5-15-2019

Essays on the Chinese Economy

Fan Duan

Louisiana State University and Agricultural and Mechanical College

ESSAYS ON THE CHINESE ECONOMY

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Economics

by
Fan Duan
B.A., Guangxi University, 2008
M.A., Zhejiang University, 2011
M.S., Louisiana State University, 2015
August 2019

Dedicated to Wenbin Duan

Acknowledgments

First of all, I would like to express my sincere gratitude to Prof. Bulent Unel, for his continuous encouragement and support at Louisiana State University. He always encourages and teaches me to become a better thinker and researcher. Without his great knowledge, sharp insight, and patience, this dissertation would not have been possible.

I thank my committee members, Professors Areendam Chanda and Fang Yang, for their helpful comments and suggestions. I am also indebted to Mr. Charles Roussel for helping me to become a qualified instructor. I also thank to my friends Swarup Joshi, Anindo U. Sarker, and Ishita Tripathi for making my graduate study more enjoyable.

I am grateful to my wife, Jiayi Wu, who has supported this great adventure from the very beginning. Her love makes me a better and stronger person. I have found great support from my family during my time at Louisiana State University. I have my special thanks to my father, Gejun Duan, my mother, Meilian Wang, and all members of my family for their selfless love.

Table of Contents

List of Tables

List of Figures

Abstract

This dissertation includes three essays which contribute to the literature on economic growth, innovation, and international trade in China. Chapter 2 examines the implication of early economic prosperity, measured by the city-level population density in 1776, for modern day China. It shows that uneven economic performance across cities in the modern-day China can be traced to variation in living standards in 1776. Chapter 3 investigates the effects of innovation and human capital on firms' export decisions over the 2003-2011 period, and it shows that more innovative and skill-intensive firms are more likely to export and stay in the export market. Finally, Chapter 4 investigates the impact of devoting resources to R&D and workers' training on development of new products. I find that investment in R&D and workers' training have positive and statistically highly significant effects on the introduction of new goods.

Chapter 1. Introduction

Since the launch of reforms in 1978, China has been the most rapidly growing economy in the world. In terms of total output produced, China now has the second largest economy in the world. However, economic performance has been highly uneven across its different regions. Several inland provinces have been falling behind the prosperous coastal ones. In 2010, for example, the income per capita in Shanghai was more than five times that of Gansu. This unequal performance has been a challenge to policy makers, and researchers have identified several factors that might drive it. In this dissertation, I investigate how key factors such as early development, international trade, and innovation have influenced economic development and growth in China.

Chapter 2 investigates implications of early-stage development for present economic development in China.[1] Specifically, this chapter argues that the wide economic differences across regions in China can be traced to their early developmental levels. Using the city-level population densities across 227 cities in 1776 as a measure of the early economic prosperity, and it appears that population density is a strong predictor of the current development indexes such as night-light density, GDP per capita, average years of schooling, and trade openness. In other words, economically prosperous cities from about two centuries ago tend to be brighter, richer, more educated, and more open today. It is remarkable to observe that despite two centuries of massive changes in economics, politics, and social order, the more prosperous cities of the Qing empire are still experiencing higher standards of living today.

Chapter 3 investigates the factors that explain international trade behavior at the firm level. International trade has played an important role in boosting China's economy for decades. According to the merchandise trade data published by WTO, China's exports in

[1]This chapter previously appeared as Duan, Fan, and Bulent Unel, Persistence of cities: Evidence from China, *Review of Development Economics*, 2019, *23(2)*, 663 – 676, reprinted by permission.

2011 were about one hundred times greater than in 1980. To investigate the determinants of the fast growth in international trade, this chapter employs a manufacturing industries data set which covers more than two million observations the 2003 – 2011 period. This chapter complements and extends the existing literature by considering the effects of two important factors on export status: innovation, which is measured by the output value of new products, and investment in human capital. The decision to export is indicated by the first-time export entry and exit decision, to mitigate the concern over reverse causality. Following Berand and Jensen (1999), I find that the more innovative a firm is, the more likely it will start exporting, and less likely it will be to stop exporting. It also finds that the impact of innovation on trade behavior is more substantial in high technology, high capital-intensity industries.

Chapter 4 investigates determinants of innovation at the firm-level. Using the insights from Romer's (1990) model and utilizing the firm-level data used in the previous chapter, it investigates the impact of investment in R&D and human capital on introducing new products. I find that investment in R&D and workers' training have positive and statistically highly significant effects on the introduction of new goods. The econometric model predicts that a 10-percent increase in the R&D investment increases sales of new products in the next year about 3.5 percent. Similarly, a 10-percent increase in spending on workers' training increases sales by 0.5 percent. Results also indicate that the impact of R&D on new products is comparable across state owned enterprises (SOE) and non-SOEs. However, the impact of investment in workers' training is higher among non-SOEs. The analysis at the industry-level yields that the impact of R&D and human capital is stronger in high-tech industries.

Chapter 2. Persistence of Development in China

2.1 Introduction

China now has the second largest economy in the world, but the living standards vary significantly across its regions (Unel and Zebregs 2009). In 2010, for example, the income per capita in Shanghai was more than five times that of Gansu. Previous studies have mainly focused on identifying factors that may have driven unequal performance across regions after the reforms began in 1978. The factors include reallocation of resources from agriculture to manufacturing and service (Brandt et al. 2008), enterprise restructuring and privatization (Dong et al. 2006, Jefferson et al. 2008, Hsieh and Klenow 2009), financial reorganization (Yi 2010), and globalization (Unel and Zebreg 2009, Sun and Heshmati 2010).

These studies, however, have paid limited attention to the fact that living standards differed significantly across regions even before the reforms. In 1978, for example, the income per capita in Shanghai was about 3.5 times that of Gansu, suggesting a persistence of cross-regional differences in living standards. This paper argues that the wide economic differences across regions in China can be traced to their early developmental levels. More precisely, we use the city-level population densities across 227 cities in 1776 of the Qing dynasty as our measure of the early economic prosperity, and show that it is a strong predictor of the current development indexes such as night-light density, per capita GDP, the average years of schooling, and trade openness. That is, economically prosperous cities from about two centuries ago tend to be brighter, richer, more educated, and more open today. It is remarkable to observe that despite two centuries of massive changes in economics, politics, and social order (as we shall discuss in the next section), many of the more prosperous cities of the Qing empire are still enjoying higher standards of living today.

Before moving further, we want to discuss why we choose the population density in 1776 as a proxy for the early development. We acknowledge the fact that the relationship between population density and economic development is quite complex, and that more densely populated areas do not necessarily imply higher economic prosperity. For example, a higher than sustainable level of population may eventually limit population growth since resources are scarce. However, Acemoglu et al. (2002) argue that during these early times, only relatively prosperous areas could support dense population. Using the data from colonized countries in 1500, they provide evidence that population density is closely associated with urbanization and income per capita. For lack of a better alternative, following Acemoglu et al. (2002), we consider population density as a good proxy for the economic prosperity in the late eighteenth-century China.

We consider 1776 for two reasons. First, this is the earliest year for which have the most comprehensive and reliable data (Cao 2001). The data include 227 cities from 24 provinces (Figure 2.1), which cover the core and frontier regions of China (known as China Proper).[1] Cao (2000) has city-level population data from 1393 of the Ming dynasty, but this survey contains a smaller set of cities and is subject to measurement errors. Second, the Qing dynasty had expanded geographically, economically, and politically until the late eighteen century, and 1770s mark the height of Qing power under the Qianlong Emperor (1736–1795). Pomeranz (2000) argues that the living standard during the second half of 1700s was likely higher than that in other parts of the world.

In measuring the current development level, we consider night-light density, output per capita, average years of schooling, and trade openness. Except for the night-light density, other measures of development have been extensively used in the literature. In an interesting paper, Henderson et al. (2012) use satellite data on night lights to augment official growth measures. They use changes in night lights as a measure of economic growth, and show that light growth is strongly correlated with GDP growth.[2] Using light density as a measure of

[1] Cao (2001) also has population survey in 1820, which covers 233 cities. Using the population density in 1820 yields qualitatively the same results.

[2] Following Henderson et al. (2012), several studies have used night lights to analyze regional inequality within and across countries (e.g., Lessmann and Seidel 2016, Henderson et al. (2017). See Donaldson and Storeygard (2016) for a comprehensive survey about the applications of satellite data in economics.

Figure 2.1: Provinces and Cities Used in the Analysis

development is particularly relevant to our analysis, because in China, a significant amount of production is carried out in the informal sector (usually not counted by officials) and the Chinese official statistics are subject to serious measurement errors (Young 2005, Clark et al. 2017).

This paper first relates to a growing literature that investigates the effects of historical variables on contemporary development. A full account of this literature is beyond the scope of this paper, however, we touch upon a few influential studies here.[3] In two seminal papers, Acemoglu et al. (2001 & 2002) argue that past institutions created by various types of colonization policies shaped current institutions. Bockstette et al. (2002) argue that a longer history of statehood might be favorable to economic development. They derive a state antiquity index, and show that it is strongly correlated with current institutional quality, income per capita, and growth. Comin et al. (2010) show that 1500 AD technology is a

[3]See Spolaore and Wacziarg (2013) for an extensive review of this literature.

strong and robust predictor of per capita income and technology adoption today. These studies use country-level aggregate data, whereas we use city-level data to investigate the impact of early development on various development indicators in China.

One strand of this literature documents the persistence of regional development. Davis and Weinstein (2002) consider the Allied bombing of Japanese cities during the second World War, and find that most cities returned to their relative position in the distribution of city sizes within about 15 years. Henderson et al. (2017) consider the 119 European cities in 10 modern European countries in 1500, and show that only 15 of them have fewer than 50,000 people today.[4] Finally, Chanda and Ruan (2017) construct a measure of urban population density in 1850 for more than 2,000 sub-national regions across 135 countries, and find strong evidence of persistence in regional development. Since our analysis focuses on China, we investigate persistence at a more detailed regional level. In this way, we can also control for differences in institutions and cultures across the provinces.

Our paper also joins a large body of work that investigates the determinants of China's development. One area of this literature focuses on market-oriented reforms implemented in the late 1970's and their effects on China's current economic performance.[5] Another area evaluates China's economic development from a longer time perspective. In this area, some studies investigate why China fell behind Western Europe despite being economically comparable until the 19th century.[6] Our paper relates to studies that link China's recent economic performance to the historical antecedents. Brandt et al. (2014) argue that deep

[4]Other studies that have also documented the persistence of cities include Eaton and Eckstein (1997), Bleakley and Lin (2012), and Jedwab et al. (2017).

[5]Zhu (2012) provides a comprehensive review of this literature, and proposes several policies (e.g., reforming financial sector) that can further improve the economic performance.

[6]Pomeranz (2000) argues that the easy access to coal supplies and ports (which led to the widespread use of steam engines) and the trade with the New World were the driving forces behind the divergence between Europe and China. Another view is that Western Europe's well-functioning markets operated under its inclusive and allocative institutions provided impetus for the industrial revolution (Acemoglu and Robinson 2014). However, Shiue and Keller (2007) argue that the performance of markets in China and Western Europe overall was comparable in the late eighteenth century. In a recent survey, Brandt et al. (2014) evaluate several explanations put forward for the great divergence, and consider institutions as a key factor for why China fell behind Europe.

historical roots surrounding China's present institutions and its past accumulation of skill have had a profound impact on recent Chinese development. Keller and Shiue (2007) show that the degree of integration of rice markets in the 1720s is a good predictor of income per capita in the 1990s. Similarly, Keller et al. (2013) analyze China's long-run trade performance, focusing on Shanghai, and find that the levels of present trade are strongly correlated with that in the 1870s. Our study complements these papers by showing that more prosperous cities of the 1770s still have higher living standards today.

The rest of the paper is organized as follows. The next section introduces a brief history of China, especially emphasizing the last two centuries. Section 3 discusses the data and provides summary statistics. Section 4 describes the econometric methodology that we employ, and presents the results. Section 5 concludes.

2.2 A Brief History

The Qing dynasty, the last imperial dynasty of China, was founded by the semi-nomadic Manchus from northeast of the Great Wall in 1644.[7] During the first half of their ruling period, the Manchus extended their rule over a vast area (covering Central Asia, Mongolia, and Tibet), and doubled the Ming dynasty's population, reaching 300 million or more by 1800 (Rowe 2010). The successful reigns of the Kangxi (1662–1722) and Qianlong (1736–1795) emperors display a period when progressive economic and social reforms are implemented. The empire went through a commercial revolution, in which interregional trade led to a rapid urbanization of rural areas and increased economic prosperity. Although foreign trade was severely regulated (e.g., all foreign trade coming into China was confined to Guangzhou) until the mid nineteenth century, production in agriculture, mining, and manufacturing increased,

[7] Rowe (2010) provides a comprehensive account of the Qing dynasty. Brandt et al. (2014) examine the long-run evolution of China's economy, and investigate roots of China's recent economic progress in the distant past. Their analysis not only covers the Qing dynasty, but also its predecessors the Song and Ming dynasties.

and the middle class expanded. Several new cities were created during this period (Rowe 2010). The average standard of living during the high point of the Qing dynasty (i.e., the second half of the 1700s) was likely higher than that in other parts of the world, including Western Europe (Pomeranz 2000).

By the early nineteenth century, the empire had been challenged and weakened by several factors such as rapidly growing population, limited reserves of food, deteriorated public infrastructure, outmoded industry policies, corruption of officials, and foreign incursions (Brandt et al. 2014). After emperor Jianqing's death in 1820, the problems exacerbated further, and the imperial government had to deal with multiple domestic and foreign adversaries until its collapse in 1911 (Naughton 2006). Beginning with the Opium War with Britain in 1839, China fought six major wars against foreign powers, and lost each of them. In addition, there were several internal uprisings such as the Taiping and the Boxer rebellions during the 1860s and 1890s.

Immediately after China became a republic in 1912, it collapsed into political instability and civil war created by rival military regimes. The warlord era continued until 1927 when the Nationalist Party unified the nation (Naughton 2006). China had relatively peaceful and prosperous years until Japan's invasion in 1937. By 1935, textile mills in China produced 8 percent of the world's cotton yarn (Brandt et al. 2014). Soon after the Sino-Japanese War (which continued until 1945), the country plunged into another civil war between nationalists and communists, resulting in the Communist Party's victory in 1949.

The Communist Party, under Mao Zedong's leadership, sought to implement a socialist *big-push* development strategy, where the government controlled the economy, owning all large factories, channeling investment toward heavy industry, mandating allocation of resources and output, and setting prices (Brandt et al. 2014). Although this strategy brought some success in early years, overall it was not a sustainable strategy and its short-term development plans sometimes ended up with tragic failures. During the Great Leap Forward (GLF), the big-push strategy intensified by transferring enormous amount of resources from agriculture to heavy industry. This created a serious shortage in food reserves, and the problem worsened when a full-blown famine hit in 1960, leaving about 25-30 million dead by the end of 1961 (Naughton 2006). Attempts to revive the economy after the GLF were short-lived because Mao purged anyone who was critical of his policies (Naughton 2006).

