal, when number of competitors increases 1, the amount of new product sales will decrease 0.22 percent for a product innovator, and for an average firm, its probability of being a product innovator will decrease 0.0251. We only find a significant effect of firm size, measured by log of net value of total assets, in Data 2000. Other things equal, when net value of total assets increases 1%, for an average product innovator, its new product sales will increase 0.19%, and for an average firm, its probability of being a product innovator will increase 0.235.

To sum up the results in Table 12, number of skilled human capital, management team's average schooling and R&D aa have positive effects on both amount of new product sales for product innovator and on the probability of being a product innovator. General manager's postgraduate degree have positive and significant effects on the amount of new product sales for product innovator and the probability of being a product innovator only in Data 2003. Management team's average age and number

Table 13: New Products Counts Regression (Dependent Variable: Number of New Products)

Panel A: Human Capital Variables	Year 2000(Data 2000) NB (1)	Marginal Effect (2)	Year 2002(Data 2002) NB (3)	Marginal Effect (4)
Number of highly educated workers(Hundred)	0.0996**	0.459**	0.0978**	0.623**
	(0.0390)	(0.201)	(0.0388)	(0.254)
General Manager's tenure (years)	0.0686**	0.317**	0.0288*	0.183*
	(0.0281)	(0.146)	(0.0165)	(0.107)
General Manager's postgraduate degree	0.0653	0.301	0.574***	3.653***
dummy (=1 if has a postgraduate degree)	(0.327)	(1.516)	(0.182)	(1.168)
Management team's average age	0.0139	0.0640	-0.0234	-0.149
	(0.0186)	(0.0881)	(0.0179)	(0.114)
Management team's average schooling	-0.209**	-0.964**	0.192***	1.225***
	(0.0843)	(0.442)	(0.0589)	(0.394)
Panel B: R&D and Firm Characteristics				
Log(average R&D in previous two years)	0.0652***	0.301***	0.0779***	0.496***
	(0.0165)	(0.0744)	(0.0126)	(0.0840)
Number of competitors	-0.00948	-0.0437	0.00981	0.0625
	(0.0903)	(0.416)	(0.0612)	(0.390)
Firm size(log(net value of total assets))	0.0174	0.0801	-0.0235	-0.150
	(0.0640)	(0.295)	(0.0554)	(0.352)
Firm age (year)	-0.00338	-0.0156	0.00534	0.0340
	(0.00669)	(0.0311)	(0.00658)	(0.0418)
Shareholding firms dummy	1.048***	4.836***	0.232	1.476
	(0.346)	(1.846)	(0.226)	(1.461)
State-owned firms dummy	0.568*	2.619*	0.251	1.601
	(0.309)	(1.516)	(0.243)	(1.568)
Foreign invested firms dummy	1.439***	6.640***	-0.372	-2.367
	(0.335)	(2.063)	(0.227)	(1.450)
Constant	1.749		0.0187	
	(1.373)		(1.228)	
Industry dummies	Yes	Yes	Yes	Yes
City dummies	Yes	Yes	Yes	Yes
lnalpha	1.377***		1.203***	
	(0.0892)		(0.0650)	
Number of observations	551	551	824	824

Standard errors in parentheses: * p <0.10, ** p <0.05, *** p <0.01. Marginal effect for dummy variable is from 0 to 1. Column (1) Presents the coefficient from negative binomial regression (NB). Column (2) presents the marginal effect of NB. Marginal effects are calculated at mean. Columns (3) and (4) are similar but for different dataset, Data 2002.

of competitors have negative and significant effects on the amount of new product sales for product innovator and the probability of being a product innovator only in Data 2003. We find that firm size has the two significant and positive effects only in Data 2000.

Table 13 presents the results of new product counts regression. Column (1)

Presents the coefficient from negative binomial regression (NB). Column (2) presents the marginal effect of NB. Marginal effects are calculated at mean. Columns (3) and (4) are similar but for different dataset, Data 2002. We also tried a zero-inflated negative binomial model and it has similar power of fitting, measured by the difference between fitted value and actual value. Therefore, we only present negative binomial and its marginal effect here. The dependent variable is number of new products. City dummies and industry dummies are control for fixed effect.

Highly educated workers has a positive effect on new product counts in both datasets and the effects are comparable, indicating that our results are robust. Specifically, other things equal, for an average firm, when number of highly educated workers increases 100, the number of new products will increase 0.459 in Data 2000 and 0.623 in Data 2002. Also, this implies that the effect of skilled human capital on new product counts are larger in small cities, less developed areas. General manager's tenure has a positive effect on number of new products introduced. When general manager's tenure increases 1 year, other things equal, for an average firm, its new products introduced will increase 0.317 in Data 2000 and 0.183 in Data 2002. We can see that general manger's tenure has a larger effect in Data 2000, more developed areas. Consistent with results using two other product innovation measures, general manager's postgraduate has a positive and significant effect both statistically and economically only in Data 2002. Specifically, other things equal, for an average firm, when its general manager has a postgraduate than not, its number of new products will increase 3.653, a very large effect.

It is weird to find that management team's average schooling in Data 2000 has a negative and significant effect. Since we use the same set of explanatory as in Table 11 and Table 12, it is very likely that the problem is caused by the dependent variable, the number of new products. It might not be very appropriate to serve as a product innovation measure since the value of new products can vary a lot. The same amount of new products can be introduced by a firm with management team with high education level,or a firm with management team with low education level. But the value of new products introduced by the first firm might be much higher than that by the second firm. The idea is that value matters a lot. That's also why all the coefficients make sense in Table 11 and Table 12 where proportion measure (based on value measure) and value measure are used. To further investigate the relationship between number of new products introduced and average years of schooling, we plot their relationship in these two datasets in Figure 10. We can see that the relationship between number of new products introduced and management team's average years of schooling is almost concave in Data 2000 but it is monotonically increasing in Data 2002. Thus, the reason why we get negative sign in Data 2000 is that the relationship between number of new products introduced and management team's average years of schooling is not monotonic. In fact, when we add a squared term of management team's average years of schooling into the regression, we find that the coefficient of management team's average years of schooling is positive and the coefficient of its squared term is negative. Though both terms are insignificant but they are jointly significant.

Similarly, we also find a positive and significant effect of R&D on new product counts. More specifically, other things equal, for an average firm, when R&D increase 1 percent, the number of new products will increase 0.301 in Data 2000 and 0.496 in Data 2002. Moreover, we find that positive fixed effects of shareholding firms, state-owned firms and foreign invested firms in Data 2000. In addition, the fixed effect of foreign firms is largest, and then followed by shareholding firms and the fixed effect of state-owned firms have the least fixed effects. Other things equal, for an average firm, when it is foreign invested firm, its new product count will be 6.64 more than a private domestic firm; when it is shareholding firm, its new product count will be 4.836 more than a private domestic firm; when it is a state owned firm, its new

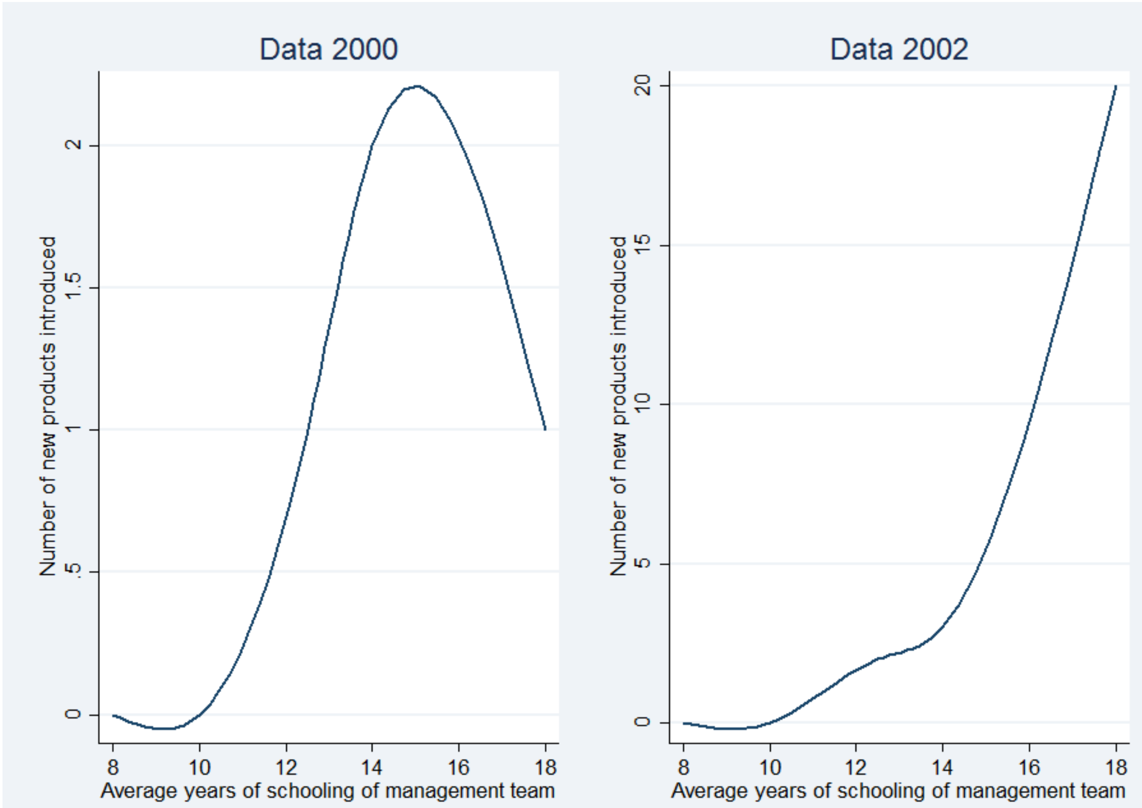

Figure 10: Relationship between Number of New Products Introduced and Management Team's Education.

product count will be 2.619 more than a domestic private firm. This is consistent with Hu (2001).