Deng Xiaoping took control of the Party in 1978, and gradually implemented reforms to improve economic conditions, moving the country towards a more market-oriented economy. Reforms included rural liberalization, introduction of a dual-track pricing system, restructuring of state-owned enterprises and privatization, and expanding trade and foreign investment (Chow 2007). In the late 1970s, China was one the poorest countries in the world, its per capita income was comparable to the Sub-Saharan African countries. Today, China has the second largest economy in the world, and in terms of per capita income it ranks among middle-income countries.

Current living standards, however, vary substantially across regions. More importantly, cross-regional development differences also persist. As we shall show below, more prosperous cities of the Qing empire are still enjoying higher living standards today. Thus, despite the substantial changes in economics, politics, and social order that China has gone through over the last two centuries, as briefly outlined above, the relative ranking of its regional development has remained mostly the same.

2.3 Data and Descriptive Statistics

Our analysis uses data on 227 cities from 24 provinces as shown in Figure 2.1, which cover about 40 percent of China's mainland, account for more than 80 percent of the total population, and generate about 90 percent of China's total output. Our sample size is determined by the availability of population data from the Qing dynasty.[8] The data on city population

[8] As shown in Figure 2.1, the population data are not available for all cities in these provinces. For example, we have data only from two cities in Hainan province.

in 1776 are from Cao (2001) who compiled from historical archives, and the 2010 population data are from the City Statistical Yearbook (2011). In matching city boundaries in 2010 with that in the Qing dynasty we use the concordance tables in Xue (2001), and some boundary information from Liang (2008). Cities whose boundaries changed significantly are excluded from our analysis. We also note that some cities over time split into several new cities (e.g., Ningxia now has 5 cities). In such cases, we combined all of the present-day cities to match with the parent city in the Qing dynasty. Cao (2001) also reports the 1820 population survey, which covers 233 cities. However, using the population density in 1820 yields qualitatively the same results.

We use the population density in 1776 as our measure of the early economic development in China. After identifying each city's boundary, we use the Gridded Population of the World (GPWv4, 2016) database developed by the Center for International Earth Science Information Network at Columbia University to measures surface areas in square kilometers. The land area obtained from this database excludes deserts and permanent ice/water areas. Dividing the population by the calculated land area, we obtain the number of people living in each square kilometer as a measure of population density.

This paper investigates the impact of the early development on today's living standards. The standard approach is to use income per capita as a proxy for prosperity. However, our preferred measure for the present living standards is night-light density measured from outer space during 2010. We use this proxy for several reasons. First, we have data on output produced (i.e. GDP), but we do not have income data. These two figures are highly correlated, but correlation is far from perfect at the sub-national level. For example, people may live in one city, but may work in another city. Second, city-level GDP data reported by the Chinese officials are subject to serious measurement errors (Young 2005). For instance, some of the economic activity (especially, in rural regions) is conducted in the informal sector, which is not fully counted by the official statistics. Finally, one needs PPP-like conversion rates to obtain a comparable production/income figures across regions.

Henderson et al. (2012) were the first to use night lights as a proxy to measure economic growth, especially using them to augment official income growth measures. We also use the average night-time light density during 2010 as a measure of development. Satellite data on night lights are from National Geophysical Data Center (NGDC).[9] Following Henderson et al. (2017), we use the radiance-calibrated version of the light data. The advantage of using this new dataset is that information about low light places are less distorted and all topcoding is removed.[10] Lights data are distributed as a grid of pixels (0.86 square kilometer), and the total amount of lights in each city is obtained by aggregating the light intensity in each grid across the land area. Dividing the light amount by the city's land area, we obtain the average light intensity.

We also investigate the impact of early development on the average GDP per capita in 2010 using Provincial Statistical Yearbook (2011). We use GDP data from 2010, but the analysis based on the data from any other available recent years yields very similar results. We also consider two other development indicators in 2010: the average years of schooling and trade openness. Data on the average years of schooling and trade are from the 2011 Population Census of China and the Monthly Custom Statistics (2011), respectively. All variables (population density, GDP per capita, average years of schooling, trade openness) are measured at city-level. In our regression analysis, we also use the average annual temperature and precipitation in each city over the 2000–2010 period as well as their variations as additional controls. We obtain these data from the China Meteorological Data Service Center.

[9]The NGDC is a part of the National Oceanic and Atmospheric Administration, which obtains the raw satellite data on night lights from the United States Airforce Defense Meteorological Satellite Program (DMSP) that have been recording the intensity of lights with their sensors.

[10]The results based on the earlier version of light data (which range from 0 to 63) are very similar.

Table 2.1: Summary Statistics on Key Variables

Province	Population Density 1776	Light Density 2010	GDP/Capita (1,000 Yuan) 2010	Average Schooling 2010	Trade/GDP (Percent) 2010
Anhui	194.00	5.61	21.51	8.26	12.84
	(152.22)	(4.94)	(14.47)	(0.78)	(11.81)
Beijing	122.16	32.38	112.21	11.71	144.57
	—	—	—	—	—
Chongqing	56.59	3.46	23.99	9.10	10.61
	—	—	—	—	—
Fujian	100.34	5.84	39.84	9.01	50.10
	(78.80)	(6.91)	(90.19)	(0.63)	(61.03)
Gansu	34.05	1.28	14.99	9.03	12.34
	(50.38)	(0.77)	(90.00)	(1.60)	(21.76)
Guangdong	105.97	11.30	49.22	9.61	103.45
	(74.28)	(15.96)	(33.66)	(0.98)	(88.39)
Guangxi	34.60	2.67	20.99	8.76	10.92
	(14.01)	(1.42)	(70.04)	(0.57)	(16.30)
Guizhou	32.15	1.87	13.51	7.61	5.60
	(18.65)	(1.66)	(56.96)	(0.88)	(6.85)
Hainan	231.20	16.61	30.24	10.42	32.77
	(143.55)	(0.72)	(28.17)	(0.43)	(27.49)
Hebei	98.98	11.35	28.24	9.40	13.56
	(70.81)	(6.78)	(14.05)	(0.37)	(7.10)
Henan	145.89	7.90	24.52	8.92	4.84
	(63.87)	(6.26)	(10.39)	(0.74)	(2.97)
Hubei	100.08	3.49	27.37	9.18	11.24
	(76.18)	(4.65)	(15.78)	(0.97)	(8.74)
Hunan	69.33	2.38	25.18	9.15	6.01
	(23.29)	(1.99)	(15.74)	(0.59)	(4.59)
Jiangsu	314.87	24.73	53.18	9.30	75.34
	(274.87)	(19.08)	(23.74)	(0.77)	(72.55)
Jiangxi	98.47	2.53	21.16	8.85	15.50
	(96.75)	(1.98)	(11.62)	(0.63)	(15.11)
Ningxia	26.10	4.97	26.69	8.09	7.85
	—	—	—	—	—

Table 2.2: Summary Statistics on Key Variables – Continued

Province	Population Density 1776	Light Denisty 2010	GDP/Capita (1,000 Yuan) 2010	Average Schooling 2010	Trade/GDP (Percent) 2010
Shaanxi	38.68	5.19	26.12	9.52	8.04
	(43.65)	(4.07)	(12.59)	(1.21)	(7.88)
Shandong	183.31	14.75	41.97	8.96	31.83
	(66.41)	(5.12)	(20.42)	(0.71)	(26.14)
Shanghai	339.77	63.12	74.63	10.73	145.47
	—	—	—	—	—
Shanxi	77.72	6.97	25.88	9.72	9.47
	(55.22)	(2.53)	(10.69)	(0.57)	(10.99)
Sichuan	48.84	3.33	19.97	8.55	13.12
	(40.33)	(4.56)	(12.52)	(0.72)	(13.14)
Tianjin	124.46	44.50	93.66	10.4	75.87
	—	—	—	—	—
Yunnan	20.59	2.40	16.11	7.76	12.22
	(11.17)	(1.75)	(89.36)	(0.83)	(14.03)
Zhejiang	220.18	14.75	49.67	8.78	63.47
	(154.91)	(10.80)	(14.90)	(0.67)	(28.26)
	84.02	5.91	33.25	9.04	50.14
	(100.29)	(8.57)	(23.43)	(0.98)	(61.50)

Notes: Satellite data on night lights are from National Geophysical Data Center. The data on population in 1776 are from Cao (2001) and Liang (2008); GDP, population, and the average years of schooling are from City and Provincial Statistical Yearbooks (2011) and the 2010 Population Census of China. Trade data from the Monthly Custom Statistics (2011).

Tables 2.1 and 2.2 report the summary statistics on these key variables across provinces (and numbers in parentheses are the standard deviations). Beijing, Chongqing, Ningxia, and Shanghai do not have any standard deviations, because each of these cities represents the whole province. GDP per capita is in 1000s of Yuan, and trade openness is measured by (Imports+Exports)/GDP and expressed in percent. Note that for each key variable, there is a substantial variation across provinces. Beijing and Shanghai are usually at the

Table 2.3: Correlation Among Key Variables

	Pop. Density	Light Density	GDP/Capita	Schooling	Openness
Pop Density	1.000	0.722	0.505	0.430	0.528
Light Density		1.000	0.740	0.578	0.654
GDP/Capita			1.000	0.715	0.624
Schooling				1.000	0.511
Openness					1.000

Notes: All variables are measured in logs.

top of the list in each variable, while Gansu, Ningxia, and Yunnan ranked at bottom. The development has not been uniform across cities. Beijing, which was relatively more densely populated in 1776, is now more bright, richer, more educated and open. Anhui was also densely populated in 1776, but now is less bright, poor, less educated, and less open. Table 2.3 reports correlation across these variables, and note that the population density in 1776 is positively correlated with all other indicators, and the correlation is especially strong with population density and light density in 2010. Note also that night-light density is strongly correlated with per capita income and openness.

2.4 Empirical Implementation

2.4.1 Econometric Specification

We index cities by c and provinces by p, and use the following model to assess the relation between the early development and the present one:

$$Y_c = \alpha_p + \beta \text{Popden}_{c,1776} + X_c + \varepsilon_c, \tag{2.1}$$

where Y_c denotes log value of the following development measures: the average night-light density, population density, GDP per capita, years of schooling, or trade openness in 2010. The dummy variable α_p equals one if city c is in province p and zero otherwise, $\text{Popden}_{c,1776}$ represents log population density in 1776, X_c is a set of control variables, and ε_c is the error term. The coefficient of interest is β.

Province fixed effects (α_p) are included to control for geographical differences across regions to some extent. However, as shown in Figure 2.1 and supported by standard deviations reported in Tables 2.1 and 2.2, there are still substantial geographic variation across cities within several provinces. Population density is likely to be higher in cities that are closer to the province's capital, where more public goods are provided. In addition, certain geographic characteristics that were not useful in the past may turn out to be beneficial later on, or vice versa. For example, until the mid-nineteenth century, foreign trade was severely regulated and all foreign trade coming into China was confined to one port in Guangzhou. During the last forty years, however, several new ports constructed, and the cities closer to these ports are more likely to develop due to international trade and investment. Omitting these factors will obviously bias our estimates. Therefore, we add a set of controls (denoted by X) which includes distance to the nearest port, distance to the capital city of each province, the mean levels of temperature (Celsius) and rainfall (mm) between 2000 and 2010, the corresponding average annual variations in temperature and precipitation, and a river dummy that equals one if one of the four major rivers (Huai, Xi, Yangzi, and Yellow) passes through the city.[11] We use log values of all continuous control variables.

We use heteroskedasticity-robust standard errors clustered at the province level. However, to minimize any potential problems in inference that may stem from a small number of clusters (24 provinces), we obtain $p-$values associated with a test of significance for each coefficient using the wild bootstrap t-procedure (developed by Cameron et al. 2008) clustered

[11] We do not include latitude in our regressions, since it is highly correlated with temperature and precipitation. Including latitude does not have any significant impact on the coefficient of $\text{Popden}_{c,1776}$. There are several other very small rivers running through China Proper, but the above rivers have been considered most important ones in China's development, especially, during the Qing Dynasty (Rowe 2010).

at the province-level with 100,000 replications. Results based on robust standard errors (i.e., not clustered) are qualitatively similar to those reported here.

2.4.2 Results

Before presenting our results about the impact of early development on today's living standards, it is interesting to investigate the degree of persistence in population density. The first two columns in Table 2.4 report the effect of the 1776 population density on that of 2010. All regressions include province fixed effects, and the p-values are shown in square brackets. The point estimate in column 1 is about 0.6 and statistically highly significant (at the 0.1-percent level), but the point estimate decreases by 20 percent when we include additional controls. According to column 2, had the population density been 1 percent higher in 1776, the average population density in each city would have been about 0.5 percent higher in 2010. The distance to the nearest port and the distance to province's capital have a negative effect on the population density, whereas average temperature has a positive effect, as expected. Other controls do not have any significant impact on the population density.[12]

Columns 3 and 4 in Table 2.4 report the results for our first development measure: night-light intensity. Column 3 shows the effect of early development on light density in 2010 without controls. The estimated coefficient on Popden_{1776} is about 0.7 and statistically highly significant. However, the estimate falls to 0.550 once we add the control variables. The point estimate in column 4 implies that had the population density been 1 percent higher in 1776, the average night-light density in each city would have been about 0.5 percent higher in 2010. Note that estimates on distance to the nearest port and distance to province's capital are negative and highly significant, whereas coefficients on other controls are insignificant.

[12]The estimated coefficient is less than 1, suggesting that there has been some convergence across the cities. To test this formally, we run equation (1), where the dependent variable is the average *annual* growth rate of the population density in city c. In this case, β measures the speed of convergence [if any] across cities. We find that $\hat{\beta} \approx -0.002$ with a p-value of 0.002, indicating a convergence in population density across cities.

Table 2.4: Impact of Population Density in 1776 on China's Development in 2010

	Popden_{2010}		Light Density		GDP per Capita		Schooling		Trade Openness	
Variable	1	2	3	4	5	6	7	8	9	10
Popden_{1776}	0.610***	0.491***	0.697***	0.550***	0.174**	0.125*	0.042***	0.024**	0.539***	0.315***
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.038]	[0.067]	[0.007]	[0.023]	[0.002]	[0.005]
Dist. to Port		−0.132*		−0.245***		−0.154**		−0.016*		−0.362***
p-value		[0.088]		[0.000]		[0.012]		[0.093]		[0.004]
Dist. to Cap.		−0.075***		−0.157***		−0.086***		−0.026***		−0.193***
p-value		[0.001]		[0.000]		[0.000]		[0.000]		[0.001]
Temp avg		1.146**		−0.090		−0.656**		−0.137*		1.370*
p-value		[0.017]		[0.771]		[0.043]		[0.086]		[0.081]
Temp sdv		−0.371		−0.604		−0.267		−0.022		−0.550
p-value		[0.661]		[0.473]		[0.160]		[0.426]		[0.429]
Rainfall avg		0.101		−0.140		−0.376		−0.106		−1.490
p-value		[0.859]		[0.775]		[0.428]		[0.551]		[0.309]
Rainfall sdv		−0.052		−0.108		−0.148		0.034		0.389
p-value		[0.859]		[0.729]		[0.471]		[0.524]		[0.434]
River		−0.041		0.070		0.189*		0.019		−0.143
p-value		[0.512]		[0.522]		[0.054]		[0.109]		[0.534]
Observations	227	227	227	227	227	227	227	227	225	225
Adj. R^2	0.740	0.778	0.745	0.772	0.485	0.645	0.519	0.755	0.577	0.651

Notes: All regressions include province fixed effects. All continuous variables (including dependent variable) are measured in logs. Numbers in square brackets are p-values obtained from wild bootstrapping (100,000 replications) clustered at the province level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

In columns 5 and 6, we repeat the same exercise where the dependent variable is log GDP per capita. Note that the coefficient on the Popden_{1776} in column 5 becomes smaller and less significant once the controls are included: the effect of early development decreases from about 0.175 to 0.125. Thus, had the population density been 1 percent higher in 1776, the average per capita GDP in each city would have been about 0.13 percent higher in 2010. Consistent with expectation, the distance to the nearest port, the distance to province's capital, and average temperature have a negative effect on GDP per capita, while access to a river has a positive effect. Note that our regressions do not include schooling and openness, because they are outcome variables, and hence considered bad controls (Angrist and Pischke 2009).

Columns 7 and 8 report results from regressions where the dependent variable is the log average years of schooling in 2010. Including controls reduces the estimated coefficient on Popden_{1776} from 0.042 to 0.024, and the latter estimate is significant only at the 5-percent level. The estimate in column 8 implies that if the population density were 10 percent higher in 1776, the average years of schooling in 2010 would have been 0.25 percent higher, which further implies that the average years of schooling in 2010 would have been about 9 years (since it is about 8.95 years). Estimates in column 8 indicate that the average years schooling is higher in places closer to a port and a capital.

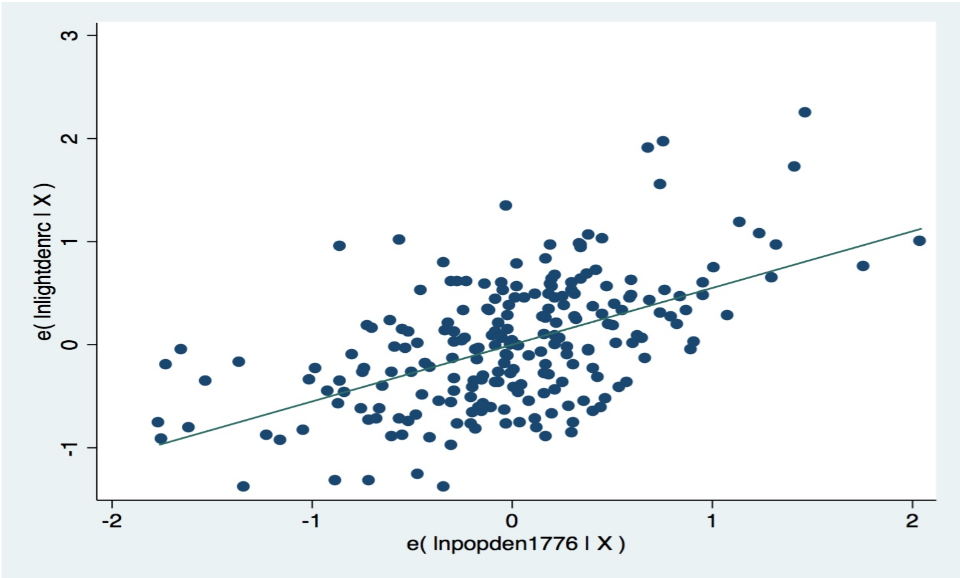
Finally, we also investigate implications of early development for trade openness. According to columns 9 and 10 in Table 2.4, early development has a positive and highly significant effect on trade openness in 2010. Note that adding controls to our model reduces the coefficient of interest more than 40 percent. This is not surprising because cities closer to coast are more likely to be exposed to international trade. According to column 10, if the population density were 1 percent higher in 1776, 2010 trade openness would have been 0.315 percent higher. We also consider specifications where export/GDP and import/GDP are separately regressed on Popden_{1776} and controls. Estimated coefficients are 0.320^{**} [0.043] and 0.431^{**} [0.025] for exports and imports, respectively.

In each specification, the estimated coefficient on Popden_{1776} falls by at least 20 percent when we include additional controls. Estimates in Table 2.4 suggest and our further analysis confirms that the reduction mainly stems from including distance to the nearest port and the distance to province's capital into regressions. Intuitively, people prefer to live in cities that are closer to capital so that they can benefit from more public goods served there, and the cities closer to the ports benefit from the economic prosperity generated from there. Thus, excluding these variables from our regressions erroneously would assign their impact on the current development to the population density in 1776.

To check whether the results are driven by any particularly influential observations, we plot the partial regression results obtained from model (1) for each outcome variable. Figure 2.2, 2.3, 2.4 and 2.5 represent these plots. Note that they do not show any apparent outliers. Our sample includes five provinces each having only one city: Beijing, Chongqing, Ningxia, Shanghai, and Tianjin. We repeat the analysis without these cities, and our point estimates and their significance are almost identical to those reported in Table 2.4. As a further robustness check, we also consider the impact on these development indicators measured in 2000. In measuring night lights in 2000, we use an older version where density ranges between 0 and 63. Results based on the year 2000 are mostly the same as those in Table 2.4, and are available upon request. In sum, our analysis implies that more densely populated cities in the 1776 Qing dynasty are still more densely populated and also more developed today in the sense that they are brighter, richer, more educated, and more open.

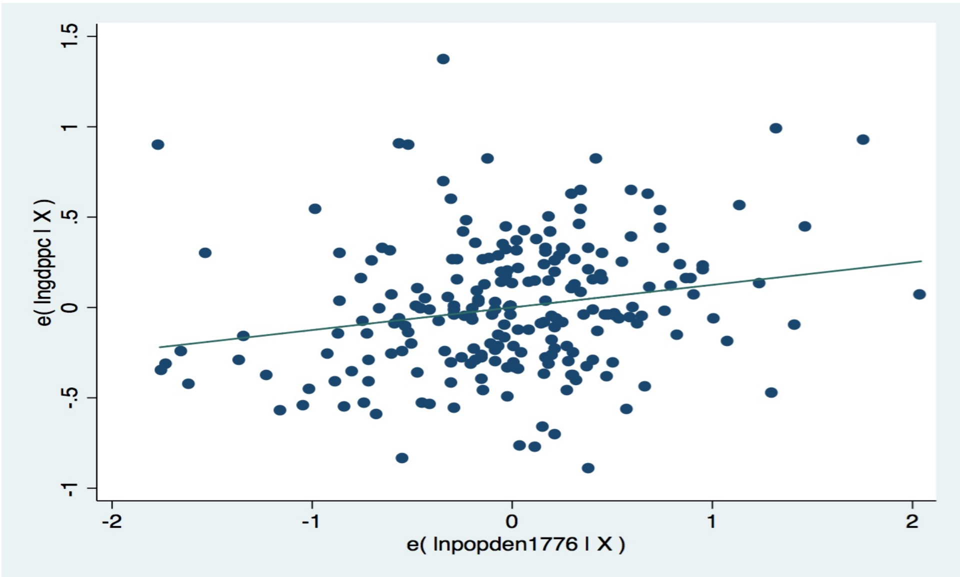
2.5 Conclusion

This paper sheds light on our understanding of the substantial variation in across-region living standards in today's China. Using insights from the recent empirical studies in the comparative development literature, the paper argues that early development is a strong and robust predictor of today's living standards. Using historical records on populations from



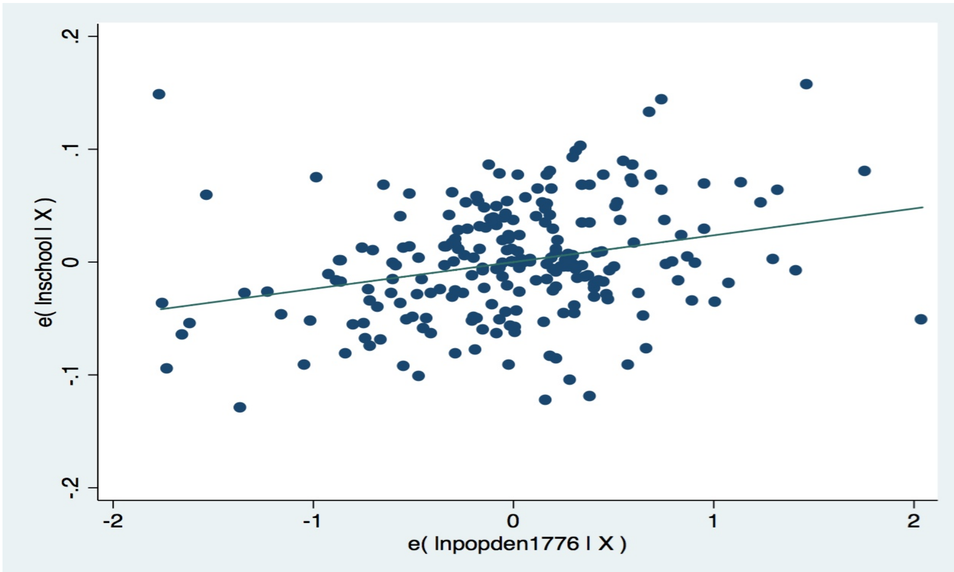
Note: X includes all controls and province fixed effects in equation (1).

Figure 2.2: Partial Regression Plots for Each Outcome: Night-Light Density



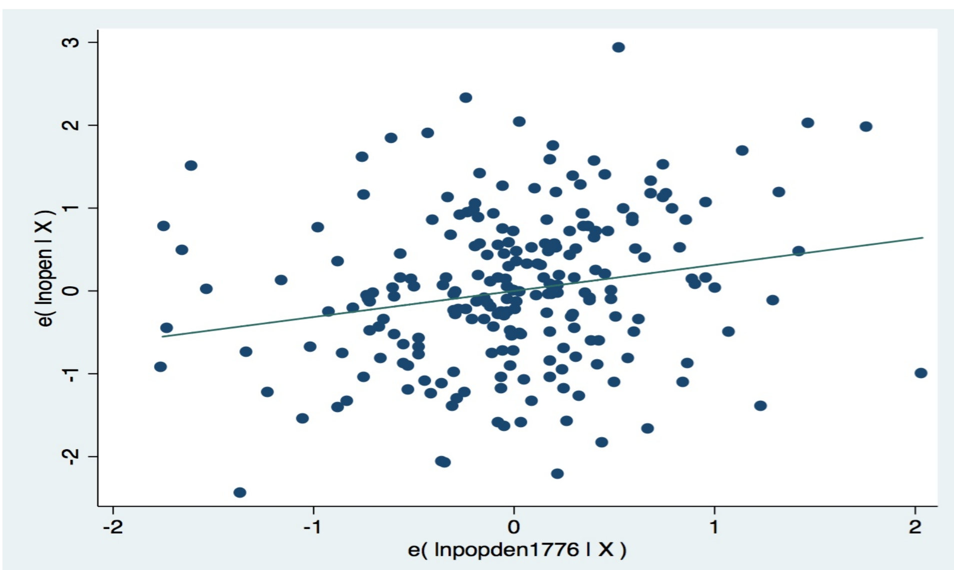
Note: X includes all controls and province fixed effects in equation (1).

Figure 2.3: Partial Regression Plots for Each Outcome: GDP per Capita



Note: X includes all controls and province fixed effects in equation (1).

Figure 2.4: Partial Regression Plots for Each Outcome: Schooling



Note: X includes all controls and province fixed effects in equation (1).

Figure 2.5: Partial Regression Plots for Each Outcome: Trade Openness

Cao (2000 & 2001), we calculated the population density of about 230 cities in 1776 from the Qing dynasty, and use it as a proxy for the early development. We then investigated how the current development level in these cities were associated with their early development, and found that more prosperous cities in 1776 are still enjoying better living standards today in the sense that they are brighter, richer, more educated, and more open to international trade. Our results are robust to the choice of control variables and years.

Chapter 3. Effects of Innovation and Training on Export Decision in China

3.1 Introduction

A large empirical literature shows the existence of large differences across firms/establishments even in the same narrowly defined sectors. It documents that firms differ in terms of output produced, labor productivity, types of worker employed, wages paid, investment in R&D, and export status. Some of these studies have further shown that exporters are better than non-exporters: exporters are larger, more productive, more capital-intensive, employ more able workers, pay higher wages, and invest in technology (Bernard and Jensen 1999, Bustos 2011).[1] Several plant- and firm-level studies have found evidence of self-selection into export markets, where only better firms export.

This paper investigates the determinants of export status of Chinese manufacturing firms over the 2003 – 2011 period. This paper complements the existing literature by considering two additional important factors that can affect the decision to export. Previous studies show that firms investing in technology (i.e., R&D spending) are more likely to export, suggesting that technologically more advanced firms are more likely to *begin* exporting. This paper uses a direct measure of technology: output value of new products introduced in each year. This paper also uses data on spending on workers' training, a proxy for investment in human capital. I investigate how these factors affect the likelihood of exporting as well as how these factors are important for firms' first-time decision to exit from export markets. Following Bernard and Jensen (1999), I consider the determinants of first-time entry to and first-time exit from export markets. This approach mitigates the reverse causality problem in the present setting. Investigating determinants of export decision in a dynamic model is left for a future research.

[1] The literature on this subject is vast. Bernard et al. 1997, Melitz and Redding 2014 provide comprehensive and complementary reviews of this literature.

The findings of this study can be summarized as follows. First, technology level and job training have a positive and significant impact on the likelihood of beginning to export. Specifically, the likelihood of an innovative firm to export is about five percentage points higher than that of non-innovative firms. Similarly, compared to firms that do not invest in workers' training, the likelihood of a firm that invests in job training to export is 0.4 percentage points higher. Job training and technology level have a negative and significant impact of likelihood of exiting from export markets as well. The likelihood of an innovative firm to exit from exporting is about 3.6 percentage points lower than that of non-innovative firms, and the likelihood of a firm that invests in job training to exit from export markets is about 0.4 percentage points lower. Second, investigating across different ownership, these effects are generally more substantial among private firms (including foreign affiliates). Finally, exploring heterogeneity across industries reveals that the impact of innovation and job training on export status is dominant among high-tech industries: they significantly increase (decrease) the likelihood of entering to (exiting from) export markets. Estimates on other controls are consistent with previous studies. For example, firms with higher market shares are more likely to export and less likely to exit from export markets.

This paper methodologically builds on Berand and Jensen (1999), who examine whether good firms become exporters or exporting improves firm performance. Using firm-level data from the U.S. manufacturing sector over the 1984 – 1992 period, they find that good firms become exporters, but the benefits of exporting are less clear. Specifically, Bernard and Jensen find that firm size and wages increase the probability of exporting. Bernard and Jensen (2004) extend their analysis to address determinants of entry and exit in the export market by U.S. plants. They find that plant characteristics (such as size, productivity, wages paid) strongly increase (decrease) the probability of entry into (exit from) exporting. Other studies have investigated additional factors that can affect firm export status, including entry costs (Robert and Tybout 1997, Clerides et al. 1998) and investment in technology (Aw et al. 2008 & 2011). This analysis extends these studies by focusing on two new factors that can affect firms' export status: innovation capacity and investment in human capital in the form of job training.