Table 14 presents the results of IV estimation. Columns (1)-(4)present the coefficients from zero-inflated beta model (ZIbeta), the zero-inflated part of ZIbeta, Tobit model and negative binomial (NB) model respectively. Columns (5)-(8) are similar but for different dataset, Data 2002. Endogenous variables are number of highly educated workers, and general manager's postgraduate degree dummy. Instruments for number of highly educated workers are number of applicants for the skilled positions, vacant weeks of skilled position before filled, city average and industry average both excluding firm itself. Note that for Data 2002, we have no information on applicants and vacant weeks, so we only use city average and industry average excluding firm itself. For general manager's postgraduate degree dummy, we use city

Table 14: IV Estimation Results

	Year 2000(Data 2000)				Year 2002(Data 2002)			
	ZIbeta	ZIbeta(=0)	Tobit	NB	ZIbeta	ZIbeta(=0)	Tobit	NB
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of highly educated workers(Hundred)	0.0182	-0.0893	0.215*	0.0714*	0.0279*	-0.0925**	0.374***	0.123***
	(0.0199)	(0.0716)	(0.115)	(0.0425)	(0.0149)	(0.0431)	(0.0915)	(0.0390)
General Manager's tenure (years)	-0.00273	0.0402	-0.152	0.0241	0.0142	-0.00152	0.0394	0.0302*
	(0.0178)	(0.0338)	(0.110)	(0.0306)	(0.0108)	(0.0209)	(0.0641)	(0.0165)
General Manager's postgraduate degree	-0.00355	0.338	-0.833	-0.596**	-0.0231	-0.920***	2.445***	0.629***
dummy (=1 if has a postgraduate degree)	(0.204)	(0.353)	(1.168)	(0.289)	(0.113)	(0.263)	(0.656)	(0.181)
Management team's average age	-0.00150	-0.00413	0.00254	-0.0194	-0.0125	0.0386**	-0.157**	-0.0242
	(0.0139)	(0.0238)	(0.0827)	(0.0205)	(0.0114)	(0.0188)	(0.0615)	(0.0177)
Management team's average schooling	0.0424	-0.195**	0.706**	-0.0604	-0.000631	-0.131**	0.467**	0.208***
	(0.0567)	(0.0993)	(0.322)	(0.0857)	(0.0409)	(0.0626)	(0.204)	(0.0568)
Panel B: R&D and Firm Characteristics								
Log(average R&D in previous two years)	0.0298**	-0.0773***	0.313***	0.0596***	0.0144*	-0.0864***	0.305***	0.0781***
	(0.0127)	(0.0208)	(0.0755)	(0.0174)	(0.00803)	(0.0138)	(0.0454)	(0.0128)
Number of competitors	-0.0923	0.0561	-0.161	0.0795	-0.0661*	0.0945	-0.430**	0.0128
	(0.0644)	(0.106)	(0.367)	(0.0965)	(0.0365)	(0.0672)	(0.213)	(0.0609)
Firm size(log(net value of total assets))	0.0127	-0.0742	0.623**	0.111	-0.0376	0.0959	-0.0225	-0.0583
	(0.0551)	(0.0896)	(0.291)	(0.0745)	(0.0328)	(0.0677)	(0.209)	(0.0575)
Firm age (year)	0.00801	-0.00755	0.0302	0.00377	0.00259	0.00363	-0.0146	0.00475
	(0.00631)	(0.00990)	(0.0311)	(0.00801)	(0.00399)	(0.00747)	(0.0232)	(0.00654)
Shareholding firms dummy	0.333	0.0164	0.176	0.873**	0.115	-0.241	0.774	0.213
	(0.247)	(0.410)	(1.415)	(0.384)	(0.140)	(0.251)	(0.806)	(0.228)
State-owned firms dummy	-0.311	0.495	-1.930	0.249	0.134	-0.319	0.632	0.233
	(0.259)	(0.413)	(1.372)	(0.369)	(0.165)	(0.267)	(0.846)	(0.246)
Foreign invested firms dummy	0.452**	0.416	-0.454	0.712**	0.0128	0.151	-0.711	-0.378*
	(0.211)	(0.399)	(1.357)	(0.349)	(0.173)	(0.272)	(0.881)	(0.230)
Residual1	-0.0555	-0.0812	0.0991	0.134	-0.0443	0.206	-0.752***	-0.205*
	(0.198)	(0.407)	(0.998)	(0.280)	(0.0321)	(0.143)	(0.277)	(0.121)
Residual2	0.951	-1.073	-0.431	-1.327	0.296	-0.210	-0.147	-0.327
	(1.661)	(2.464)	(9.066)	(1.779)	(0.309)	(0.773)	(1.806)	(0.471)
Constant	-1.206	3.246*	-13.61**	0.243	0.0577	-0.438	1.146	-0.0922
	(1.163)	(1.937)	(6.532)	(1.577)	(0.760)	(1.359)	(4.432)	(1.185)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
lnalpha				1.098***				1.192***
				(0.0993)				(0.0649)
Number of observations	339	339	339	347	799	799	799	810

Standard errors in parentheses: * p <0.10, ** p <0.05, *** p <0.01. Marginal effect for dummy variable is from 0 to 1. Columns (1)-(4)present the coefficients from zero-inflated beta model (ZIbeta), the zero-inflated part of ZIbeta, Tobit model and negative binomial (NB) model respectively. Columns (5)-(8) are similar but for different dataset, Data 2002. Endogenous variables are number of highly educated workers, and general manager's postgraduate degree dummy. Instruments for number of highly educated workers are number of applicants for the skilled positions, vacant weeks of skilled position before filled, city average and industry average both excluding firm itself. Note that for Data 2002, we have no information on applicants and vacant weeks, so we only use city average and industry average excluding firm itself. For general manager's postgraduate degree dummy, we use city and industry average excluding firm itself as instruments. For all the models, we use control function technique. Residual1 is the residual from the first stage of number of highly educated workers, and residual2 is the residual from the first stage of general manager's postgraduate degree dummy.

and industry average excluding firm itself as instruments. For all the models, we use control function technique. Residual1 is the residual from the first stage of number of highly educated workers, and residual2 is the residual from the first stage of general manager's postgraduate degree dummy. We can see that in columns (1)- (6), none of the residuals are significant, indicating that we cannot reject that the suspected endogenous variables, number of highly educated workers and general manager's postgraduate degree dummy, are exogenous. Thus, for them, the original model still hold. For columns (7) and (8), residual1, the residual from first stage of number of highly educated workers, is significant, indicating that we can reject that the suspected endogenous variable is exogenous. Compared columns (7) and (8) with their original model, we find that the main results still holds and the effect of human capital indicators even has slightly larger effects. Therefore, even after we take endogeneity into account, our results still hold.

4.6 Conclusion

In this chapter, we examine how product innovation is determined in a firm from a human capital point of view. Three product innovation measures are used, new product sales proportion, new product sales and new product count. For skilled human capital, we find that it also tends to have a positive effect across the three measures, though it is not significant in regression using new product proportion as product innovation measure. When new product sales proportion and new product sales are used as product innovation measures, we find that management team's average years of schooling has a positive effect in determining product innovation, however, its effect in Data 2000 when new product count used as product innovation measure is negative and significant. This indicates that though the results from the three regressions are generally consistent, but they still differ. When compare the results using different innovation measures, we should be very cautious. Another thing is that different from results in last chapter where R&D has a positive effect only in Data 2003, less developed areas, we find that R&D has a very significant and positive effect on product innovation no matter which product innovation measure is used. This indicates that R&D is still important in promoting product innovation and policies supporting investment in R&D is still important. Our results still hold when endogeneity is considered. Notably, in all three models, we all find that general manager's postgraduate degree has a large and significant effect in less developed areas but insignificant at all in more developed areas.

Why postgraduate has a large and positive effect only in Data 2002, less developed areas across all product innovation measures? One possibility is that there are too less managers with postgraduate degree in less developed areas, but we can see from descriptive statistics that the percentage of managers holding a postgraduate degree is 16% in Data 2000 and 17% in Data 2002, even larger than in more developed areas. Thus, this cannot be the reason. Another possibility is that the market development in less developed areas is still low and in this situation, a general manager with a postgraduate degree might have more insight in making decision. The relationship between market environment and product innovation is discussed in Li and Atuahene-Gima (2001). They explored how environmental factors (competition, institutional support, and environmental turbulence) moderate the product innovation-performance relationship in new technology ventures in China. They argued that environment factors are important to product innovation since resource dependence theory suggested that managers interpret demands and dependencies in their environment prior to making strategic choices and instituting adjustments to organization strategies.

CHAPTER V

FIRM-LEVEL HUMAN CAPITAL AND PRODUCTIVITY: EVIDENCE FROM CHINA

5.1 Introduction

According to OECD, China was the second largest R&D spender in 2012, allocating 294 billion dollars compared to top-spending America at around 454 billion dollars that year. In fact, China is forecast to over take the European Union and the United States in R&D by 2019.At the same time, according to Li et al.(2009), China's total real human capital increased from 26.98 billion yuan in 1985 to 118.75 billion yuan in 2007 and the average annual growth rate is around 6.78%. Does this mean that China has finished the transition from Made in China to Innovated in China? Probably not. Despite success in some areas, notably high-speed rail, solar energy, supercomputing and space explorations, China is still far from an innovation country. For example, not a single Chinese company is on the list of 100 most innovative companies by Thomson Reuters until 2013 and there's only one Chinese company, Huawei, enters in the list in 2014 while there are around 40 percent from America. Why it looks that there's still very few innovation in China? One possible reason is that China do a lot of incremental innovation rather than radical innovation. If this is the case, then we should see a positive relationship between R&D, human capital and productivity.

This paper is an attempt to understand the effects of R&D spending and firm-level human capital on productivity and the most important determinants of productivity in Chinese manufacturing firms. We first present estimates of production function and TFP is the residual from production function. We then study the effects of R&D, skilled human capital, management characteristics, firm characteristics on TFP. We find that the elasticity of skilled human capital across both datasets is much larger than that of less skilled human capital. The number of skilled human capital and R&D have a positive and significant effect on productivity. Firm age has a significant and negative effect on TFP and the effects are almost the same across both datasets. Other things equal, a state-owned firms will have a lower TFP. Finally, we find a positive effect of foreign invested firms, indicating that foreign ownership means a higher productivity.