Previous studies have used R&D spending as a proxy for technology, arguing that investment in R&D implies higher productivity and more innovation, which in turn increases firms' likelihood of exporting. There is a large literature that studies the importance of R&D for international trade (see Grossman and Helpman 1991 for an earlier account). This paper measures the technology capacity of firms with the output value of new goods produced. I present results using R&D spending, but (to the best of my knowledge) this is the first paper that uses firm innovation level as a determinant for its export status. Using output of new products as a proxy for the firm's technology level has two main advantages. First, unlike R&D spending (which is an investment to technology production), I directly measure firms' technology levels. Second, R&D data are often subject to serious measurement errors. Firms wishing to evade taxes and/or get subsidies for their investment in technology have incentives to report higher R&D spending.

These findings are consistent with the predictions of a large theoretical literature that has studied the role of technology and human capital on productivity and trade. The hallmark of endogenous growth theory is that profit seeking entrepreneurs use human capital and existing technology to introduce new goods (Romer 1990, Grossman and Helpman 1991). These insights, combined with the recent literature on firm heterogeneity, imply that firms investing in technology and human capital are more likely to export. These models also emphasize the dynamic interaction between these factors and trade. For example, Atkinson and Burstein (2011), Bustos (2011), and Unel (2013), among many others show that technologically advanced firms will self-select into export markets, and export in turn induces firms to adopt better technologies. Unel (2015) and Dinopoulos and Unel (2017) show the similar dynamic effect between human capital and trade. The present paper mainly focuses on how these factors affect entry into and exit from foreign markets. Exploration of how trade affects firm investment in technology and human capital is left for a future work.

This paper also relates to a large body of literature that investigates the factors that affect Chinese firms' trade decisions. China is among the world's leading countries in international trade and has grown dramatically since the implementation of market-oriented reforms in the late 1970s. According to World Trade Organization (WTO), China's total merchandise trade has increased by more than 180 times in nominal value since the 1980s. When exploiting factors that have accelerated China's export, productivity and efficiency advancement are both considered to be crucial (Manova 2010, Fan et al. 2015, Egger and Keuschnigg, 2017). The Chinese research and development expenditure is only second to that of the United States since the 2000s. Reports by World Intellectual Property Organization (WIPO) reveal that there are particular interests in encouraging productivity improvement and R&D activities in China. The total research and development expenditure in China has grown from about 0.56% of GDP in 1996 to approximately 2.07% of GDP in 2015 (World Bank 2017). Thus, investigating how technology has affected the decision to export is important.

The rest of the paper is organized as follows. The next section introduces the data employed in my analysis and provides summary statistics. Section 3 describes the econometric methodology. Section 4 presents the results. This section also explores the heterogeneity across firms based on their ownership and industry. Section 5 presents analysis using R&D investment as a proxy for firm technology, and Section 6 concludes.

3.2 Data

This paper uses a large Chinese firm-level dataset that cover all manufacturing firms either are state-owned enterprises or with an annual revenue of at least five million Chinese Yuan between 2003 and 2011. The dataset (also known as *China Annual Survey of Industrial Firms*) is compiled by China's National Bureau of Statistics. The survey has information about firms' location, industry (at the 4-digit level), age, ownership structure, output produced, revenues from sales, number of workers, book-value of capital, R&D spending, and export. A novel feature of this survey is that it includes data on output of *new* products and annual spending on workers' training. Thus, unlike previous studies that use R&D spending as a proxy for technology, it is possible to directly measure technology output from the data, and relate it to firm export status. Spending on job training is used as a direct measure of investment in human capital. The industry-level Producer Price Index (PPI) from China's National Bureau of Statistics is used to convert nominal values to real values.

The original dataset contains more than 2.8 million observations. However, some observations have serious measurement errors stemming from self-reporting and weak quality control in early survey years (Cai and Liu 2009; Nie, Jia, and Yang 2012), and thus one needs to clean them before conducting any analysis. Dropping all observations with non-unique firm ID, missing location or industry classification, and low reliability (privately-owned firm sales less than five million Chinese Yuan, or negative R&D spending, etc.). All observations with measurement errors (e.g., the total value of net fixed assets is less than the total value of assets) or only one year of observations are dropped. This cleaning process results in about two million observations. About 50% of firms provide four or more years of observation in this study. Thus, the dataset is an unbalanced panel.

Table 3.1 provides the summary statistics on key variables. The first three rows show the fraction of firms that export, produce new products, provide on-the-job training, and invest in R&D. The last row shows the relative market share (measured by its revenue relative to the industry's average revenue). As a comparison with SOEs[2], non-SOEs[3] are more likely to export, but less likely to introduce new products, invest in R&D, and provide job training. Table 3.2 provides information about exporting, the value of new products, spending on job training, and R&D expenditure as a fraction of total output. Similar to Table 3.1, SOEs invest more in job training, R&D, and produce more new products. Tables 3.3 and 3.4 give similar statistics for exporters and non-exporters. Consistent with several other studies

[2]SOEs include state-owned enterprises and collective-controlled enterprises in this paper.

[3]Non-SOEs include private-controlled enterprises, Hong Kong-Macao-Taiwan resident-controlled enterprises and foreign-controlled enterprises in this paper.

Table 3.1: Summary Statistics on Key Variables – Share in Total Number of Firms

	All	Non-SOE	SOE
Export	0.264	0.276	0.189
New Product	0.113	0.071	0.182
Job Training	0.263	0.279	0.356
R&D	0.070	0.071	0.122
Market Share	1.000	0.938	1.518

(Melitz and Redding 2014), exporters are more innovative, invest more in R&D, and have higher market shares across all groups.

Table 3.2: Summary Statistics on Key Variables – Share in Total Output

	All		Non-SOE		SOE	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Export	0.1410	2.7364	0.1371	0.3729	0.0765	8.4040
New Product	0.0554	0.2030	0.0170	0.0936	0.0847	0.2398
Training	0.0015	0.0111	0.0014	0.0103	0.0033	0.0174
R&D	0.0004	0.0027	0.0003	0.0021	0.0008	0.0054

Table 3.3: Summary Statistics on Key Variables – Exporters

	All		Non-SOE		SOE	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
New Product	0.0689	0.2125	0.0316	0.1237	0.1445	0.2752
Job Training	0.0004	0.0019	0.0003	0.0017	0.0006	0.0023
R&D	0.0021	0.0114	0.0019	0.0106	0.0065	0.0203
Market Share	1.7099	6.7353	1.5465	6.5535	3.9775	12.1438

Table 3.4: Summary Statistics on Key Variables – Non-exporters

	All		Non-SOE		SOE	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
New Product	0.0504	0.1991	0.0114	0.0785	0.0707	0.2285
Job Training	0.0004	0.0029	0.0003	0.0023	0.0008	0.0059
R&D	0.0013	0.0110	0.0012	0.0101	0.0026	0.0166
Market Share	0.7371	1.9715	0.7063	1.9581	0.9404	2.5353

3.3 The Empirical Model

I investigate the factors that affect the likelihood of entering into or exiting from export markets by estimating the following linear probability model:

$$Y_{ijct} = \beta_n \mathrm{NP}_{it} + \beta_h \mathrm{JT}_{it-1} + \theta X_{it} + \eta_j + \eta_c + \eta_t + \varepsilon_{it}, \tag{3.1}$$

where i denotes firm, j denotes industry, c denotes city, and t denotes year. Y_{it} is a dummy variable that equals one if firm i exports for the first time in year t, zero otherwise. All observations after the first year of export are dropped. When I address the determinants of exiting from export markets, Y_{it} equals one if firm i stop exporting, zero otherwise. Similarly, all observations after the first year exiting from exporting are dropped. This analysis covers the 2003 – 2011 period and thus does not include firms founded prior. In this case, defining exporters and non-exporters is more complex. More details about defining these first-time export and exit indicators are provided in appendix.

Variable NP represents a dummy variable that equals one if the firm produces new products, zero otherwise. Similarly, JT_{it-1} is an indicator variable that equals one if firm i provides job training, zero otherwise. Results are also presented when these variables are continuous: the share of new products in total output, and total spending on job training normalized by the total output. Results based on the latter specifications are qualitatively similar to that based on the former ones. Note the use of the lagged value of JT, because it takes time for job training to impact production.

Variable X represents the set of controls, including the firm's market share (measured by its revenue relative to the industry's average revenue), its age, and its owners. There are five different types of ownership: state-owned enterprises, collective enterprises, private enterprises, enterprises controlled by residents in Hong Kong-Macao-Taiwan, and foreign-controlled enterprises. In some specifications, this analysis focuses on only one type of ownership (e.g., private and foreign controlled enterprises).

I include city and industry fixed effects (η_c and η_j) to control for any time invariant city and industry specific factors that can affect export decision, and year fixed effects (η_t) to control for common shocks to economies. The year fixed effects are at the provincial level, i.e. each province in each year has a different fixed effect. Finally, ε_{it} denotes the error term. Heteroskedasticity robust standard errors are clustered at the 4-digit industry-level to mitigate the potential serial correlation in the error term (Bertrand et al. 2004). Clustering data at the city-level yields qualitatively similar results.

3.4 Results

This section reports results based on equation (3.1). Section 3.1 presents results using the full sample. Results show that firms producing new products and investing in their workers' training are more likely to export, and less likely to quit from exporting. Section 3.2 restricts analysis to using only non-SOEs to see how much results change. Finally, in Section 3.3, I investigate heterogeneity across industries.

3.4.1 Using the Full Sample

Table 3.5 shows the impact of innovation (measured by introduction of new products) and job training on the likelihood of selling goods in foreign markets (i.e., exporting). As mentioned in the previous section, the sample only considers the first-time exporters. In addition, NP (New Products) and JT (Job Training) are indicator variables. Estimates on them are all positive and highly statistically significant. According to the last column, the likelihood of an innovative firm to export is about five percentage points higher than that of non-innovative firms. Similarly, compared to firms that do not invest in workers' training, the likelihood of a firm both exporting and investing in job training is 0.4 percentage points higher. Estimates on controls are mostly consistent with previous studies. Bigger firms (measured by their market shares) are more likely to export, as foreign owned or Hong Kong-Macao-Taiwan based firms do. Older, collective-owned firms are less likely to export. Private firms (relative to SOEs) are less likely to begin exporting.

Table 3.6 reports results when the dependent variable equals one if firm exits from exporting, zero otherwise. Once a firm stops exporting, the subsequent years are excluded from the analysis. Thus, it investigates the determinants of exiting foreign markets for the first time. Estimated coefficients on new products are negative and highly significant, i.e. innovative firms are less likely to exit from foreign markets. The likelihood of an innovative firm to exit from exporting is about 3.6 percentage points lower than that of non-innovative firms. Similarly, compared to firms that do not invest in human capital, firms that invest in job training are less likely to exit from foreign markets – the likelihood of a firm that invests in job training to stop exporting is about 0.4 percentage points lower. Estimated coefficients on controls imply that bigger firms and foreign-owned firms (including ownership from Hong Kong-Macao-Taiwan) are less likely to exit from exporting.

Table 3.5: Determinants of Exporting (NP, JT: Discrete Variables)

	1	2	3	4
New Product (NP)	0.1472***	0.0566***	0.0506***	0.0511***
	(0.0067)	(0.0032)	(0.0031)	(0.0031)
Job Training (JT)		0.0064***	0.0038***	0.0043***
		(0.0007)	(0.0007)	(0.0007)
Market Share			0.0111***	0.0095***
			(0.0006)	(0.0006)
Age			−0.0006***	−0.0005***
			(0.0001)	(0.0001)
Collective				−0.0107***
				(0.0014)
Private				−0.0035**
				(0.0015)
HK-Macao-TW				0.0359***
				(0.0033)
Foreign				0.0589***
				(0.0040)
Adj. R^2	0.3630	0.0467	0.0542	0.0617
Obs	527,093	269,172	269,172	269,172

Notes: All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table 3.6: Determinants of Exiting from Exporting (NP, JT: Discrete Variables)

	1	2	3	4
New Product (NP)	-0.0218^{***}	-0.0423^{***}	-0.0331^{***}	-0.0357^{***}
	(0.0011)	(0.0021)	(0.0020)	(0.0021)
Job Training (JT)		-0.0117^{***}	-0.0066^{***}	-0.0080^{***}
		(0.0010)	(0.0009)	(0.0010)
Market Share			-0.0081^{***}	-0.0071^{***}
			(0.0005)	(0.0005)
Age			-0.0009^{***}	-0.0010^{***}
			(0.0001)	(0.0002)
Collective				0.0036
				(0.0038)
Private				0.0059
				(0.0040)
HK-Macao-TW				-0.0225^{***}
				(0.0052)
Foreign				-0.0210^{***}
				(0.0060)
Adj. R^2	0.3191	0.3360	0.3395	0.3419
Obs	371,754	183,164	183,161	183,161

Notes: All regressions include city fixed effects, 4-digit industry fixed effects, year fixed effects and province-year fixed effects. Numbers in parenthesis are standard deviation, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

In Tables 3.7, new products (NP) and job training (JT) are modeled as continuous variables. Here, NP represents the output value of new products normalized by the total output produced and JT represents the fraction of output spent on training of workers. Results related to new products in these tables are consistent with those in Tables 3.5 and 3.6. Although spending on job training has a positive and significant impact on entry to the export market, it has no significant effects on exiting from the export market. Note that estimates on this variable have large standard errors, suggesting a significant amount of variation/noise in this way of measurement. In sum, firms that are more innovative and invest more in workers' training are more likely to export, and less likely to exit from foreign markets.

Table 3.7: Determinants of Export Status (NP, JT: Continuous Variables)

	Entry	Exit
New Product (NP)	0.0696***	−0.0519***
	(0.0065)	(0.0034)
Job Training (JT)	0.2634***	0.0345
	(0.0888)	(0.3065)
Market Share	0.0107***	−0.0088***
	(0.0006)	(0.0005)
Age	−0.0005***	−0.0011***
	(0.0001)	(0.0002)
Collective	−0.0128***	0.0085**
	(0.0015)	(0.0039)
Private	−0.0056***	0.0112***
	(0.0015)	(0.0042)
HK-Macao-TW	0.0335***	−0.0155***
	(0.0032)	(0.0053)
Foreign	0.0565***	−0.0142**
	(0.0039)	(0.0062)
Adj. R^2	0.0581	0.3399
Obs	269,166	183,160

Notes: All regressions include city fixed effects, 4-digit industry fixed effects, year fixed effects and province-year fixed effects. Numbers in parenthesis are standard deviation, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

3.4.2 Heterogeneity by Ownership

The benchmark results reported in the previous section include all observations. However, according to Tables 3.1, 3.2, 3.3 and 3.4, non-SOEs and SOEs differ from each other substantially. In addition, data on SOEs are usually subject to measurement errors (Young 2003). In this section, I conduct the same analysis by considering non-SOEs and SOEs separately. The former set includes private enterprises, foreign-controlled enterprises, and those controlled by residents of Hong Kong, Macao, and Taiwan, while the latter includes collectives and state-owned enterprises.