The study on the R&D —productivity link is prevalent. Most of those studies were based on the simplified specification in terms of total factor productivity and R&D. For example, in Griliches (1980) knowledge capital enters into production function in the same way as capital and labor do. Using a large sample of industrial firms, he finds that R&D elasticity amounts to 0.07 for manufacturing as a whole. Similarly, Hall and Mairesse (1995) use same production function but with corrections for R&D double counting, that is, the number of researcher is subtracted from the total number of employees and physical capital devoted to R&D laboratories from total physical capital, they found a much more higher R&D elasticity. However, this is a deficient method for innovation-productivity study since (1) innovation also involves other firm resources, mainly a firm's skilled human capital and (2) they ignore the non-R&D innovation which is usually important to firm productivity.

Theoretically, pioneer work by Nelson and Phelps (1966) argued that human capital can promote TFP growth by facilitating technology spillover. Romer (1990) first formally modeled how human capital can enhance productivity through technology innovation. Empirically, a lot of studies find that human capital have a positive and significant effect on TFP (Vandenbussche, Aghion, and Meghir, 2006).

Generally, there are three types of creative activities that do not require R&D but very important to firm productivity. First, Kim and Nelson (2000) found that many imitative activities, including reverse engineering, do not require R&D, and the imitation is mainly dependent on the firm's technical personnel and engineers. Second, firms can make minor modifications or incremental changes to products and processes, relying on engineering human capital. Moreover, Hansen and Serin (1997) noted that the innovation process in low-and medium-technology sectors is more related to adaptation and learning by doing, based on design and process optimization, rather than from R&D. Third, firms can combine existing knowledge in new ways, for example in industrial design and engineering projects (Grimpe and Sofka, 2009). Due to the large share of firms that innovate without performing R&D, we can conclude that studies that only focus on R&D without including firm-level human capital should not be enough to fully explain innovation differences across firms.

Moreover, we can regard R&D capital as a specific form of human capital associated with innovation (Benhabib and Spiegel,1994). In fact, research workers' salaries constitute a sizable percentage of total expenditures. Thus, in a firm to get productivity improved, R&D-oriented workers needs to work with other skilled workers. Like R&D, a high level of human capital affects the ability of firms to learn and absorb new information, and also allows tangible inputs to be used more effectively.

In addition, management team can have important effects on firm productivity. Penrose (2009) argues that although markets set prices that influence resources allocation, those within the firms make decisions on what activities the firm will be involved in, how those activities will be performed, what resources are required, which resources are allocated to different activities and, ultimately, which resources are used. As a consequence, internal processes and insights rather than external market prices and cost signals will greatly influence a firm's growth (Darroch, 2005). Though there's a lot of studies on management and innovation (Lin et al., 2011), there's quite few studies on management and productivity. Among them, Singell (1993) links managerial skills and firm production directly by examining how the experience of major league baseball managers affects both team and individual player performance. He shows that with team skills held constant, the probability of winning depends on the baseball-specific human capital of the manager.

A constant problem for human capital study is how to measure it. Human capital of a firm is defined as the knowledge and skills of its professionals that can be used to produce professional services. In Romer model , skilled workers or human capital is explicitly included in the model. Pennings, Lee and Witteloostuijn(1998) use firm tenure, industry experience, and graduate education to capture firm-level human capital. Ballot, Fakhfakh and Taymaz (2001) constructs measures of a firm's human capital stock based on firms' past and present training expenditures to study its effect on productivity.

In this paper, we use the number of highly educated workers (college or above, mainly technical and engineering workers and managerial personnel) to measure firm-level skilled human capital. We use the number of less educated workers (high school or below, mainly production workers and auxiliary production workers) to measure less skilled human capital. We use General Manager's graduate degree dummy, tenure, management team's average age and schooling to study the effect of managerial human capital in a firm on productivity. In production function, skilled human capital and less skilled human capital can enter directly. Moreover, we further study how skilled human capital, R&D and managerial human capital influence TFP.

The chapter is organized as follows. Section 5.2 presents empirical strategy. Section 5.3 introduces the data. In Section 5.4, we present our main results and interpret the findings. Section 5.5 concludes.

5.2 Methodology

Total factor productivity (TFP) as an alternative to study output, has attracted more and more attention (e.g., Romer, 1993; Prescott, 1998). The origins of total factor productivity (TFP) can be traced back to Solow (1957) in aggregate study.

Nowadays, the availability of firm-level data provides researchers an opportunity to use TFP to study firm-level productivity. In those studies (for example, Griliches and Regev,1995), output, usually measured as deflated sales or value added, to be a function of the inputs the firm employs and its productivity. The measure of TFP is then obtained as the residual in this functional relationship. The production function is usually assumed of Cobb-Douglas form

$$Y_{it} = A_{it} K_{it}^{\beta_k} L_{it}^{\beta_l} \tag{40}$$

where Y_{it} if firm is output at time t; Output is typically measured as value added per year, deflated for price changes in time-series studies. It can also, however, be measured as physical units of output per year or gross value of output per year. Here, following Fleisher et al.(2011), we use value-added. A_{it} is its productivity; K_{it} is its capital input; L_{it} is its labor input; β_k and β_l are the elasticity of capital and labor respectively. Ideally, inputs should be measured in terms of *services* of of the input per unit of time, but such data are generally not available.

We can transform it into log form

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it} \tag{41}$$

where $\log A_{it} = \omega_{it} + \varepsilon_{it}$ known to firm, but ε_{it} is not and it is assumed to be random, and ω_{it} is TFP, which is usually known to firms themselves but not to economists. When firm has information about ω_{it} and choose inputs, K_{it} and L_{it}, then K_{it} and L_{it} are very likely to be correlated with ω_{it} This is called simultaneity (Mendershausen, 1938; Marschak and Andrews, 1944).

Studies which try to solve the simultaneity can be divided in four strands. First is IV method. We need to find instruments which are correlated with inputs K_{it} and L_{it} but uncorrelated with $\omega_{it} + \varepsilon_{it}$. Combining equation (40) with constraint $wL + rK = I$, where w is the wage for labor, r is interest rate and I is total budget, a profit-maximizing firm will get that

$$K = \frac{I\beta_k}{r(\beta_k + \beta_l)} \tag{42}$$

and

$$L = \frac{I\beta_l}{w(\beta_k + \beta_l)} \tag{43}$$

From the above two equations, we can see that each of the inputs of L and K is a function of input price, w and r. Thus, natural instruments for L and K should be input prices, which are likely to be uncorrelated with ω_{it}if input market is competitive. But in reality, they are usually unavailable.

Second, if we have panel data, we can consider fixed effects (FE). Suppose that

$$\omega_{it} = \eta_i + \delta_t + \omega_{it}^* \tag{44}$$

where ω_{it}^* is assumed to be uncorrelated with K_{it} and L_{it}, and is assumed to be not serially correlated and is strictly exogenous.

Plug equation (44) into equation (41),we then get that

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \eta_i + \delta_t + \omega_{it}^* + \varepsilon_{it} \tag{45}$$

In equation (45), η_i can be easily got rid of using fixed effects method. If we further assume that $E(\delta_t | k_{it}, l_{it}, \eta_i) = 0, t = 1, 2..., T$, that is, strict exogeneity, we then get that a consistent FE estimator. However, this is a too strong assumption.

The third strand is control function approach. Control function approach is very similar to usual two-stage least square. We need to find a proxy variable, different from instrument in IV method, proxy variable needs to be correlated with error term as closely as possible. The different is that we have to put the residual from the first stage into the second stage and then we regress on the residual from the first stage and all the other exogenous variables. The method proposed by Olley and Pakes (1996, OP thereafter) and then extended by Levinsohn and Petrin (2003, LP thereafter) is belong to this category. OP approach use investment as proxy variable. In our datasets, there are more than one third of the firms have zero investment and thus the investment proxy might not respond to the productivity shock very smoothly, violating the consistency condition (Petrin, Poi and Levinsohn, 2004). Thus, LP propose to use intermediate input or material, m_{it} as proxy. LP can solve the corner solution problem encountered by OP, but as Ackerberg, Caves and Frazer (2006, ACF) pointed out that OP/LP approach may suffer from identification problem due to the collinearity among inputs, and proposed an alternative estimator based on dynamic panel, which is our preferred approach. Since ACF is based on LP, we will begin our presentation from LP. The production function LP considered is

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \varepsilon_{it} \tag{46}$$

where m_{it} is an intermediate input such as electricity, fuel, or materials and ω_{it} is productivity, or TFP. When firms make its material input decision, it will depend on productivity level, ω_{it}, and state variable, k_{it}. It is chosen at the time production takes place, and ACF call it as "perfectly variable" input. Thus, LP assume the following intermediate input demand function

$$m_{it} = f_t(\omega_{it}, k_{it}) \tag{47}$$

In equation (47), f is indexed by t to indicate that input prices to vary across time but not across firm. Note that l_{it} is not included in equation (47), and thus labor is also a "perfectly variable" input, and it is chosen simultaneously with m_{it}, implying that the choice of l_{it} will have no impact on the choice of m_{it}.But ACF believe that labor is "less variable" than materials. Thus, a firm's material input demand at time t will now directly depend on the l_{it} chosen prior to it, and thus

$$m_{it} = f_t(\omega_{it}, k_{it}, l_{it}) \tag{48}$$

Inverting equation (48) for ω_{it} and substituting into the production function results in a first stage equation

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + f_t^{-1}(m_{it}, k_{it}, l_{it}) + \varepsilon_{it} \tag{49}$$

Clearly, β_l cannot be identified in the first stage, but we can obtain an estimate, $\hat{\Phi_{it}}$, of the composite term

$$\Phi_t(m_{it}, k_{it}, l_{it}) = \beta_k k_{it} + \beta_l l_{it} + f_t^{-1}(m_{it}, k_{it}, l_{it}) \tag{50}$$

Equation (50) represents output net of the untransmitted shock, ε_{it}. They also assume that ε_{it} evolves according to a first order Markov process between these subperiods $t-1$, $t-b$ where $0 < b < 1$, and t, i.e., $p(\omega_{it}|I_{it-b}) = p(\omega_{it}|\omega_{it-b})$ and $p(\omega_{it-b}|I_{it-1}) = p(\omega_{it-b}|\omega_{it-1})$, where I_{it-b} is investment at $t-b$ for firm i. Thus, we will have

$$\omega_{it} = E[\omega_{it}|I_{it}] + \xi_{it} = E[\omega_{it}|\omega_{it-1}] + \xi_{it} \tag{51}$$

where ξ_{it} is mean independent of all information known at $t-1$. Since k_{it} was decided at $t-1$, then the second stage moment condition is

$$E[\xi|k_{it}] = 0 \tag{52}$$

Moreover, l_{it} is also uncorrelated with ξ_{it}, thus the final second stage moment is

$$E[\xi_{it}| \begin{matrix} k_{it} \\ l_{it-1} \end{matrix}] = 0 \tag{53}$$

Thus,

$$E[\xi_{it} \cdot (\begin{matrix} k_{it} \\ l_{it-1} \end{matrix})] = 0 \tag{54}$$

This is the moments condition we need to estimate to get the coefficients of β_k and β_l.

To implement the whole procedure, we use STATA procedure proposed by Professor Jagadeesh Sivadasan. We set number of bootstrap replications to 400. The bootstrap samples do block bootstrapping, i.e., entire time series for each firm is sampled in blocks, which allows the errors to be correlated within plants. Amoeba optimization routine is used and the maximum number of iterative steps that should be done is set to be 500.

After we get the TFP from the production function estimation, we can start to study the determinants of TFP. We now specify the TFP determination equation as

$$\log(TFP_i) = \alpha_0 + \alpha_1 \log(HC_i) + \alpha_2 \log(RD_i) + \alpha_3 X_i + \varepsilon_i \tag{55}$$

Where TFP_i is firm is total factor productivity (TFP) which we get from the residual of the production function; HC_i is the human capital measure, here we use number of highly educated workers and managerial human capital; RD_i is firm's R&D input of last period. X_i is other firm characteristics.

We use the information on the number of skilled (highly educated) workers, whether General Manager has a postgraduate degree, General Manager's tenure in current firm, management team's average age, and the average years of schooling of management team as our human indicators. We expect that all the indicators have positive effects on productivity except General Manager's tenure and management team's average age.

Another important variable is R&D spending by the firm. We use the average of lagged one and lagged two periods R&D (we only have three years in our data). There are two rationales justify our R&D measure. First, R&D expenditure is a long-term investment. Thus, only including contemporaneous R&D cannot capture a firm's real innovation efforts. In this point of view, an average R&D over years rather than R&D of a certain year is a better innovation input measure for the firm. Second, contemporaneous R&D is very likely to be endogenous. That is, there is a possible correlation between unobserved innovation productivity shocks and R&D level. Thus, we exclude current R&D from the averages to lessen endogeneity. Though there are many studies on productivity—R&D relationship but there are few on the endogeneity of R&D. Among them, Jaffe (1986) treated firm R&D as endogenous and used corresponding industry average as an instrument. Following their method, we use both industry and city averages as instruments.

Firm size is measured by the log of total assets rather than the log of total sales to lessen the correlation between firm size and other variables. Intuitively, firms with large size might have economics of scale or/and scope. Generally, we expect a positive effect of firm size and when human capital is considered. We use two approaches to study the effect of market environment on innovation. First, we include market share of each firm in our model to account for a firm's market position. Second, we use two datasets, one from metropolitan cities and the other from provincial middle cities, to examine how productivity of firms in different markets, a more advanced one and a less advanced one differs.

5.3 Data Description

The datasets are still the same. Here, we directly present statistics summary for the full sample in Table 15 (Data 2000) and Table 16 (Data 2002). In Data 2000, the average firm has around 915-945 workers, of which around 180 skilled workers while in Data 2002, the mean value of total employment is around 735-745 workers of which 160 skilled workers. The average sales in Data 2000 is around 255-335 Million RMB while the mean of total sales in Data 2002 is around 200-270 Million RMB. The average total net assets in Data 2000 is 102.8 Million RMB while it is 96.52 in Data 2002. When it comes to the material cost, the mean in Data 2000 is around

110-140 Million RMB while it is around 115-150 Million RMB. From here we can see that firms in more developed area tends to be larger and have more skilled workers and have more revenue. However, the differences are not so large so that the two datasets are still comparable. We can see that there's very large difference in R&D between two datasets, with around 15-19 Million RMB in Data 2000 and around 2-4 Million RMB in Data 2002. In addition, there's little difference in General Manager's education and experience and the firm's age.

We use value added as the dependent variable and it is defined as total sale less total material costs (including raw materials, energy and other). We can see that the average of value added is larger in Data 2000 since cities in Data 2000 are metropolitan cities and are much more developed. Capital is defined as net value of assets in that year, including buildings, production machinery and equipment and other. Material is defined as total material cost, including raw materials, energy and other. Our human capital measures follow Fleisher et al.(2011). We use two measures, skilled human capital and less skilled human capital. Skilled human capital is defined as the number of skilled or highly educated workers. Skilled workers mainly consist of technical and engineering workers and managerial personnel. Less skilled human capital is defined as the number of less skilled workers. Less skilled workers mainly including basic production workers and auxiliary production workers. On average, there are around 180 skilled workers in more developed cities and around 160 skilled workers in less developed cities. We can see the difference is not so big. In both datasets, the number of less skilled workers is much larger, around twice, than that of skilled workers. To lessen endogeneity, we use lagged one period of R&D.

Finally, we divided all the firms into four ownership categories: shareholding companies, state-owned companies, foreign invested firms and other firms. We can see that there are 29% shareholding firms in Data 2002 while 16% shareholding firms in Data 2000. The deepening economic revolution might account for this change. There

Table 15: Descriptive Statistics for Productivity Study (Data 2000)

	Year	Observation #	Mean	Std. Dev.	Min	Max
Panel A: Outputs, Inputs and Human Capital Variables						
Value added (Million RMB)	2000	580	83.021	267.80	0.50	2967.24
	1999	579	67.49	213.37	0.47	3664.58
	1998	573	55.27	163.04	0.11	2723.20
Number of highly educated workers in firm	2000	582	181	316	10	4133
	1999	579	173	270	10	2496
	1998	582	178	267	10	2731
Number of less educated workers in firm	2000	582	535	786	11	7166
	1999	579	527	782	13	8298
	1998	582	559	851	10	10109
Capital (Million RMB)	2000	583	128.63	333.00	0.33	3539.07
	1999	583	121.20	309.59	0.32	3524.55
	1998	581	110.80	287.65	0.31	3411.28
Material Cost (Million RMB)	2000	580	159.62	482.09	0.12	4965.79
	1999	579	130.34	385.59	0.20	4398.96
	1998	573	112.10	349.45	0.38	4296.90
Value of total sales (Million RMB)	2000	583	246.82	707.07	3.00	7545.01
	1999	583	201.60	573.61	3.01	6439.07
	1998	579	169.38	495.95	3.04	6303.52
Total number of employees	2000	583	904	1403	50	17098
	1999	582	887	1385	50	18466
	1998	581	918	1433	50	18059
Years of schooling of General Manager (GM)	2000	581	14.01	2.29	5.00	18
Years of GM holding the position	2000	582	5.78	4.51	1.00	30
GM's postgraduate dummy (=1, postgraduate)	2000	581	0.15	0.36	0.00	1
Management team's average age	2000	573	36.28	6.63	18.00	54
Management team's average schooling	2000	574	11.86	1.49	8.00	18
Panel B: R&D and Firm Characteristics						
R&D expenditure by firm (Million RMB)	2000	563	9.64	55.95	0	830.29
	1999	570	7.64	57.26	0	1035.00
	1998	569	7.08	55.24	0	1041.21
Net value of total assets (Million RMB)	2000	582	81.17	201.27	0.20	2015.65
Firm age	2000	583	18.15	17.41	0.00	92
Shareholding firms dummy	2000	583	0.16	0.36	0.00	1
State-owed firms dummy	2000	583	0.24	0.43	0.00	1
Foreign invested firms dummy	2000	583	0.39	0.49	0.00	1

are 22% foreign invested firms in Data 2002 while there are 39% foreign invested firms in Data 2000. We can see that compared to middle-sized provincial cities, foreign investors prefer metropolitan cities in China.

5.4 Results

We first present production function estimation results in Table 17. From column (1)-(4), they are estimated using OLS, FE, LP and ACF respectively for Data 2000.

Table 16: Descriptive Statistics for Productivity Study (Data 2002)

	Year	Observation #	Mean	Std. Dev.	Min	Max
Panel A: Outputs, Inputs and Human Capital Variables						
Value added (Million RMB)	2002	798	97.99	313.80	0.57	4293.94
	2001	791	89.16	326.00	0.41	5811.52
	2000	788	76.18	237.78	0.40	2846.78
Number of highly educated workers in firm	2002	879	153	277.46	10	4281
	2001	877	154	300.35	10	5383
	2000	874	155	314.63	10	6086
Number of less educated workers in firm	2002	879	474	942.05	14	14434
	2001	877	474	937.26	11	12611
	2000	874	472	955.22	14	14559
Capital (Million RMB)	2002	880	121.08	339.12	0.38	3791.49
	2001	875	107.42	295.92	0.39	3307.45
	2000	870	99.17	272.82	0.30	3157.44
Material cost (Million RMB)	2002	798	135.95	498.46	0.022	5220.85
	2001	791	110.82	393.33	0.028	4546.09
	2000	788	99.07	359.63	0.021	4233.15
Value of total sales (Million RMB)	2002	884	221.65	698.60	3.11	7458.79
	2001	881	189.41	585.40	3.09	6033.07
	2000	876	166.59	495.75	3.01	4972.78
Total number of employees	2002	884	713	1246.50	50	15500
	2001	883	726	1307.88	50	19906
	2000	879	720	1334.04	50	22044
Years of schooling of General Manager (GM)	2002	878	14.12	2.23	5	18
Years of GM holding the position	2002	877	5.90	4.48	1	23
GM's postgraduate dummy (=1, postgraduate)	2002	878	0.16	0.37	0	1
Management team's average age	2002	854	36.47	5.33	20	51
Management team's average schooling	2002	858	12.11	1.51	8	18
Panel B: R&D and Firm Characteristics						
R&D expenditure by firm (Million RMB)	2002	879	3.21	18.34	0	371.16
	2001	868	2.79	19.67	0	438.07
	2000	867	1.85	9.04	0	119.00
Firm age	2002	885	15.99	14.32	2	52
Shareholding firms dummy	2002	885	0.29	0.46	0	1
State-owned firms dummy	2002	885	0.26	0.44	0	1
Foreign invested firms dummy	2002	885	0.21	0.41	0	1

Similarly, we have column (5)-(8) for Data 2002. In production function, we use value added as dependent variable. The number of skilled workers, the number of less skilled workers, and capital are explanatory variables in production function. Material cost is used as proxy variable in LP and ACF. Our preferred models are LP and ACF.