Table 3.8: Determinants of Export Status, Non-SOE

	Discrete		Continuous	
	Entry	Exit	Entry	Exit
	1	2	3	4
New Product (NP)	0.0529***	−0.0350***	0.0736***	−0.0515***
	(0.0036)	(0.0022)	(0.0075)	(0.0036)
Job Training (JT)	0.0049***	−0.0070***	0.3796**	0.1560
	(0.0008)	(0.0009)	(0.1716)	(0.3297)
Market Share	0.0112***	−0.0068***	0.0123***	−0.0083***
	(0.0007)	(0.0006)	(0.0008)	(0.0006)
Age	−0.0006***	−0.0012***	−0.0005***	−0.0013***
	(0.0001)	(0.0004)	(0.0001)	(0.0004)
HK-Macao-TW	0.0382***	−0.0277***	0.0380***	−0.0261***
	(0.0030)	(0.0023)	(0.0030)	(0.0022)
Foreign	0.0609***	−0.0268***	0.0607***	−0.0254***
	(0.0036)	(0.0032)	(0.0036)	(0.0032)
Adj. R^2	0.0659	0.3318	0.0627	0.3299
Obs	215,367	167,571	215,367	167,570

Table 3.9: Determinants of Export Status, SOE

	Discrete		Continuous	
	Entry	Exit	Entry	Exit
	1	2	3	4
New Product (NP)	0.0467***	−0.0462***	0.0577***	−0.0566***
	(0.0043)	(0.0047)	(0.0087)	(0.0095)
Job Training (JT)	0.0014	−0.0185***	0.0823*	−1.6776**
	(0.0009)	(0.0041)	(0.0493)	(0.7372)
Market Share	0.0061***	−0.0091***	0.0073***	−0.0125***
	(0.0005)	(0.0013)	(0.0006)	(0.0013)
Age	−0.0003***	−0.0005***	−0.0003***	−0.0006***
	(0.0000)	(0.0001)	(0.0000)	(0.0001)
Collective	−0.0061***	0.0042	0.0079***	0.0116***
	(0.0012)	(0.0042)	(0.0013)	(0.0042)
Adj. R^2	0.0363	0.4471	0.0292	0.4429
Obs	53,792	15,554	53,792	15,554

Notes: All regressions include city fixed effects, 4-digit industry fixed effects, year fixed effects and province-year fixed effects. Numbers in parenthesis are standard deviation, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table 3.8 reports results for non-SOEs. In columns 1 and 2, new products and job training are modeled as discrete variables; in the last two columns they are modeled as continuous variables. Estimates and their significance in columns 1 and 2 are very similar to those reported in Tables 3.5 and 3.7. Except for job training, the other estimates in the last two columns are similar to those in Tables 3.6 and 3.7. Note that estimated coefficient on JT in the last column is positive but highly insignificant, suggesting substantial variation across observations. In sum, private firms that are innovative and invest in workers' training are more likely to begin exporting, and less likely to quit exporting.

Table 3.9 reports results for SOEs. New products have a positive and highly significant impact on entry to the export market, and note that point estimates are smaller than the corresponding estimates in Table 3.8. The impact of job training on entry is positive, but not always significant, which may stem from relatively small sample size. In addition, the estimated coefficient on JT is small compared to that in Table 3.8. According to columns 2 and 4, estimated coefficients on NP and JT are negative and statistically significant. The analysis shows that SOEs that are innovative and invest in workers' training are more likely to begin exporting, and less likely to exit from exporting.

3.4.3 Heterogeneity by Industry

This section explores heterogeneity across industries. To this end, firms are grouped into four broad industries as specified in Table 3.10. Industries 1 and 2 are considered low-tech industries with low capital intensity, whereas industries 3 and 4 are high-tech with high capital intensity. These industries are constructed based on the *Industrial Classification for National Economic Activities in China* (2011). As mentioned earlier, only manufacturing industries are included.

In Table 3.11, new products and job training are modeled as discrete variables. Columns 1 and 2 report results using only data from industry 1. Estimated coefficients on NP are positive and significant for entry to export market, and negative and significant for exit from foreign markets. That is, innovative firms are more likely to export and less likely to exit from foreign markets. Compared to Tables 3.5 and 3.6, the estimated coefficient on NP for entry (exit) is substantially smaller (larger in absolute terms). Although job training has a negative and significant impact on the exit from foreign markets, its impact on entry is positive and insignificant. Estimated coefficients on control are mostly the same as those reported in Tables 3.5 and 3.6.

Columns 3 and 4 report the results using industry 2. New products and job training

Table 3.10: Industries

Industry	Sub-industry
Industry 1	Agricultural products and food manufacturing, Beverage manufacturing, Tobacco manufacturing
Industry 2	Textile manufacturing, Clothing manufacturing, Leather and feather products manufacturing, Lumber and bamboo products manufacturing, Furniture manufacturing, Paper products manufacturing
Industry 3	Chemical product manufacturing, Pharmaceutic manufacturing, Rubber and plastic products manufacturing, Metal products manufacturing
Industry 4	General equipment manufacturing, Special-purpose equipment manufacturing, Stationary manufacturing, Automobile manufacturing, Other transportation manufacturing, Computer manufacturing

have a significant and positive effect on the decision to export, and dissuade firms from abandoning exporting. The impact of NP on exporting in industry 2 is the strongest among all four industries, while its influence on exit is the weakest. These results show that the introduction of new products in industry 2 could help firms become exporters, however, its usefulness in helping firms stay in the international market is relatively limited compared to other industries. A possible explanation for this finding is that industry 2 is more likely to roll out new products because the textile and furniture manufacturing industries change designs more frequently at lower cost, regardless of whether firms are exporting.

Columns 5 and 6 show the estimates using industry 3, whereas columns 7 and 8 report the results using industry 4. These industries are the high-tech sectors with a high capital intensity. The results report that new products help firms to begin exporting by more than four percentage points, while it can significantly lower the likelihood to quit exporting by about 3.5 percentage points in industry 3. JT demonstrates considerably more impact on the export decision in both industries. The estimated coefficients on JT in columns 5 to 8 show that spending on job training has a substantial impact on encouraging exporting and discouraging exit from foreign markets. The estimated coefficients on entry are about six times of that in columns 1 and 3. These findings are not surprising given that high-tech industries use more sophisticated technologies that demand more human capital.

I implement the same exercises using NP and JT as continuous variables, and results are given in Table 3.12. Estimated coefficients on NP for entry and exit are consistent with those reported in Tables 3.7 and 3.11. That is, innovative firms are more likely to export and less likely to exit from foreign markets. Th estimates on JT, however, are less precise. In industry 2 and 4, there is a positive and significant impact on entry to exporting, whereas in the other two industries the impact is insignificant. The impact on exit from foreign markets is usually insignificant. Note that when the estimates are insignificant (as is the case for exit), they are also less precisely estimated (i.e., standard errors are high). This suggests substantial variation across observations.

In sum, the impact of innovation and job training on exporting varies substantially across industries. Innovative firms are more likely to export and less likely to stop exporting, and the likelihood is substantially higher in high-tech industries. Job training turns out to be more important in high-tech industries, i.e. high-tech firms that invest in workers' training are more likely to export and less likely to exit from export markets. Estimates on controls are mostly consistent with benchmark results. Firm size is positively correlated with entry and negatively with the exit. Firms with ties to Hong Kong, Macao, Taiwan or foreign countries are also more likely to begin to export and less likely to exit from foreign markets.

Table 3.11: Determinants of Export Status (NP, JT: Discrete Variables), Industry

	Industry 1		Industry 2		Industry 3		Industry 4	
	Entry	Exit	Entry	Exit	Entry	Exit	Entry	Exit
	1	2	3	4	5	6	7	8
New Product (NP)	0.0357***	−0.0807***	0.0642***	−0.0288***	0.0408***	−0.0345***	0.0546***	−0.0301***
	(0.0049)	(0.0189)	(0.0107)	(0.0043)	(0.0067)	(0.0044)	(0.0045)	(0.0024)
Job Training (JT)	0.0008	−0.0127***	0.0080***	−0.0054***	0.0046***	−0.0090***	0.0058***	−0.0092***
	(0.0023)	(0.0035)	(0.0024)	(0.0020)	(0.0014)	(0.0031)	(0.0012)	(0.0016)
Market Share	0.0034***	−0.0079***	0.0154***	−0.0084***	0.0082***	−0.0075***	0.0124***	−0.0054***
	(0.0012)	(0.0018)	(0.0023)	(0.0010)	(0.0010)	(0.0010)	(0.0011)	(0.0005)
Age	−0.0004***	−0.0012***	−0.0008	−0.0007**	−0.0003*	−0.0011**	−0.0004***	−0.0011***
	(0.0001)	(0.0003)	(0.0005)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0001)
Collective	−0.0037	0.0021	−0.0189***	0.0098	−0.0076***	−0.0023	−0.0117***	−0.0019
	(0.0030)	(0.0103)	(0.0059)	(0.0087)	(0.0024)	(0.0077)	(0.0027)	(0.0068)
Private	−0.0054**	0.0183**	−0.0027	0.0199***	0.0012	−0.0018	−0.0030	−0.0064
	(0.0026)	(0.0091)	(0.0072)	(0.0072)	(0.0027)	(0.0065)	(0.0028)	(0.0047)
HK-Macao-TW	0.0143**	−0.0132	0.0462***	−0.0008	0.0327***	−0.0318***	0.0451***	−0.0281***
	(0.0066)	(0.0103)	(0.0085)	(0.0086)	(0.0068)	(0.0081)	(0.0063)	(0.0053)
Foreign	0.0497***	−0.0104	0.0616***	0.0047	0.0485***	−0.0387***	0.0688***	−0.0259***
	(0.0111)	(0.0087)	(0.0091)	(0.0104)	(0.0092)	(0.0078)	(0.0053)	(0.0045)
Adj. R^2	0.0646	0.5768	0.0829	0.2198	0.0353	0.3519	0.0593	0.2866
Obs	26,073	11,755	40,253	55,261	51,605	25,285	74,664	50,229

Notes: All regressions include city fixed effects, 4-digit industry fixed effects, year fixed effects and province-year fixed effects. Numbers in parenthesis are standard deviation, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table 3.12: Determinants of Export Status (NP, JT: Continuous Variables), Industry

	Industry 1		Industry 2		Industry 3		Industry 4	
	Entry	Exit	Entry	Exit	Entry	Exit	Entry	Exit
	1	2	3	4	5	6	7	8
New Product (NP)	0.0534***	−0.1230***	0.0961***	−0.0425***	0.0570***	−0.0393***	0.0657***	−0.0488***
	(0.0151)	(0.0211)	(0.0172)	(0.0101)	(0.0113)	(0.0102)	(0.0103)	(0.0047)
Job Training (JT)	−0.0289	1.5125	1.2849***	0.5329	0.1266	−0.3672*	0.2544***	0.4470
	(0.2075)	(2.6261)	(0.5175)	(0.6301)	(0.1416)	(0.2127)	(0.0936)	(0.9286)
Market Share	0.0041***	−0.0107***	0.0167***	−0.0095***	0.0094***	−0.0096***	0.0144***	−0.0069***
	(0.0012)	(0.0018)	(0.0022)	(0.0010)	(0.0010)	(0.0010)	(0.0012)	(0.0006)
Age	−0.0004***	−0.0012***	−0.0007	−0.0007**	−0.0003*	−0.0012***	−0.0004***	−0.0012***
	(0.0001)	(0.0003)	(0.0005)	(0.0003)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
Collective	−0.0043	0.0085	−0.0210***	0.0142*	−0.0095***	0.0015	−0.0158***	0.0063
	(0.0030)	(0.0100)	(0.0060)	(0.0084)	(0.0024)	(0.0075)	(0.0027)	(0.0068)
Private	−0.0058**	0.0246**	−0.0053	0.0245***	−0.0006	0.0022	−0.0071***	0.0111**
	(0.0026)	(0.0097)	(0.0077)	(0.0069)	(0.0026)	(0.0066)	(0.0026)	(0.0047)
HK-Macao-TW	0.0139**	−0.0026	0.0437***	0.0047	0.0306***	−0.0256***	0.0403***	−0.0211***
	(0.0066)	(0.0097)	(0.0090)	(0.0083)	(0.0066)	(0.0081)	(0.0061)	(0.0052)
Foreign	0.0496***	−0.0025	0.0590***	0.0100	0.0461***	−0.0327***	0.0638***	−0.0191***
	(0.0111)	(0.0089)	(0.0094)	(0.0101)	(0.0091)	(0.0079)	(0.0052)	(0.0044)
Adj. R^2	0.0622	0.5723	0.0805	0.2187	0.0323	0.3497	0.0532	0.2846
Obs	26,073	11,755	40,253	55,261	51,604	25,285	74,659	50,228

Notes: All regressions include city fixed effects, 4-digit industry fixed effects, year fixed effects and province-year fixed effects. Numbers in parenthesis are standard deviation, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

3.5 R&D Expenditure

R&D is also broadly applied as an indicator for innovation. In this section, I report the impact of R&D on a firm's export decision by employing an R&D investment (RD) at time $t-1$ instead of the introduction of a new product at time t in the model. Since R&D is an input into the innovation, there is a strong correlation between R&D and the introduction of new goods. Consequently, to prevent multicollinearity, I exclude NP from my analysis. In Tables 3.11 and 3.12, columns 1 and 2 report the results when R&D and JT are modeled as discrete variables, whereas columns 3 and 4 show the results when they are modeled as continuous variables (R&D Expenditure/Output).

According to column 1 in Table 3.13, the likelihood of a firm investing in R&D to export is about 1.1 percentage points higher than firms that do not have any R&D investment. For exiting the export market, the estimated coefficient on R&D is negative and highly significant, i.e. firms investing in R&D are less likely to exit from foreign markets. The likelihood of a firm investing in R&D to exit from exporting is about 0.9 percentage points lower than that of firms that do not invest in R&D. It is worth noting that estimated coefficients on JT in columns 1 and 2 are statistically significant and have the expected signs: firms investing in their workers are more (less) likely to enter (exit from) foreign markets. Note the estimated coefficients on JT in columns 1 and 2 are almost identical to those reported in column 4 of Tables 3.5 and 3.6. In addition, estimates on controls are mostly the same as those reported in columns 4 in Tables 3.5 and 3.6.

In the continuous case, the estimated coefficients on R&D are positive and highly significant for entry to export, and negative and statistically significant for exit from foreign markets. The estimate in column 3 implies that a 10 percent increase in R&D spending increases the likelihood of exporting by five percent. Similarly, the estimate in column 4 implies that a 10 percent increase in R&D spending decreases the likelihood of exiting from foreign market by three percent. Estimates on JT are significant at the 10 percent level only for the entry to export, and its impact on exit from foreign markets is positive but insignificant. These estimates by and large are similar to those reported in Table 3.7. Estimates on controls are also largely the same as those in Table 3.7.

In sum, consistent with previous studies that used R&D expenditure as a proxy for technology (Aw et al. 2008 & 2011), I find that it has a positive and significant effect on entry to export, and negative and significant impact on the exit from foreign markets. Thus, firms investing in R&D are more likely to export and less likely to stop exporting. This analysis further shows that using R&D instead of NP does not have any significant impact on estimated coefficients on job training and controls. Job training has a significant positive (negative) effect on entry (exit) when it is modeled as a discrete variable.

3.6 Conclusion

This paper investigates the impact of innovation and investment in human capital on the likelihood of firms' entry to and exit from foreign markets. To this end, I use a detailed firm-level database for Chinese manufacturing firms over the 2003 – 2011 period. The data include variables (such as firms' location, industry, output, number of employees, assets, R&D spendings, and export) that previous studies have widely used. In addition, the data include how much new products that each firm produces in each year, and how much each firm spends on workers' training. I use the first variable as proxy for firm innovative capacity, and the latter one as a proxy for investment in human capital.

My benchmark analysis shows that innovative firms are more likely to export, and less likely to exit from exporting. I also show that a similar conclusion holds for job training as well: firms providing job training are more likely to export, and less likely to exit from foreign markets. I then extend my analysis to investigate any heterogeneity across firms. To this end, I consider heterogeneity by ownership and industry, respectively. In the former case, I consider privately owned firms (which also include firms owned by foreigners and residents of Hong Kong, Macao, and Taiwan) and state-owned enterprises (which include collective firms as well) separately. For both samples, I find that firms that are innovative and invest in workers' training are more likely to begin exporting, and less likely to quit exporting, However, these effects are stronger among private firms.

Exploring heterogeneity across industries, I first partition the set of firms into four broad industries. Two of these industries (which mainly include food, beverage, tobacco, textile, wood, etc.) are low-skill and less-capital intensive, and the other two (which mainly include chemical products, equipment and machinery, transport equipment, etc) are high-skill and capital intensive industries. Consistent with expectations, my analysis shows that innovation and job training has a more substantial impact on firm export status in the high-tech industries. That is, firms that are more innovative and invest in workers' training are substantially more likely to export, and less likely to exit from foreign markets.

In this paper, I investigated the determinants of first-time export and first-time to exit from export. Analyzing the problem in a dynamic framework where firms can enter to and exit from foreign market multiple times will provide a better understanding of the dynamics of export participation decisions. In addition, several studies have shown that exporting induces firms to invest more in R&D and human capital (Atkinson and Burstein 2008, Bustos 2011, and Unel 2013). In a dynamic model, investigating how exporting in turn affects creation of new products and their investment in job training will shed more light on our understanding of the nexus between exporting and innovation and human capital formation.

Table 3.13: Determinants of Export Status, RD and JT

	Discrete		Continuous	
	Entry	Exit	Entry	Exit
	1	2	3	4
R&D (RD)	0.0112***	−0.0090***	0.1641***	−0.1129**
	(0.0016)	(0.0013)	(0.0370)	(0.0095)
Job Training (JT)	0.0045***	−0.0083***	0.2669*	1.0062
	(0.0007)	(0.0010)	(0.0891)	(0.7372)
Market Share	0.0104***	−0.0084***	0.0112***	−0.0095***
	(0.0006)	(0.0005)	(0.0006)	(0.0005)
Age	−0.0005***	−0.0010***	−0.0005***	−0.0011***
	(0.0001)	(0.0002)	(0.0000)	(0.0002)
Collective	−0.0120***	0.0079**	−0.0133***	0.0105***
	(0.0015)	(0.0039)	(0.0015)	(0.0039)
Private	−0.0048***	0.0107***	−0.0062***	0.0134***
	(0.0015)	(0.0040)	(0.0015)	(0.0042)
HK-Macao-TW	0.0344***	−0.0162***	0.0328***	−0.0126**
	(0.0032)	(0.0052)	(0.0032)	(0.0053)
Foreign	0.0571***	−0.0146**	0.0557***	−0.0112*
	(0.0039)	(0.0060)	(0.0038)	(0.0061)
Adj. R^2	0.0573	0.3396	0.0568	0.3391
Obs	269,157	183,150	269,127	183,135

Notes: All regressions include city fixed effects, 4-digit industry fixed effects, year fixed effects and province-year fixed effects. Numbers in parenthesis are standard deviation, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Chapter 4. Effect of R&D on Innovation in China

4.1 Introduction

The hallmark of endogenous growth models is that economic growth is driven by the research and development (R&D) efforts of profit maximizing entrepreneurs, with the implication that an increase in resources devoted to R&D increases the growth rate of output per worker. Key studies include Romer (1990), Grossman and Helpman (1991), Aghion and Howitt (1992), and Jones (1995). In Romer (1990), technological progress is measured by the introduction of new goods (i.e., horizontal differentiation), whereas in Grossman and Helpman (1991) and Aghion and Howitt (1992) by the introduction of higher quality goods. Jones (2007) provides a comprehensive review of this literature.

These models also revived a large literature in industrial organization (IO) that empirically investigates the impact of R&D on output growth (Griliches 1979 & 1991). The main finding of this literature is that R&D (measured by its total spending) has a significant impact on output growth. However, these studies have mainly investigated the impact of R&D on either output or total factor productivity (TFP) – the latter is assumed to be a proxy for technology. In this chapter, I contribute to this empirical literature by directly investigating the impact of R&D on introduction of new goods. Specifically, using a novel firm-level database from China over the 2003 – 2011 period, I measure the effect of R&D investment on introduction of new goods. The database, also known as *China Annual Survey of Industrial Firms*, is compiled by China's National Bureau of Statistics and provides rich information about "large-scale" manufacturing firms, such as firm location, its industry, output produced, number of workers, total fixed assets, firm ownership, etc. It also provides information about sales from new products produced, firm R&D investment, and spending on workers' job training. Using a panel data model, I then examine the effect of R&D investment and workers' training on the introduction of new products.

Data on new products provided in this database report the value of the output of newly launched products each year. The National Bureau of Statistics in China loosely define a new product as *either products based on groundbreaking technology, new idea, or new process; or products based on significant renovation in structure, materials or technique, and shows the recognizable innovation progression and productivity growth which could be distinguished from the previous, exiting stock of technology.* On one hand, it is important to note that the understanding of *new product* is independent between firms, which may promote measurement error and heterogeneity concerns. On the other hand, the credibility of these observations is endorsed by each firm with further validation by the local Bureau of Statistics and Tax Bureau.

Information about the new products introduced is in the form of total sales from new products. Consequently, I do not know how many new products are produced in each period. Therefore, I measure the value of innovation in two ways. In my main analysis, I measure it as a continuous variable: log values of the total sales of new products. I also present results based on an indicator variable: if a firm sells a new product in a given year, the value of innovation equals one in that year, and zero otherwise. Results based on the second approach are qualitatively similar to the first. I measure R&D activity as total R&D spending, and investment in human capital as total spending on workers' training. Normalizing these variables by total output produced yields similar results.

I find that investment in R&D and workers' training have positive and statistically highly significant effects on the introduction of new goods. Specifically, a 10-percent increase in R&D spending increases sales of new products in the next year by about 3.5 percent. Similarly, a 10-percent increase in spending on workers' training increases sales by 0.5 percent. I also find that the impact of R&D on new products is comparable across SOE and non-SOEs. However, the impact of investment in workers' training is higher among non-SOEs. I then extend my analysis to investigate whether these effects differ across industries. It seems the impact of R&D is stronger in high-tech industries, where the bulk of R&D is done. I also find that the effect of workers' training is stronger in high-tech industries.

This study contributes to endogenous growth theory by quantifying implications of investment in R&D and human capital on introduction of new goods. In a path-breaking paper, Romer (1990) develops the first R&D based growth model in which entrepreneurs invest in R&D to develop new products. His model shows that product variety expands if resources devoted to R&D and human capital increase. Grossman and Helpman (1991) and Aghion and Howitt (1992) develop quality-ladder growth models, in which growth is driven by rising product quality through investment in R&D. Although these models exhibit a scale effect (Jones 1995a & 1995b), later models without scale effects deliver the same conclusion that resources devoted to R&D drive economic growth. My analysis is more related to the Romer model by empirically investigating the impact of R&D on introducing new products.

This paper also contributes to a large empirical literature that investigates the effects of R&D on economic growth. These studies are done at the firm, industry, or regional level, and early studies usually use data from developed countries. Estimates from this literature indicate that R&D has a significant impact on productivity and growth. As emphasized above, these studies usually use effects on productivity measured by output per worker or TFP. Hall et al. (2010) provides an excellent review of this literature, and Keller (2010) reviews studies that investigate R&D spillovers across regions. Unlike studies in this literature, my analysis focuses on the effects of R&D and human capital investment on product development. In addition, I investigate in a developing country context.

Finally, this paper contributes to a large body of research on China that investigates the determinants of its growth. Researchers have considered several factors behind the rapid growth in China, but the effects of R&D have drawn limited attention. If R&D plays a significant role on growth, policy makers should subsidize R&D investment. Policy implications of the nexus between R&D and innovation are important for the Chinese economy. R&D expenditure in China has been growing since its economic reform in 1980s. Table 4.1 provides information about R&D expenditure (% of GDP) between 2003 and 2011. As

Table 4.1: Research and Development Expenditure (% of GDP) and Growth Rate Each Year (%) in China and the World, 2003 – 2011

	Level		Growth Rate (%)	
	China	World	China	World
2003	1.120	2.051	5.909	−0.003
2004	1.215	2.003	8.445	−2.337
2005	1.308	1.985	7.650	−0.855
2006	1.369	1.996	4.635	0.510
2007	1.373	1.965	0.324	−1.537
2008	1.445	2.022	5.223	2.889
2009	1.662	2.059	15.050	1.823
2010	1.710	2.040	2.878	−0.898
2011	1.775	2.034	3.828	−0.311

Notes: Data is based on *World Development Indicators: Science and technology* from World Bank.

shown in columns 1 and 3 of Table 4.1, based on the data of *World Development Indicators* from World Bank, the research and development expenditure (% of GDP) in China has been increasing from 1.1% in 2003 to 1.8% in 2011. During this period, the world average R&D expenditure (% of GDP) did not change much and remained around 2%. Columns 2 and 4 in Table 4.1 show that the growth rate of R&D expenditure (% of GDP) in China has been consistently and significantly higher than that of the rest of world. Over the period of 2003 to 2011, China experienced an average growth rate of 6%, while the world simple average was about -0.1%.

This chapter is organized as follows. The next section introduces the data employed in analysis and provides summary statistics. Section 3 describes the econometric methodology, and Section 4 presents and discusses the results. This section also explores the heterogeneity across firms based on their ownership and industry. Section 5 offers concluding remarks.

4.2 Data

The data used in this paper is from the China Annual Survey of Industrial Firms, compiled by China's National Bureau of Statistics with the assistance of local Bureaus of Statistics. This data set consists of state-owned enterprises and the large-scale non-state-owned enterprises. For this particular data set, any non-state-owned enterprise which generates at least five million current Chinese Yuan of annual revenue is defined as a large-scale company. Such a non-SOE is requested by the local Bureau of Statistics to self-report its brief financial statements annually. The observations used for this analysis are from the period 2002 to 2011, providing about 2.8 million observations in total. Each observation shows a firm's basic information (e.g., firm's ID, location, products, industry at 4-digit level, ownership structure), balance sheet (assets and liabilities), income statement (total revenue, revenue of new product, R&D expenditure, spending on workers' job training, administration expense, etc.) and cash flow statement in a specific year. The observation is individual legal entity-based instead of group company-based. For example, if a parent group company has three subsidiaries (independent legal entities) qualified to be included in the data set in current year, one should see the separated observations of each subsidiary instead of a consolidated report of the parent group.

As mentioned above, the original data set includes about 2.8 million observations; however, due to the weak quality control when it was collected and complied in early survey years, there exist data measurement issues which have been confirmed by previous research. Cai and Liu (2009) and Nie, Jia, and Yang (2012) both indicate that there are potential problems in this data set, such as non-unique firm IDs, missing basic information like location or industry classification, and low reliability in some observations (e.g., privately-owned firm with sales less than five million Chinese Yuan, or negative R&D expenditure, etc.). Following their suggestions to clarify and clean the questionable observations, about two million observations remaining during the reporting period. About 50% of firms provide four or more years of observations in the processed sample. Thus, the data set is an unbalanced panel.

All of the nominal financial variables are converted into their real value by using the industry-level Producer Price Index (PPI) from National Bureau of Statistics in China. Here, the real value of R&D expenditure and its variations provide the investment in research activities as the input, while the real output value of a new product and its variations indicate the innovation as outcome. The salary of R&D researchers is included in R&D expenditure based on *Chinese Accounting Standards*. The output of a new product is self-reported by the enterprise. There is no proof of further data validation provided for this specific term. It is worthy to note that different enterprises may have divergent understandings about the definition of new products in real-world practice. Therefore, measurement errors in the first-hand data might still exist in this cleaned data set.

Tables 4.2 and 4.3 provide the summary statistics on key variables, including the output of new products, R&D expenditure, and spending on workers' job training. Table 4.2 reports the statistics of these variables in continuous logarithmic form. The statistics for the full sample, SOEs, and non-SOEs are provided, respectively. Table 4.3 reports the statistics of these variables in the variation of discrete variables. In each of these tables, the statistics of the full sample, the sub-sample including only SOEs[1] and the sub-sample of non-SOEs[2] are shown, respectively. As shown in Tables 4.2 and 4.3, the SOEs are generally bigger in size of output, invest more in R&D activities and workers' job training, and introduce more new products than those of non-SOEs.

To explore the heterogeneity across industries, we create four industry categories with different levels of R&D intensity. To this end, firms are grouped into four broad industries as specified in Table 4.4. These industries are constructed based on the *Industrial Classification for National Economic Activities in China (2011)*. Table 4.4 also provides the R&D intensity,

[1]SOEs include state-owned enterprises and collective-controlled enterprises in this paper.

[2]Non-SOEs include private-controlled enterprises, Hong Kong-Macao-Taiwan resident-controlled enterprises and foreign-controlled enterprises in this paper.

Table 4.2: Summary Statistics on Key Variables

	All		SOEs		non-SOEs	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Output	10.516	1.385	10.548	1.858	10.512	1.298
New Product	0.777	2.626	1.755	3.892	0.624	2.331
R&D	0.654	1.996	1.210	2.751	0.574	1.849
Job Training	1.183	1.739	1.891	2.168	1.082	1.644
Age	8.719	10.476	18.170	16.278	7.262	8.363

Notes: The variables (expect for *Age*) are in the logarithmic form of real value (price in 2011).