We can see that all factors in OLS and FE are significant and have positive effect on value added across both datasets. Also, we notice that in LP, only the number of

Table 17: Production Function Estimates

	Data 2000				Data 2002			
	OLS	FE	LP	ACF	OLS	FE	LP	ACF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log(Number of highly educated workers)	0.319***	0.247*	0.250***	0.222	0.421***	0.297**	0.320***	0.454***
	(0.0360)	(0.133)	(0.0432)	(0.195)	(0.0335)	(0.117)	(0.0522)	(0.0973)
Log(Number of less educated workers)	0.0926***	0.334***	-0.00958	0.000807	0.118***	0.338***	0.0110	0.245***
	(0.0310)	(0.125)	(0.0374)	(0.166)	(0.0294)	(0.110)	(0.0431)	(0.0893)
Log(Capital)	0.448***	0.412***	0.565***	0.640***	0.391***	0.147**	0.479***	0.477***
	(0.0228)	(0.0861)	(0.126)	(0.199)	(0.0211)	(0.0575)	(0.0734)	(0.0701)
Constant	3.192***	2.531***			3.428***	5.227***		
	(0.155)	(0.885)			(0.135)	(0.645)		
Number of observations	1724	1724	1724	1706	2351	2351	2351	2351
Adjusted R^2	0.553	0.136			0.553	0.129		

(1) Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
(2) We use value-added as dependent variable.
(3) LP is the method proposed in Levinsohn and Petrin (2003). ACF is the method proposed in Ackerberg, Caves and Frazer (2006). Both methods are popular in production function estimation.
(4) For ACF, we set number of bootstrap replications to 400. The bootstrap samples do "block bootstrapping", i.e., entire time series for each firm is sampled in blocks, which allows the errors to be correlated within plants. Amoeba optimization routine is used and the maximum number of iterative steps that should be done is set to be 500.

skilled workers and capital are significant and the number of less skilled workers has a positive coefficient but not significant in both datasets. Though the coefficient of the number of skilled human capital in the two datasets differs not so much, 0.233 versus 0.320, the coefficients of the number of less skilled workers and capital differ a lot, with less skilled workers 0.0699 versus 0.0110 and capital 0.730 versus 0.479. The coefficients differs even much more in ACF. The coefficient of the number of skilled workers in Data 2002 is around two times as that of it in Data 2000. The coefficient of the number of less skilled workers even has a negative coefficient though not significant in Data 2000 and it is 0.245 and significant at 1% in Data 2002. The coefficient of capital in Data 2000 is 0.762 and is 0.477 in Data 2002 and both are significant at 1%. From the results, we can see that LP and ACF produce very similar results in both datasets and capital in Data 2000 has a much higher coefficient. Our results are consistent with Fleisher et al.(2011) in that their estimates of the elasticity of skilled human capital is also around 0.3, but they get a higher elasticity of less skilled human capital.

Table 18 presents TFP regressions for Data 2000 and Data 2002 respectively. All the models are estimated using OLS but using different TFP measure. Columns (1) and (3) use TFP calculated from LP method. Columns (2) and (4) use TFP calculated from ACF method. We can see that the number of skilled workers has a positive coefficient across all models and both datasets and it is significant in all models except column (4). The significant and positive effect of it in all models indicates that skilled workers still have a productivity effect even when it has been taken into account in production function. For example, skilled workers are important to innovation and thus have very important effect on productivity.

We can get that R&D tends to have a positive effect on TFP across all models and both datasets. The positive relationship of R&D is very easy to understand since more R&D usually means higher absorptive capability as well as innovative ability

Table 18: TFP Estimation Results

	Year 2000(Data 2000)		Year 2002(Data 2002)	
	LP	ACF	LP	ACF
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)
Number of highly educated workers (Hundred)	0.0362***	0.0380***	0.0586***	0.0109
	(0.0116)	(0.0117)	(0.0215)	(0.0147)
General Manager's tenure (years)	0.00490	0.00485	-0.0133	-0.0133*
	(0.00901)	(0.00904)	(0.00813)	(0.00784)
General Manager's postgraduate degree	0.208**	0.205*	0.0351	0.0609
dummy (=1 if has a postgraduate degree)	(0.105)	(0.106)	(0.0935)	(0.0893)
Management team's average age	-0.0159**	-0.0162**	-0.0151*	-0.0131*
	(0.00657)	(0.00659)	(0.00798)	(0.00777)
Management team's average schooling	0.0147	0.0167	0.0593**	0.0622**
	(0.0323)	(0.0325)	(0.0271)	(0.0264)
Panel B: R&D and Firm Characteristics				
Log(average R&D in previous two years)	0.0149**	0.0150**	0.0195***	0.0153**
	(0.00716)	(0.00718)	(0.00611)	(0.00592)
Market share	0.00272	0.00269	0.00573***	0.00638***
	(0.00250)	(0.00252)	(0.00203)	(0.00196)
Firm size (Log(net value of total assets))	-0.0674**	-0.131***	-0.0554*	-0.171***
	(0.0296)	(0.0297)	(0.0293)	(0.0267)
Firm age (year)	-0.00536*	-0.00565*	-0.00837**	-0.0124***
	(0.00290)	(0.00291)	(0.00336)	(0.00328)
Shareholding firms dummy	0.474***	0.477***	0.0617	0.0433
	(0.134)	(0.135)	(0.105)	(0.103)
State-owned firms dummy	-0.176	-0.172	-0.453***	-0.454***
	(0.117)	(0.118)	(0.121)	(0.119)
Foreign invested firms dummy	0.462***	0.450***	0.149	0.176
	(0.124)	(0.125)	(0.111)	(0.107)
Constant	3.624***	3.528***	3.763***	2.898***
	(0.582)	(0.585)	(0.545)	(0.542)
City Dummies	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes
Number of observations	527	527	732	732
Adjusted R^2	0.261	0.263	0.220	0.277

Standard errors in parentheses: * p <0.10, ** p <0.05, *** p <0.01. Dependent variable in column (1) is TFP from LP approach; dependent variable in column (2) is TFP from ACF. Columns (3) and (4) are similar but for different dataset, Data 2002.

and both are very important resources of firm productivity (Griffith et al., 2004). But R&D is only significant in Data 2002 indicating that currently the change in R&D has a much more important effect on productivity than in Data 2000. This is consistent with what we find when use patent application to study innovation. The underlying implication might be that in less developed area (Data 2002), R&D is still not enough and it still constraints firm's development. Thus, change in R&D can bring about large changes in both patent application and firm's productivity.

Also, we find a positive effect of General Manager's tenure in current firms in Data 2000 but a negative effect in Data 2002 and they are all not significant at 5%. This might suggest that in developed area, General Manager's experience are playing its role and is good for firm's productivity since he then will know better about the firm and this will help him to make better decisions. However, in less developed area, General Manager's (GM) tenure is bad for firm's productivity since being under the same General Manager's leading for a long time might bring about lack of flexibility for firms. Unexpectedly, we find a positive coefficient of General Manager's postgraduate degree dummy across all models and both datasets. This indicates that General Manager with a postgraduate degree will promote a firm's productivity. More interesting, we find that the effect is significant in Data 2000 but insignificant in Data 2002, and the effect in Data 2000 is much larger than that in Data 2002.

Management team has a negative and significant effect on productivity across all models and both datasets and the effects in two datasets differ very little. Our results are consistent with our intuition and previous management studies. Older executives tend to be more conservative (Hambrick and Mason, 1984) and empirical studies have found that older top managers tend to be risk averse (Barker and Mueller, 2002) and follow lower-growth strategies (Child, 1974). One reason is that older executives have less of the physical and mental stamina needed to implement organizational

Table 19: IV Estimation Results for TFP

	Year 2000(Data 2000)		Year 2002(Data 2002)	
	LP	ACF	LP	ACF
Panel A: Human Capital Variables	(1)	(2)	(3)	(4)
Number of highly educated workers (Hundred)	0.0308***	0.0326***	0.0645***	0.00952
	(0.0111)	(0.0112)	(0.0211)	(0.0154)
General Manager's tenure (years)	0.00394	0.00449	-0.0136*	-0.0135*
	(0.0108)	(0.0108)	(0.00794)	(0.00765)
General Manager's postgraduate degree	0.163	0.159	-0.0310	-0.00667
dummy (=1 if has a postgraduate degree)	(0.121)	(0.122)	(0.0998)	(0.0946)
Management team's average age	-0.00653	-0.00673	-0.0153**	-0.0131*
	(0.00836)	(0.00841)	(0.00774)	(0.00755)
Management team's average schooling	0.0869**	0.0913**	0.0608**	0.0634**
	(0.0407)	(0.0408)	(0.0263)	(0.0256)
Panel B: R&D and Firm Characteristics				
Log(average R&D in previous two years)	0.00836	0.00835	0.0195***	0.0157***
	(0.00822)	(0.00824)	(0.00597)	(0.00580)
Market share	0.00243	0.00247	0.00583***	0.00648***
	(0.00295)	(0.00297)	(0.00198)	(0.00191)
Firm size (log(net value of total assets))	-0.0477	-0.112***	-0.0590**	-0.168***
	(0.0330)	(0.0333)	(0.0282)	(0.0260)
Firm age (year)	-0.00322	-0.00369	-0.00822**	-0.0123***
	(0.00353)	(0.00355)	(0.00327)	(0.00319)
Shareholding firms dummy	0.540***	0.543***	0.0634	0.0464
	(0.160)	(0.161)	(0.102)	(0.1000)
State-owned firms dummy	-0.202	-0.198	-0.459***	-0.456***
	(0.148)	(0.149)	(0.117)	(0.115)
Foreign invested firms dummy	0.480***	0.466***	0.153	0.178*
	(0.147)	(0.148)	(0.108)	(0.104)
Constant	2.231***	2.113***	3.784***	2.863***
	(0.724)	(0.731)	(0.524)	(0.524)
City Dummies	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes
Number of observations	336	336	732	732
Adjusted R^2	0.249	0.247	0.220	0.276
Tests of endogeneity (F test)	2.825*	2.724*	2.056	2.090
p-value	(0.0608)	(0.0672)	(0.129)	(0.125)
Overidentification test (Score chi2)	7.421	7.626	0.357	0.228
p-value	(0.115)	(0.106)	(0.837)	(0.892)