Table 4.3: Summary Statistics on Key Variables

	All		SOEs		non-SOEs	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
New Product	0.0861	0.2806	0.1817	0.3856	0.0710	0.2569
R&D	0.1115	0.3147	0.1864	0.3895	0.1007	0.3010
Job Training	0.4146	0.4926	0.5454	0.4979	0.3959	0.4890

Notes: The variables are discrete variables.

which is defined as the ratio of R&D expenditure to total output. Industry 1 includes agricultural products and related industries, which generally requires less technology and capital investment. Industry 2 is composed of the textile, leather and wood product, and related industries, with similar technology and capital requirements to industry 1. As shown in Table 4.4, industries 1 and 2 both have lower R&D intensities compared to industries 3 and 4. Industry 3 includes the industries producing rough machining and semi - finishing of the material products. High capital density with the application of mature technology is the characteristic of industry 3. Industry 4 includes sectors such as machinery, equipment and IT products industries, which has highest requirement for capital investment among the four industries. Additionally, industry 4 applies relatively more advanced and newer technology

Table 4.4: Industries and its R&D Intensity (RD/Y * 100 %)

Industry	Sub-industry	Mean	S.D.
Industry 1	Agricultural products and food manufacturing, Beverage manufacturing, Tobacco manufacturing	0.08	0.73
Industry 2	Textile manufacturing, Clothing manufacturing, Leather and feather products manufacturing, Lumber and bamboo products manufacturing, Furniture manufacturing, Paper products Manufacturing	0.05	0.49
Industry 3	Chemical product manufacturing, Pharmaceutic manufacturing, Rubber and plastic products manufacturing, Metal products manufacturing	0.13	1.01
Industry 4	General equipment manufacturing, Special-purpose equipment manufacturing, Stationary manufacturing, Automobile manufacturing, Other transportation manufacturing, Computer manufacturing	0.25	1.43

in general, which is implied by the highest R&D intensity among these four industries.

As mentioned earlier, only manufacturing industries are included. Table 4.5 reports the number of enterprises in each industry in 2003, 2005, 2007, 2009 and 2011. As reported in Table 4.5, the number of enterprises in each industry has been increasing over the period between 2003 and 2011, the high-tech industry 4 has been growing fastest. Industry growth is generally tied to the high-tech and high capital-intensity sectors.

Table 4.5: Number of Enterprise in Each Industry Category, 2003 to 2011

	2003	2005	2007	2009	2011
Industry 1	2,353	3,159	27,344	24,165	25,729
Industry 2	1,594	3,677	68,746	56,791	78,401
Industry 3	5,053	11,787	105,675	88,506	115,032
Industry 4	3,761	11,482	100,191	89,520	111,866

4.3 The Empirical Model

The following model is used to estimate the impact of R&D and human capital on product expansion:

$$\ln(\mathrm{New}_{it}) = \alpha_1 \ln(\mathrm{RD}_{it-1}) + \alpha_2 \ln(\mathrm{RD}_{it-2}) + \beta \ln(\mathrm{JT}_{it-1}) + \theta X_{it} + \eta_j + \eta_c + \eta_{pt} + \varepsilon_{it}, \quad (4.1)$$

where i denotes firm, j industry, c city, p province, and t year. New_{it} represents the output value of new products produced by firm i in year t. I also present results in the appendix when New_{it} is a categorical variable that equals one if firm i produces new products in year t, zero otherwise. Variable RD represents the R&D expenditure, and JT denotes spending on workers' job training. Note that I include the lagged values of RD and JT in the model, because it takes time for R&D activities and job training to impact production. I include only JT_{it-1}, because higher lags do not have any significant effects on outcome.

Variable X represents the set of controls, including the firm's last year output, its age, and its ownership structure. There are five different types of ownership: state-owned enterprises, collective enterprises, private enterprises, enterprises controlled by residents in Hong Kong-Macao-Taiwan, and foreign-controlled enterprises. In some specifications below, I only consider a subset of these different ownerships (e.g., SOE). Firm age is measured as the number of years between firm establishment and the year of the observation in the sample.

City and industry fixed effects (η_c and η_j) are included to control for any time invariant city and industry specific factors that can affect the launch of new products. Year fixed effects (η_t) is controlled for common shocks, and note that they are measured at the provincial level. Finally, ε_{it} denotes the error term. Heteroskedasticity robust standard errors are clustered at the 4-digit industry-level to mitigate the potential serial correlation in the error term (Bertrand et al. 2004). Clustering data at the city-level yields qualitatively similar results.

4.4 Results

This section reports results based on equation (4.1). Section 4.1 presents results using the full sample. Results show that firms investing more in R&D and their workers' training are more likely to have higher output value of new product, and more likely to have innovative outcomes. In Section 4.2, we investigate the heterogeneity between different ownership structures. Section 4.3 restricts the heterogeneity across industries.

4.4.1 Benchmark Results

As described in equation (4.1), New, as the proxy for innovation, is the continuous logarithmic variable. The results when using the discrete variable of New are reported in Appendix. The independent variables including $R\&D$, JT and $Output$ are also continuous logarithmic variables. Table 4.6 reports the benchmark results of equation (4.1), in which the full sample is employed.

In Table 4.6, Column (1) reports the scenario when R&D in the previous two periods are used as the independent variables to explain the likelihood of producing a new product in the following year. The coefficients of R&D are significantly positive, which instructs the positive impact of R&D on innovation. Column (1) instructs that 1% more expenditure on

R&D at time *t-1* could increase the likelihood of launching new products at time t by about 0.45%. Column (1) shows that an additional 1% spending on R&D at time *t-2* could increase the likelihood of launching new products at time t by about 0.35%. The impact of R&D at one previous period is stronger than that at time *t-2*. In Column(2), the investment in job training at time *t-1* is added to the model. As shown in Column (2), the R&D expenditure in the previous period still maintains positive impact, the job train at time *t-1* shows small but positive impact as well. A 1% increase in spending on job training could increase the likelihood of producing new products at time t by around 0.17%. Furthermore, in Column (3), the variable of output value at time *t-1* is added to controlled for the size of firm. As shown in Column (3), the bigger firm shows higher likelihood to innovate. 1% more output value at time *t-1* could increase the likelihood of innovation at time t by about 0.41%. At the same time, R&D and job training still keep their significantly positive coefficients.

Column (4) shows the scenario when output value are firm age are both controlled for in equation (4.1). Column (5) shows the results when ownership is added as the control variable. As shown in Column (5), a 1% increase in R&D expenditure at time *t-1* increases the likelihood of producing new products at time t by almost 0.35%. Similarly, R&D expenditure at time *t-2* has a smaller, but still positive, impact. Similar to the findings in other configurations, job training, total output, and firm age all indicate statistically significant and positive impact. As to the ownership, as shown in Column (5), compared to state-owned enterprises, collective-owned enterprises are 0.33% less likely to produce new products, private-owned enterprises are 0.34% less likely to produce new products, Hong kong-Macao-Taiwan resident-owned enterprises and foreign-owned enterprises are about 0.55% and 0.56% less likely to produce new products at time t, respectively. In all, the results indicate that R&D and job training are helpful to encourage firm innovation in China.

Table 4.6: Impact of R&D on the Introduction of New Product

	1	2	3	4	5
$R\&D_{t-1}$	0.4476***	0.3988***	0.3553***	0.3520***	0.3476***
	(0.0269)	(0.0248)	(0.0232)	(0.0232)	(0.0230)
$R\&D_{t-2}$	0.3540***	0.3284***	0.2926***	0.2877***	0.2832***
	(0.0212)	(0.0214)	(0.0215)	(0.0214)	(0.0215)
JT_{t-1}		0.1729***	0.0759***	0.0656***	0.0523***
		(0.0189)	(0.0146)	(0.0142)	(0.0136)
$Output_{t-1}$			0.4121***	0.4156***	0.4202***
			(0.0311)	(0.0308)	(0.0317)
Age				0.0116***	0.0097***
				(0.0020)	(0.0020)
Collective					−0.3319***
					(0.0661)
Private					−0.3418**
					(0.0724)
HK-Macao-TW					−0.5463***
					(0.1297)
Foreign					−0.5556***
					(0.1610)
Adj. R^2	0.4088	0.4158	0.4324	0.4345	0.4361
Obs	23,511	23,479	23,474	23,474	23,474

Notes: R&D, Job Training and Output are continuous logarithmic variables. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

4.4.2 Heterogeneity by Ownership

The benchmark results reported in the previous section use the full sample. However, according to Tables 4.2 and 4.3, non-SOEs and SOEs differ from each other substantially in many ways. Here, as described in previous section, non-SOEs include private enterprises, enterprises controlled by residents in Hong Kong-Macao-Taiwan, and foreign-controlled enterprises. SOEs include state-owned enterprises and collective enterprises. Table 4.7 shows the results for sample of SOEs and sample of non-SOEs, respectively, with the same model setting in equation (4.1).

In Table 4.7, all three configurations for the sample of SOEs (shown in columns (1), (2) and (3)) report the similar results with those in the full sample case. The estimated coefficients of R&D and job training are statistically significant and positive across different settings, with or without further controlling for output value, firm age, and ownership structure. Similarly, the results based on sample of non-SOEs are consistent with those in SOEs case. It is worthy to note that R&D in different periods have different magnitudes across the SOEs and non-SOEs sample. For example, comparing the results in Columns (3) and (6), the coefficient of R&D at time *t-1* in SOEs are larger than that in non-SOEs. Meanwhile, the coefficient of R&D at time *t-2* in SOEs are smaller than that in non-SOEs. The results in the other columns also show comparable results. The estimates of job training at time *t-1* are both statistically significant and positive, as shown in all columns. Comparing the estimates of job training in Columns (3) and (6), the results imply that job training is playing a more important role in non-SOEs than SOEs, since its estimate in non-SOEs is larger that that in SOEs by more than 57% (i.e., 0.0737 / 0.0467). With or without the controls for output, firm age and ownership, the estimates of R&D and job training are consistent across models. Similar to the findings in benchmark models, output and firm age both have positive influence on producing new products.

Table 4.7: Impact of R&D on the Introduction of New Product, Heterogeneity by Ownership

	SOEs			non-SOEs		
	1	2	3	4	5	6
$R\&D_{t-1}$	0.4392***	0.3903***	0.3834***	0.2761***	0.2481***	0.2456***
	(0.0280)	(0.0263)	(0.0264)	(0.0367)	(0.0361)	(0.0358)
$R\&D_{t-2}$	0.3044***	0.2626***	0.2550***	0.3408***	0.3260***	0.3250***
	(0.0231)	(0.0234)	(0.0233)	(0.0487)	(0.0482)	(0.0480)
JT_{t-1}	0.1855***	0.0716***	0.0467***	0.1331***	0.0776***	0.0737***
	(0.0230)	(0.0177)	(0.0167)	(0.0240)	(0.0231)	(0.0230)
$Output_{t-1}$		0.4596***	0.4621***		0.2781***	0.2759***
		(0.0371)	(0.0372)		(0.0363)	(0.0370)
Age			0.0095***			0.0113***
			(0.0021)			(0.0050)
Ownership			Y			Y
Adj. R^2	0.4458	0.4636	0.4324	0.2830	0.2949	0.2961
Obs	16,410	16,406	16,406	6,971	6,970	6,970

Notes: R&D, Job Training and Output are continuous logarithmic variables. Columns (1), (2) and (3) report the results based on SOEs. Columns (4), (5) and (6) report the results based on non-SOEs. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

4.4.3 Heterogeneity by Industry

Tables 4.8, 4.9, 4.10 and 4.11 show the results for each industry. Each industry is defined in the data section. Table 4.8 reports the estimates when using the sample of industry 1, etc. Output value, firm age, and ownership structure are controlled for in each configuration. The estimated coefficients of R&D are all significantly positive across industries. With the heterogeneity by industry, we find that the R&D in SOEs is playing a more important role than in non-SOEs. Across industries, the estimated coefficients of R&D at time *t-1* for the sample of SOEs are consistently larger than those for the sample of non-SOEs. For R&D at time *t-2*, coefficients for SOEs are generally more significant and positive, except for the case of industry 4. All configurations report positive estimated coefficients for R&D and job training.

Across industries, the results show that the estimated coefficients of R&D in industry 2 are higher than in the other three industries. Industries 4, 3 and 1 follow. R&D is most influential to innovation in industry 2 and industry 4, while its impact is less substantial but still positive in industries 1 and 3. These results imply that the industries with relatively low requirements for capital investment and less application of mature technology may benefit more from R&D activities than industries requiring heavier capital investment and existing mature technology. Further, the most high-tech industries, though requiring more capital investment, are more likely to benefit from R&D than any other types of industries.

Except for some cases in industries 1 and 2, the estimated coefficient of job training are all statistically significant and positive, such as those shown in both Columns (2) and (3) in Tables 4.10 and 4.11. These results imply that the investment in workers' job training could be helpful to promote technological progress in industries 3 and 4, while its impact is statistically insignificant for industries 1 and 2 in some configurations. The results without controlling for output value, firm age and ownership, are similar to those we report.

Table 4.8: Impact of R&D on the Introduction of New Product in Industry 1

	All	SOEs	non-SOEs
$R\&D_{t-1}$	0.3482***	0.3562***	0.2859***
	(0.0628)	(0.0828)	(0.1074)
$R\&D_{t-2}$	0.2485***	0.2825***	0.2190
	(0.0649)	(0.0815)	(0.1775)
JT_{t-1}	0.0161	0.0334	0.0606
	(0.0345)	(0.0417)	(0.0876)
$Output_{t-1}$	0.2383***	0.2045*	0.2011
	(0.0753)	(0.1045)	(0.1266)
Age	−0.0061	−0.0089	−0.0040
	(0.0061)	(0.0077)	(0.0105)
Adj. R^2	0.2765	0.3136	0.0835
Obs	2,184	1,533	558

Notes: R&D, Job Training and Output are continuous logarithmic variables. Ownership is controlled for across models. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

4.5 Concluding Remarks

This chapter investigates the impact of R&D activities on innovation by using richly informative, firm-level data from manufacturing sectors in China over the 2003 – 2011 period. Innovation is measured by the total sales of new products. Spending on R&D is used a proxy for the intensity of R&D activity, and spending on workers' job training is used as a proxy for investment in human capital. The database provides firms' basic information such as firms' ID, location, industry classification, established year, and ownership structure, as well as the consolidated financial statements including balance sheet, income statement, and cash flow statement. For example, these financial statements provide the total output, output of new product, R&D expenditure, job training spending, and many other observations.

Table 4.9: Impact of R&D on the Introduction of New Product in Industry 2

	All	SOEs	non-SOEs
$R\&D_{t-1}$	0.4570***	0.5241***	0.3104***
	(0.0868)	(0.0906)	(0.1115)
$R\&D_{t-2}$	0.2002***	0.1486*	0.1599
	(0.0698)	(0.0871)	(0.1630)
JT_{t-1}	0.0420	0.0552	0.0441
	(0.0376)	(0.0565)	(0.0369)
$Output_{t-1}$	0.3284***	0.4044***	0.1798***
	(0.0668)	(0.0846)	(0.0520)
Age	0.0114**	0.0113	0.0153***
	(0.0058)	(0.0083)	(0.0058)
Adj. R^2	0.3117	0.3524	0.1671
Obs	3,158	1,834	1,243

Notes: R&D, Job Training and Output are in the logarithmic form. Ownership is controlled for across models. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Using a panel data approach, my analysis yields several interesting findings. First, I find both R&D and workers' training have positive and highly significant effects on introducing new products. Second, I find that the impact of R&D on new products is comparable across state owned enterprises (SOE) and non-SOEs. However, the impact of investment in human capital is higher among non-SOEs. Third, the impact of R&D on product innovation is stronger in high-tech industries, where the bulk of R&D is done. I also find that the effect of workers' training is stronger in the high-tech industries. My results are robust to the choice of controls and the way the new products are measured.