Standard errors in parentheses: * p <0.10, ** p <0.05, *** p <0.01.Dependent variable in column (1) is TFP from LP approach; dependent variable in column (2) is TFP from ACF. Columns (3) and (4) are similar but for different dataset, Data 2002. Endogenous variables are number of highly educated workers, and general manager's postgraduate degree dummy. Instruments for number of highly educated workers are number of applicants for the skilled positions, vacant weeks of skilled position before filled, city average and industry average both excluding firm itself. Note that for Data 2002, we have no information on applicants and vacant weeks, so we only use city average and industry average excluding firm itself. For general manager's postgraduate degree dummy, we use city and industry average excluding firm itself as instruments. We use two-stage least square (2SLS) to estimate all the models.

changes (Child, 1974). Another reason is that older managers may have greater difficulty grasping new ideas and learning new behaviors (Hambrick and Mason, 1984) because some cognitive abilities seem to diminish with age, including learning ability, reasoning, and memory. Finally, younger managers are likely to have received their education more recently than older managers, so their technical knowledge should be superior (Bantel and Jackson, 1989).

Management team's schooling has a positive effect on productivity across all models and both datasets, implying that the higher education of management team, the more innovation a firm can have. The importance of the top manager's education has been studied in a number of studies. Attained education level is always assumed to be correlated with cognitive ability, and higher levels of education should be associated with higher ability to generate (and implement) creative solutions to complex problems. Hitt and Tyler (1991) found that more educated executives have greater cognitive complexity and such cognitive complexity provides greater ability to absorb new ideas. But the effect in Data 2002 is much larger and significant (it is not significant in Data 2000).

We find that firm age has a significant and negative effect on TFP and the effects are almost the same across both datasets. State ownership has a negative effect on TFP. This makes sense intuitively since state ownership often means bureaucracy and is harm to productivity. Moreover, we find a positive effect of foreign invested firms, indicating that foreign ownership means a higher productivity. This can be easily understood since multinational firms are often firms with higher productivity. Table 4 presents IV estimation results when R&D is believed to be endogenous. All the models are estimated using two-stage least square. We use city average and industry average of corresponding R&D excluding firm's own R&D as instruments. All the first stages have very good performance. However, we don't have enough evidence to reject the tests of endogeneity. Thus, our basic regression results still hold.

5.5 *Conclusion*

This paper tries to study the effects of R&D spending and firm-level human capital on productivity in Chinese manufacturing firms to account for the effect of innovation on productivity. Different from studies, we also include firm-level human capital. The reasons why firm-level human capital is important are: (1) human capital provides foundation for R&D to induce innovation. (2) human capital is complementary to R&D. (3) skilled human capital itself can bring about innovation through learning by doing. Using firm-level data in Chinese manufacturing firms, we find that firms in more developed area operate under constant return to scale while firms in less developed area operate under decreasing return to scale. An important conclusion is that skilled human capital and R&D have a significant effect on productivity even after skilled human capital is included in production function. We find that firm age has a significant and negative effect on TFP and the effects are almost the same across both datasets. State ownership has a negative effect on TFP. This makes sense intuitively since state ownership often means bureaucracy and is harm to productivity. Moreover, we find a positive effect of foreign invested firms, indicating that foreign ownership means a higher productivity.

The different effects across the two datasets provide us more insights into firm's productivity. However, we still want to point out several limitations of our work. First, the ideal dependent variable should be inventory adjusted output, but because we don't have the related material cost information, we still choose to use the unadjusted one. This might bring about bias results. If inventory is very common for all firms, then we tends to get a downward bias. Second, we only have average education level in our datasets, but no information on education quality which might be very important in human capital study. Third, we only have three years of data, which might be not enough for us to study the TFP growth.

CHAPTER VI

CONCLUSION AND OUTLOOK

This thesis studies firm innovation in China from a human capital perspective. There are several reasons for us to study firm innovation in China from a human capital point of view. First, R&D spending in China is also an objective pursued by government and firm itself due to tax incentive or government procurement. Thus, innovation studies using R&D as the main determinants might be biased. Second, in China, learning technology introduced from advanced country is still important. Incremental and informal innovation rather than formal R&D might be more important for a lot of firms. Thus, rather than R&D, skilled human capital can better capture firm innovation. Third, according to resource-based theory and upper echelon theory, human capital is the most important thing in a firm and determines firm's behavior.

The third chapter explores firm-level innovation from a human capital point of view using patents as proxy for innovation. In the theoretical model, two firms compete with each other in a three-stage Cournot, innovation stage and production stage. Skilled human capital level can affect innovation success probability directly and via R&D level indirectly. Managerial human capital can affect firm innovation through their choice of projects and R&D level. We find that a firm's innovation is not only determined by its human capital level, firm characteristics, and its market share, but also might be affected by market environment. In the empirical study, we use two firm-level datasets from China, one from metropolitan cities and one from provincial middle cites. Human capital indicators are skilled human capital (number of highly educated workers), general manager's education and experience, and management team's education and age. We find that a firm's skilled human capital and managerial personnel's education have significantly positive effects on innovation while the management team's age has a significantly negative effect on innovation. The effect of R&D on patents is insignificant in large metropolitan cities while it is positive and significant in provincial middle cities.

In the fourth chapter, we study the relationship between firm-level human capital and product innovation. Three measures are used: new product sales proportion, new product sales, and number of new products. We find that for skilled human capital, we find that it also tends to have a positive effect across the three measures, though it is not significant in regression using new product proportion as product innovation measure. When new product sales proportion and new product sales are used as product innovation measures, we find that management team's average years of schooling has a positive effect in determining product innovation, however, its effect in Data 2000 when new product count used as product innovation measure is negative and significant. This indicates that though the results from the three regressions are generally consistent, but they still differ. When compare the results using different innovation measures, we should be very cautious. Another thing is that different from results in last chapter where R&D has a positive effect only in Data 2003, less developed areas, we find that R&D has a very significant and positive effect on product innovation no matter which product innovation measure is used. This indicates that R&D is still important in promoting product innovation and policies supporting investment in R&D is still important. Our results still hold when endogeneity is considered. Notably, in all three models, we all find that general manager's postgraduate degree has a large and significant effect in less developed areas but insignificant at all in more developed areas. The reason might be that in less developed areas, market development is much more incomplete and thus general manger's education matters more. A higher education can enable a general manager to make more insightful decision.

In the fifth chapter, the relationship between firm-level human capital, R&D and total factor productivity (TFP) is examined. Firstly, production function is estimated using methods proposed by Levinsohn and Petrin (2003, LP, thereafter) and Ackerberg, Caves and Frazer (2006, ACF) since inputs are very likely to be correlated with productivity. We then use TFP from ACF method to examine the determinants of TFP. We find that skilled human capital is important in determining TFP even when it is already included in production function. Management team's age has a negative effect on TFP. Management team's average schooling has a positive and significant effect in less developed areas, but we fail to find a significant effect in more developed areas. Notably, R&D has a positive and significant effect across both datasets. When the endogeneity of number of highly educated workers, and general manager's postgraduate degree is considered, our main results still hold.

Throughout the whole thesis, generally our results are very robust to two datasets though there are also some differences: (1) postgraduate degree are found to have a large and positive effect in Data 2002, less developed areas in patents estimation and product innovation. But postgraduate degree is found to have a positive and significant effect in Data 2002 rather than in Data 2003. I have no explanation for this. (2) highly educated workers have a much larger effect on patents in Data 2002 when patents are estimated. The reason might be that in more developed area, there are enough R&D, almost 5 times of R&D in less developed areas, and R&D can facilitate human capital exert more effect on innovation. (4) we only find a positive and significant effect of R&D in Data 2003 when patents are estimated. Since there's enough R&D in Data 2002, the effect of R&D on innovation has been absorbed by the effect of human capital. Also, the significance of the effect of R&D in Data 2003 might indicate that there's not enough of R&D in less developed areas. Thus, the implication is that to promote innovation, we should invest more in human capital in more developed areas while we need to invest more in R&D in less developed areas.

Future work should focus on estimating production function in different industries. In the current work, we find that the production function estimation results differ a lot in more developed areas and less developed areas. One possible reason might be market environment. Input prices are different in different market environment, and thus the inputs will be different. Another important reason might be that industries are not exactly the same for the two datasets and the distribution of firms in industries are also different. When we only take observations within the same industries and same firm distribution in industries, we will get more insight.

We can also enrich our work by examining how different innovation measures influence TFP. More specifically, we can check how patents, new product sales, new product proportion, and new product counts might play a role in firm's productivity. We can also try to deal with the measurement error of capital input in production function estimation. For example, we have data on depreciation of building. On the one hand, it is closely related to capital, but on the other hand, it is unlikely to be correlated with productivity. Also, besides LP and ACF, we can try to use GMM method proposed by Wooldridge (2009). Moreover, we can try to use firm's investment in machinery as proxy and use OP to estimate production function.

With our current datasets, there is also much more work to be done in the future. For example, we can study R&D cooperation since our datasets provide us the information on cooperation with local university, government research institution, private research institution and private companies. This is a very interesting topic to pursue since nowadays innovation are achieved both through in-house innovation and also outside channel and cooperation. Thus, the R&D cooperation study will largely enrich our work on firm innovation in China. Moreover, we can also study how workers of different positions affect firm innovation since firm innovation is a way of exchanging human capital.

REFERENCES

[1] Ackerberg, D., Caves, K., & Frazer, G. (2006). Structural identification of production functions. *Working paper*.

[2] Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research policy*, 31(7), 1069-1085.

[3] Acs, Z. J.,& Audretsch, D. B. (1988). Innovation in large and small firms: an empirical analysis. *The American Economic Review*, 678-690.

[4] Aghion, P., Boustan, L., Hoxby, C., & Vandenbussche, J. (2009). The causal impact of education on economic growth: evidence from US. *Unpublished paper*.

[5] Aghion and Howitt (2009). *The Economics of Growth*, MIT press.

[6] Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity: Economic and Social Factors* (pp. 609-626). Princeton University Press.

[7] Arundel, A., & Kabla, I. (1998). What percentage of innovations are patented? Empirical estimates for European firms. *Research Policy*, 27(2), 127-141.

[8] Atuahene-Gima, K., & Li, H. (2004). Strategic decision comprehensiveness and new product development outcomes in new technology ventures. *Academy of Management Journal*, 47(4), 583-597.

[9] Audretsch, D. B., & Acs, Z. J. (1991). Innovation and size at the firm level. *Southern Economic Journal*, 739-744.

[10] Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review*, 630-640.

[11] Balachandra, R., & Friar, J. H. (1997). Factors for success in R&D projects and new product innovation: a contextual framework. *Engineering Management, IEEE Transactions on*, 44(3), 276-287.

[12] Ballot, G., Fakhfakh, F., & Taymaz, E. (2001). Firms' human capital, R&D and performance: a study on French and Swedish firms. *Labour Economics*, 8(4), 443-462.

[13] Bantel, K. A., & Jackson, S. E. (1989). Top management and innovations in banking: does the composition of the top team make a difference? *Strategic Management Journal*, 10(S1), 107-124.

[14] Barker III, V. L., & Mueller, G. C. (2002). CEO characteristics and firm R&D spending. *Management Science*, 48(6), 782-801.

[15] Barro, R. J. (2001). Human capital and growth. *American Economic Review*, 12-17.

[16] Basmann, R. L. (1960). On finite sample distributions of generalized classical linear identifiability test statistics. *Journal of the American Statistical Association* 55: 650659.

[17] Baum, C.F. (2008).Modeling proportions. *Stata Journal* 8: 299303.

[18] Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34(2), 143-173.

[19] Bhattacharya, S., & Ritter, J. R. (1983). Innovation and communication: Signaling with partial disclosure. *The Review of Economic Studies*, 50(2), 331-346.

[20] Blundell, R., Griffith, R., & Van Reenen, J. (1999). Market share, market value and innovation in a panel of British manufacturing firms. *The Review of Economic Studies*, 66(3), 529-554.

[21] Booz, Allen and Hamilton (1982). *New Products Management for the 1980s*. Booz, Allen and Hamilton. Inc., New York.

[22] Bound, J., Cummins, C., Griliches, Z., Hall, B. H., & Jaffe, A. B. (1984). Who Does R&D and Who Patents? *In R & D, Patents, and Productivity* (pp. 21-54). University of Chicago Press.

[23] Brouwer, E., & Kleinknecht, A. (1999). Innovative output, and a firm's propensity to patent.: An exploration of CIS micro data. *Research Policy*, 28(6), 615-624.

[24] Calantone, R. J., Cavusgil, S. T., & Zhao, Y. (2002). Learning orientation, firm innovation capability, and firm performance. *Industrial marketing management*, 31(6), 515-524.

[25] Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data* (No. 53). Cambridge university press.

[26] Cassiman, B., & Vanormelingen, S. (2013). Profiting from innovation: Firm level evidence on markups. Available at SSRN 2381996.

[27] Chen, Y., & Schwartz, M. (2013). Product innovation incentives: Monopoly vs. competition. *Journal of Economics & Management Strategy*, 22(3), 513-528.

[28] Child, J. (1974). Managerial and organizational factors associated with company performance part I. *Journal of Management Studies*, 11(3), 175-189.

[29] Coe, D. T., & Helpman, E. (1995). International R&D spillovers. *European economic review*, 39(5), 859-887.

[30] Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative science quarterly*, 128-152.

[31] Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science*, 48(1), 1-23.

[32] Cook, D. O., Kieschnick, R., & McCullough, B. D. (2008). Regression analysis of proportions in finance with self selection. *Journal of Empirical Finance*, 15(5), 860-867.

[33] Cooper, R. G. (1975). Why new industrial products fail. *Industrial Marketing Management*, 4(6), 315-326.

[34] Covin, J. G., & Slevin, D. P. (1989). Strategic management of small firms in hostile and benign environments. *Strategic Management Journal*, 10(1), 75-87.

[35] Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, 829-844.

[36] Darroch, J. (2005). Knowledge management, innovation and firm performance. *Journal of Knowledge Management*, 9(3), 101-115.

[37] Dhongde, S. (2014) Economic Development and CBRN Terrorism. *Working paper*.

[38] Dougherty, D. (1990). Understanding new markets for new products. *Strategic Management Journal*, 59-78.

[39] Ernst, D., & Naughton, B. J. (2012). Global Technology Sourcing in China's Integrated Circuit Design Industry: A Conceptual Framework and Preliminary Findings. *East-West Center Working Paper*, (131).

[40] Fleisher, B. M., Hu, Y., Li, H., & Kim, S. (2011). Economic transition, higher education and worker productivity in China. *Journal of Development Economics*, 94(1), 86-94.

[41] Furman, J. L., Porter, M. E., & Stern, S. (2002). The determinants of national innovative capacity. *Research policy*, 31(6), 899-933.

[42] Furman, J. L., & Hayes, R. (2004). Catching up or standing still?: National innovative productivity among followercountries, 19781999. *Research Policy*, 33(9), 1329-1354.

[43] Ge, D., & Fujimoto, T. (2004). Quasi-open Product Architecture and Technological Lock-in. *Annals of Business Administrative Science*, 3(2), 15-24.

[44] Geroski, P. A. (1990). Innovation, technological opportunity, and market structure. *Oxford Economic Papers*, 586-602.

[45] Gourieroux, C., Monfort, A., & Trognon, A. (1984). Pseudo maximum likelihood methods: Applications to Poisson models. *Econometrica: Journal of the Econometric Society*, 701-720.

[46] Grabowski, H. G. (1968). The determinants of industrial research and development: A study of the chemical, drug, and petroleum industries. *The Journal of Political Economy*, 292-306.

[47] Greenhalgh, C., & Rogers, M. (2010). *Innovation, intellectual property, and economic growth*. Princeton University Press.

[48] Griffith, R., Redding, S., & Van Reenen, J. (2004). Mapping the two faces of R&D: Productivity growth in a panel of OECD industries. *Review of Economics and Statistics*, 86(4), 883-895.

[49] Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics*, 92-116.

[50] Griliches, Z. (1980). Returns to research and development expenditures in the private sector. In New developments in productivity measurement (pp. 419-462). University of Chicago Press.

[51] Griliches, Z. (1990). Patents Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28(4), 1661-1707.

[52] Griliches, Z., & Regev, H. (1995). Firm productivity in Israeli industry 1979-1988. *Journal of Econometrics*, 65(1), 175-203.

[53] Grimpe, C., & Sofka, W. (2009). Search patterns and absorptive capacity: Low and High-technology sectors in European countries. *Research Policy*, 38(3), 495-506.

[54] Grimpe, C., & Sofka, W. (2009). Search patterns and absorptive capacity: Low- and high-technology sectors in European countries. *Research Policy*, 38(3), 495-506.

[55] Hall, B. H., Griliches, Z., & Hausman, J. A. (1986). Patents and R&D: Is there a lag? *NBER working paper*.

[56] Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools (No. w8498). *National Bureau of Economic Research*.

[57] Hall, B. H., Lotti, F., & Mairesse, J. (2009). Innovation and productivity in SMEs: empirical evidence for Italy. *Small Business Economics*, 33(1), 13-33.

[58] Hall, B. H., & Mairesse, J. (1995). Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics*, 65(1), 263-293.

[59] Hall, B. H., & Ziedonis, R. H. (2001). The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995. *RAND Journal of Economics*, 101-128.

[60] Hambrick, D. C., & Fukutomi, G. D. (1991). The seasons of a CEO's tenure. *Academy of Management Review*, 16(4), 719-742.

[61] Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review*, 9(2), 193-206.

[62] Hansen, P. A., & Serin, G. (1997). Will low technology products disappear? The hidden innovation processes in low technology industries. *Technological Forecasting and Social Change*, 55(2), 179-191.

[63] Hausman, J., Hall, B. H., & Griliches, Z. (1984). Econometric Models for Count Data with an Application to the Patents-R&D Relationship. *Econometrica: Journal of the Econometric Society*, 909-938.

[64] Hitt, M. A., & Tyler, B. B. (1991). Strategic decision models: Integrating different perspectives. *Strategic Management Journal*, 12(5), 327-351.

[65] Hitt, M. A., Biermant, L., Shimizu, K., & Kochhar, R. (2001). Direct and moderating effects of human capital on strategy and performance in professional service firms: A resource-based perspective. *Academy of Management Journal*, 44(1), 13-28.

[66] Hopkins, D. S. (1980). New-product winners and losers. Conference Board.

[67] Howitt, P., & Mayer-Foulkes, D. A. (2005). R&D, Implementation, and Stagnation: A Schumpeterian Theory of Convergence Clubs. *Journal of Money, Credit and Banking*, 37(1), 147-77.

[68] Hu, A. G. (2001). Ownership, government R&D, private R&D, and productivity in Chinese industry. *Journal of Comparative Economics*, 29(1), 136-157.

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

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

[71] Jung, D. D., Wu, A., & Chow, C. W. (2008). Towards understanding the direct and indirect effects of CEOs' transformational leadership on firm innovation. *The Leadership Quarterly*, 19(5), 582-594.

[72] Katila, R. (2002). New product search over time: past ideas in their prime? *Academy of Management Journal*, 45(5), 995-1010.

[73] Kim, L., & Nelson, R. R. (Eds.). (2000). *Technology, Learning, and Innovation: Experiences of Newly Industrializing Economies*. Cambridge University Press.

[74] Kortum, S., & Lerner, J. (1999). What is behind the recent surge in patenting?. *Research Policy*, 28(1), 1-22.

[75] Kortum, S., & Lerner, J. (2000). Assessing the contribution of venture capital to innovation. *RAND Journal of Economics*, 674-692.

[76] Krueger, A. B. (1993). How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989. *The Quarterly Journal of Economics*, 108(1), 33-60.

[77] Landstrom, H., Harirchi, G., & Astrom, F. (2012). Entrepreneurship: Exploring the knowledge base. *Research Policy*, 41(7), 1154-1181.

[78] Leiponen, A., & Helfat, C. E. (2010). Innovation objectives, knowledge sources, and the benefits of breadth. *Strategic Management Journal*, 31(2), 224-236.

[79] Levin, R. C., Cohen, W. M., & Mowery, D. C. (1985). R & D appropriability, opportunity, and market structure: new evidence on some Schumpeterian hypotheses. *The American Economic Review*, 20-24.

[80] Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2), 317-341.

[81] Li, H., & Atuahene-Gima, K. (2001). Product innovation strategy and the performance of new technology ventures in China. *Academy of Management Journal*, 44(6), 1123-1134.

[82] Li, H., Fraumeni, B. M., Liu, Z., & Wang, X. (2009). Human capital in China (No. w15500). *National Bureau of Economic Research working paper*.

[83] Lieberman, M. B. (1987). Patents, learning by doing, and market structure in the chemical processing industries. *International Journal of Industrial Organization*, 5(3), 257-276.

[84] Lin, C., Lin, P., Song, F. M., & Li, C. (2011). Managerial incentives, CEO characteristics and corporate innovation in China's private sector. *Journal of Comparative Economics*, 39(2), 176-190.

[85] Link, P. L. (1987). Keys to new product success and failure. *Industrial Marketing Management*, 16(2), 109-118.

[86] Liu, X., & Buck, T. (2007). Innovation performance and channels for international technology spillovers: Evidence from Chinese high-tech industries. *Research policy*, 36(3), 355-366.

[87] Lucas Jr, R. E. (1993). Making a miracle. *Econometrica: Journal of the Econometric Society*, 251-272.

[88] Marschak, J., & Andrews, W. H. (1944). Random simultaneous equations and the theory of production. *Econometrica, Journal of the Econometric Society*, 143-205.

[89] Maskell, P., & Malmberg, A. (1999). Localised learning and industrial competitiveness. *Cambridge journal of economics*, 23(2), 167-185.

[90] Mendershausen, H. (1938). On the significance of Professor Douglas' production function. *Econometrica: Journal of the Econometric Society*, 143-153.

[91] Miller, D. (1987). The structural and environmental correlates of business strategy. *Strategic Management Journal*, 8(1), 55-76.

[92] Miller, D., De Vries, M. F. K., & Toulouse, J. M. (1982). Top executive locus of control and its relationship to strategy-making, structure, and environment. *Academy of Management Journal*, 25(2), 237-253. *M*orrison, W. M. (2014). *China's economic rise: History, trends, challenges, and implications for the United States*. Washington, DC: Congressional Research Service.

[93] Myers, S., & Marquis, D. G. (1969). *Successful industrial innovations: A study of factors underlying innovation in selected firms* (pp. 69-17). Washington, DC: National Science Foundation.

[94] Nahm, J., & Steinfeld, E. S. (2014). Scale-up nation: Chinas specialization in innovative manufacturing. *World Development*, 54, 288-300.

[95] Nee, V. (1992). Organizational dynamics of market transition: Hybrid forms, property rights, and mixed economy in China. *Administrative science quarterly*, 1-27.

[96] Nelson, R. R., & Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *The American Economic Review*, 69-75.

[97] Nelson, R. R., & Winter, S. G. 1982. *The evolutionary theory of economic change.* Gambridge, MA: Belknap Press.

[98] Nishimizu, M., & Page, J. M. (1982). Total factor productivity growth, technological progress and technical efficiency change: dimensions of productivity change in Yugoslavia, 1965-78. *The Economic Journal*, 920-936.

[99] OECD (2014), OECD Science, Technology and Industry Outlook 2014, OECD Publishing, Paris. DOI:http://dx.doi.org/10.1787/sti_outlook-2014-en.

[100] Olley, G. S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64: 1263-1297.

[101] Pakes, A., & Griliches, Z. (1980). Patents and R&D at the firm level: A first report. *Economics Letters*, 5(4), 377-381.

[102] Pakes, A., & Griliches, Z. (1984). Patents and R&D at the firm level: a first look. *In R&D, Patents, and Productivity* (pp. 55-72). University of Chicago Press.

[103] Pakes, L. E. and J. Wooldridge. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics* 11: 619632.

[104] Peng, M. W., & Heath, P. S. (1996). The growth of the firm in planned economies in transition: Institutions, organizations, and strategic choice. *Academy of Management Review*, 21(2), 492-528.

[105] Pennings, J. M., Lee, K., & Van Witteloostuijn, A. (1998). Human capital, social capital, and firm dissolution. *Academy of Management Journal*, 41(4), 425-440.

[106] Penrose, E. (2009). *The Theory of the Growth of the Firm*. Oxford University Press.

[107] Petrin, A., Poi, B. P., & Levinsohn, J. (2004). Production function estimation in Stata using inputs to control for unobservables. *Stata Journal*, 4, 113-123.

[108] Pfeffer, J. (1995). *Competitive Advantage Through People: Unleashing the Power of the Work force*. Harvard Business Press.

[109] Prescott, E. C. (1998). Needed: A theory of total factor productivity. *International Economic Review*, 39(3), 525-552.

[110] Romer, P. M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy*, 1002-1037.

[111] Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, S71-S102.

[112] Romer, P. M. (1993). Idea gaps and object gaps in economic development. *Journal of Monetary Economics*, 32(3), 543-573.

[113] Rosen, R. J. (1991). Research and development with asymmetric firm sizes. *The RAND Journal of Economics*, 411-429.

[114] Rubenstein, A.H.,A.K.Chakrabarti, R.D. OKeffe, W.E. Souder and H.C. Young (1976). Factors Influencing Innovation Success at the Project Level. *Research Management*, (May 1976), 15-20.

[115] Salomon, R. M., & Shaver, J. M. (2005). Learning by exporting: new insights from examining firm innovation. *Journal of Economics & Management Strategy*, 14(2), 431-460.

[116] Sargan, J. D. 1958. The estimation of economic relationships using instrumental variables. *Econometrica* 26: 393415.

[117] Scherer, F.M. (1965a). Size of Firm, Oligopoly and Research: A Comment. *Canadian Journal of Economics and Political Science*, 31, pp. 256-266.

[118] Scherer, F.M. (1965b). Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions. *American Economic Review*, 55, pp. 1097-1125.

[119] Scherer, F. M. (1982). Inter-industry technology flows in the United States. *Research policy*, 11(4), 227-245.

[120] Scherer, F. M. (1983). The propensity to patent. *International Journal of Industrial Organization*, 1(1), 107-128.

[121] Scherer, F. (1984a). Using linked patent and R&D data to measure interindustry technology flows. In *R&D, patents, and productivity* (pp. 417-464). University of Chicago Press.

[122] Scherer, F. M. (1984b). *Innovation and Growth: Schumpeterian Perspectives*. MIT Press Books, 1.

[123] Scherer, F. M. (1986). Innovation and growth: Schumpeterian perspectives. MIT Press Books, 1.

[124] Schumpeter, J. A. (1942). *Socialism, Capitalism and Democracy*. Harper and Brothers.

[125] Singell Jr, L. D. (1993). Managers, specific human capital, and firm productivity in Major League Baseball. *Atlantic Economic Journal*, 21(3), 47-59.

[126] Solow, R.M. (1957) Technical change and the aggregate production function. *Review of Economics and Statistics* 39: 312320.

[127] Souitaris, V. (2002). Technological trajectories as moderators of firm-level determinants of innovation. *Research Policy*, 31(6), 877-898.

[128] Spence, M. (1984). Cost reduction, competition, and industry performance. *Econometrica: Journal of the Econometric Society*, 101-121.

[129] Swann, G. P. (2009). *The Economics of Innovation: An Introduction*. Edward Elgar Publishing.

[130] Symeonidis, G. (1996), Innovation, Firm Size and Market Structure: Schumpeterian Hypotheses and Some New Themes, *OECD Economics Department Working Papers*, No. 161, OECD Publishing.

[131] The World Bank (2002). Study of Competiveness, Technology & Firm Linkages. http://microdata.worldbank.org/index.php/catalog/651.

[132] The World Bank (2003). Investment Climate Survey. http://microdata.worldbank.org/index.php/catalog/591.

[133] Thoenig, M., & Verdier, T. (2003). A theory of defensive skill-biased innovation and globalization. *American Economic Review*, 709-728.

[134] Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: Journal of the Econometric Society*, 24-36.

[135] Utterback, J. M., & Abernathy, W. J. (1975). A dynamic model of process and product innovation. *Omega*, 3(6), 639-656.

[136] Vandenbussche, J., Aghion, P., & Meghir, C. (2006). Growth, distance to frontier and composition of human capital. *Journal of Economic Growth*, 11(2), 97-127.

[137] Wooldridge, J. M. 1995. Score diagnostics for linear models estimated by two stage least squares. In *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao*, ed. G. S. Maddala, P. C. B. Phillips, and T. N. Srinivasan, 6687. Oxford: Blackwell.

[138] Wooldridge, J. (2012). *Introductory Econometrics: A Modern Approach*. Cengage Learning.

[139] Wu, J. (2013). The effects of external knowledge search and CEO tenure on product innovation: evidence from Chinese firms. *Industrial and Corporate Change*, dtt009.

[140] Xin, K. K., & Pearce, J. L. (1996). Guanxi: Connections as substitutes for formal institutional support. *Academy of Management Journal*, 39(6), 1641-1658.

[141] Zheng, Y., & Tong, S. Y. (Eds.). (2014). *China's Evolving Industrial Policies and Economic Restructuring*. Routledge.

[142] Zirger, B. J., & Maidique, M. A. (1990). A model of new product development: an empirical test. *Management Science*, 36(7), 867-883