Analysis presented in this chapter can be extended in various directions. I consider only direct effect of R&D on product innovation. However, there is a large literature that has investigated the impact of R&D spillovers on productivity. Since my panel data were unbal-

Table 4.10: Impact of R&D on the Introduction of New Product in Industry 3

	All	SOEs	non-SOEs
$R\&D_{t-1}$	0.2229***	0.2356***	0.1982***
	(0.0324)	(0.0400)	(0.0480)
$R\&D_{t-2}$	0.3081***	0.3095***	0.2110***
	(0.0313)	(0.0381)	(0.0524)
JT_{t-1}	0.0491**	0.0266	0.1079***
	(0.0206)	(0.0237)	(0.0272)
$Output_{t-1}$	0.3334***	0.3923***	0.1826***
	(0.0461)	(0.0566)	(0.0525)
Age	0.0123***	0.0132***	0.0055
	(0.0028)	(0.0031)	(0.0080)
Adj. R^2	0.3513	0.3524	0.2101
Obs	9,554	6,810	2,647

Notes: R&D, Job Training and Output are continuous logarithmic variables. Ownership is controlled for across models. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

anced, I could not include spillovers effects. But my goal is to get more data from China's National Bureau of Statistics to investigate how R&D spillovers from other firms/industries affected introduction of new products. Another extension is to investigate the impact of foreign R&D spillovers on product innovation in China. These extensions will nicely complement works done by Coe and Helpman (1995), Keller (2002), and Unel (2008), among many others.

Table 4.11: Impact of R&D on the Introduction of New Product in Industry 4

	All	SOEs	non-SOEs
$R\&D_{t-1}$	0.4068***	0.4713***	0.2531***
	(0.0285)	(0.0314)	(0.0601)
$R\&D_{t-2}$	0.2868***	0.2367***	0.4035***
	(0.0351)	(0.0362)	(0.0873)
JT_{t-1}	0.0802***	0.0867***	0.0892*
	(0.0238)	(0.0293)	(0.0527)
$Output_{t-1}$	0.5725***	0.5771***	0.4542***
	(0.0535)	(0.0560)	(0.0748)
Age	0.0103***	0.0089***	0.0217**
	(0.0033)	(0.0034)	(0.0090)
Adj. R^2	0.5168	0.5481	0.3847
Obs	8,938	6,458	2,396

Notes: R&D, Job Training and Output are continuous logarithmic variables. Ownership is controlled for across models. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Bibliography

Acemoglu, Daron, Francisco A. Gallego, and James A. Robinson, "Institutions, Human Capital, and Development," *Annual Review of Economics*, 2014, *6*, 875 – 912.

Acemoglu, Daron, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 2001, *91*, 1369 – 1401.

Acemoglu, Daron, Simon Johnson, and James A. Robinson, "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 2002, *107*, 1231 – 1294.

Aghion, Philippe and Peter Howitt, "A Model of Growth Through Creative Destruction ," *Econometrica*, 1992, 60(2), 323 – 351.

Angrist, Joshua D. and Jörn-Steffen Pischke, "Mostly Harmless Econometrics." Princeton University Press, Princeton, NJ, 2009.

Arnold, Jens Matthias, and Katrin Hussinger, "Export Behavior and Firm Productivity in German Manufacturing: A Firm-Level Analysis," *Review of World Economics*, 2005, *141(2)*, 219 – 243.

Atkinson, Andrew, and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices.," *American Economic Review*, 2008, *98(5)*, 1998 – 2031.

Atkinson, Andrew, and Ariel Burstein, "Aggregate Implications of Innovation Policy," *Journal of Political Economy*, June 2018, Forthcoming.

Aw, Bee Yan, Mark Roberts, and Daniel Yi Xu, "R&D Investments, Exporting, and the Evolution of Firm Productivity," *American Economic Review*, 2008, *98(2)*, 451 – 456.

Aw, Bee Yan, Mark Roberts, and Daniel Yi Xu, "R&D Investment, Exporting, and Productivity Dynamics," *American Economic Review*, 2011, *101(4)*, 1312 – 1344.

Bernard, Andrew, and Bradford Jensen, "Exporters, Skill Upgrading and the Wage Gap," *Journal of International Economics*, 1997, *42*, 3 – 31.

Bernard, Andrew, and Bradford Jensen, "Exceptional Exporter Performance: Cause, Effect, or Both?," *Journal of International Economics*, 1999, *47*, 1 – 25.

Bernard, Andrew, and Bradford Jensen, "Why Some Firms Export?," *Review of Economics and Statistics*, 2004, *86(2)*, 561 – 569.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-differences Estimates?," *Quarterly Journal of Economics*, 2004, *119*, 249 – 275.

Bleakley, Hoyt, and Jeffrey Lin, "Portage and Path Dependence," *Quarterly Journal of Economics*, 2012, *127*, 587 – 644.

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, 2011, *101(1)*, 304 – 340.

Bockstette, Valerie, Areendam Chanda, and Louis Putterman, "States and Markets: The Advantage of an Early Start," *Journal of Economic Growth*, 2002, *7*, 347 – 69.

Brandt, Loren, Chang-tai Hsieh, and Xiaodong Zhu, "Growth and Structural Transformation in China," *China's Great Economic Transformation*, Edited by L. Brandt and T. G. Rawski, Cambridge University Press, NY: New York, 2008.

Brandt, Loren, Debin Ma, and Thomas G. Rawski, "From Divergence to Convergence: Reevaluating the History behind China's Economic Boom," *Journal of Economic Literature*, 2014, *52*, 45 – 123.

Cai, Hongbin, and Qiao Liu, "Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms," *The Economic Journal*, 2009, *119(537)*, 764 – 795.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90*, 414 – 427.

Cao, Shuji, "The Population History of China, vol.4, 'Period of Ming Dynasty'," *Fudan University Press*, China: Shanghai, 2000.

Cao, Shuji, "The Population History of China, vol.5, 'Period of Qing Dynasty'," *Fudan University Press*, China: Shanghai, 2001.

Chanda, Areendam, C. Justin Cook, and Louis Putterman, "Persistence of Fortune: Accounting for Population Movements, There was no Post-Columbian Reversal," *American Economic Journal: Macroeconomics*, 2014, *6*, 1 – 28.

Chanda, Areendam, and Dachao Ruan, "Early Urbanization and the Persistence of Regional Disparities within Countries," *LSU Working Paper*, 2017.

Chow, Gregory, "China's Economic Transformation," Second Edition, *Wiley-Blackwell*, MA: Malden, 2007.

Clark, Hunter, Maxim Pinkovskiy, and Xavier Sala-i-Martin, "China's GDP Growth May be Understated," *NBER Working Paper*, 2017.

Clerides, Sofronis, James Tybout, and Saul Lach, "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco," *Quarterly Journal of Economics*, 1998, *113(3)*, 903 – 947.

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

Comin, Diego, William Easterly, and Erick Gong, "Was the Wealth of Nations determined in 1000 BC?," *American Economic Journal: Macroeconomics*, 2010, *2*, 65 – 97.

Davis, Donald R., and David E. Weinstein, "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 2002, *92*, 1269 – 1289.

Dinopoulos, Elias, and Bulent Unel, "Managerial Capital, Occupational Choice, and Inequality in a Global Economy," *Canadian Journal of Economics*, 2017, *50*, 365 – 397.

Donaldson, Dave, and Adam Storeygard, "The View from Above: Applications of Satellite Data in Economics," *Journal of Economic Perspective*, 2016, *30*, 171 – 98.

Dong, Xiao-yuan, Louis Putterman, and Bulent Unel, "Privatization and Firm Performance: A Comparison Between Rural and Urban Enterprises in China," *Journal of Comparative Economics*, 2006, *34*, 608 – 633.

Duan, Fan, and Bulent Unel, "Persistence of Cities: Evidence from China," *Review of Development Economics*, 2019, *23(2)*, 663 – 676.

Eaton, Jonathan, and Zvi Eckstein, "Cities and Growth: Theory and Evidence from France and Japan," *Regional Science and Urban Economics*, 1997, *27*, 443 – 474.

Egger, Peter, and Christian Keuschnigg, "Access to Credit and Comparative Advantage," *The Canadian journal of economics*, 2017, *50(2)*, 481 – 505.

Fan, Haichao, Edwin Lai, and Yao Amber Li, "Credit constraints, quality, and export prices: Theory and evidence from China," *Journal of Comparative Economics*, 2015, *43(2)*, 390 – 416.

Fisher-Vanden, Karen, and Gary H. Jefferson, "Technology Diversity and Development: Evidence from China's Industrial Enterprises," *Journal of Comparative Economics*, 2008, *36*, 658 – 672.

Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, "Do Institutions Cause Growth?," *Journal of Economic Growth*, 2004, *9*, 271 – 303.

Griliches, Zvi, "Issues in Assessing the Contribution of Research and Development to Productivity Growth," Bell Journal of Economics," *Bell Journal of Economics*, 1979, 10(1), 92 – 116.

Griliches, Zvi, "The search for R&D spillovers," *The Scandinavian Journal of Economics*, 1992, 94, 29 – 47.

Grossman, Gene, and Elhanan Helpman, "Innovation and Growth in the Global Economy," *MIT Press Cambridge*, 1991.

Hall, Bronwyn, Jacques Mairesse, and Pierre Mohnen, "Measuring the Returns to R&D," *Handbook of the Economics of Innovation*, 2010, 2, 1033 – 1082.

Hansson, Pa, and Nan Nan Lundin, "Exports as an Indicator on or Promoter of Successful Swedish Manufacturing Firms in the 1990s," *Review of World Economics*, 2004, *140*, 415 – 445.

Henderson, Vernon J., Adam Storeygard, and David N. Weil, "Measuring Economic Growth from Outer Space," *American Economic Review*, 2012, *102*, 994 – 1028.

Henderson, Vernon J., Tim Squires, Adam Storeygard, and David N. Weil, "The Global Distribution of Economic Activity: Nature, History, and the Role of Trade," *Quarterly Journal of Economics*, 2017, Forthcoming.

Hsieh, Chang-Tai, and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124*, 1403 – 1448.

Jedwab, Remi, and Adam Storeygard, "Economic and Political Factors in Infrastructure Investment: Evidence from Railroads and Roads in Africa 1960 - 2015," *Working Paper*, 2017.

Jones, Charles, "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 1995, 113(4), 759 – 784.

Jones, Charles, "Time Series Tests of Endogenous Growth Models," *The Quarterly Journal of Economics*, 1995, 110(2), 495 – 525.

Jones, Charles, "Handbook of Economic Growth," *Handbook of Economic Growth*, 2005, 1B, 1063 – 1111.

Keller, Wolfgang, "Geographic Localization of International Technology Diffusion," *American Economic Review*, 2002, *92(1)*, 120 – 142.

Keller, Wolfgang, and Carol H. Shiue, "Market Integration and Economic Development: A Long-run Comparison," *Review of Development Economics*, 2007, *11*, 107 – 123.

Keller, Wolfgang, "International Trade, Foreign Direct Investment, and Technology Spillovers," *Handbook of the Economics of Innovation*, 2010.

Keller, Wolfgang, Ben Li, and Carol H. Shiue, "Shanghai's Trade, China's Growth: Continuity, Recovery, and Change Since the Opium Wars," *IMF Economic Review*, 2013, *61*, 336 – 37.

Lessmann, Christian, and André Seidel, "Regional Inequality, Convergence, and its Determinants -–A View from Outer Space," *European Economic Review*, 2017, *92*, 110 – 32.

Liang, Fangzhong, "Statistics of Hukou, Farmland and Land Tax in the History of China," *Zhonghua Book Company*, China: Beijing, 2008.

Manova, Kalina, "Volatility and Growth: Credit Constraints and the Composition of Investment," *Journal of Monetary Economics*, 2010, *57*, 246 – 265.

Melitz, Marc, and Stephen Redding, "Heterogeneous Firms and Trade," *Handbook of International Economics*, 2014, *4(4)*, 1 – 54.

Naughton, Barry J., "The Chinese Economy: Transition and Growth." *MIT Press*, MA: Cambridge, 2006.

Nie, Huihua, Ting Jiang, and Rudai Yang, "A Review and Reflection on the Use and Abuse of Chinese Industrial Enterprises Database," *World Economy*, 2012, 5, 142 – 158.

Pomeranz, Kenneth, "The Great Divergence: China, Europe, and the Making of the Modern World Economy," *Princeton University Press*, NJ: Princeton, 2009.

Roberts, Mark, and James Tybout, "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 1997, *87(4)*, 545 – 564.

Romer, Paul, "The Problem of Development: A Conference of the Institute for the Study of Free Enterprise Systems," *The Journal of Political Economy*, 1990, *98(5),2*, S71 – S102.

Rowe, William T., "China's Last Empire: the Great Qing," *Harvard University Press*, MA: Cambridge, 2010.

Spolaore, Enrico, and Romain Wacziarg, "How Deep are the Roots of Economic Development?," *Journal of Economic Literature*, 2013, *51*, 325 – 369.

Sun, Peng and Almas Heshmati, "International Trade and its Effects on Economic Growth in China," *IZA Discussion Papers*, 2010.

Romer, Paul, "Endogenous Technological Change," *Journal of International Trade and Economic Development*, 2008, 17, 105 – 133.

Unel, Bulent, "R&D Spillovers through Trade in a Panel of OECD Industries," *Journal of International Trade and Economic Development*, 2008, 17, 105 – 133.

Unel, Bulent and Harm Zebregs, "The Dynamics of Provincial Growth in China: A Nonparametric Analysis," *IMF Staff Papers*, 2009, *56*, 239 – 62.

Unel, Bulent, "The Interaction Between Technology Adoption and Trade When Firms are Heterogeneous," *Review of International Economics*, 2013, *21*, 797 – 808.

Unel, Bulent, "Human Capital Formation and International Trade," *The B.E. Journal of Economic Analysis & Policy*, 2015, *15*, 1067 – 1092.

Xue, Guoping, "Zhongguo Gu Jin Di Ming Dui Zhao Biao," *Shanghai Lexicographic Publishing House*, China: Shanghai, 2010.

Yi, Gang, "The Intrinsic Logic of China's Banking Industry Reform," *Transforming the Chinese Economy*, Edited by C. Fang, *Brill Online Publications*, 2010.

Young, Alwyn, "Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period," *Journal of Political Economy*, 2003, *111*, 1220 – 1261.

Zhu, Xiaodong, "Understanding China's Growth: Past, Present, and Future," *Journal of Economic Perspectives*, 2012, *26*, 103 – 124.

Appendix A. The Definitions of Dependent Variables

In this appendix, the definitions of dependent variables, first-time export entry and exit decision, are introduced. *first-time export entry* is a dummy variable which indicates a firm's first-time export if it equals 1. More details about this first-time export indicator are as follows. First, for firms founded before 2003, since the data before 2003 is not available in this research, I define *first-time export entry* following these rules: if the firm did not export during two previous observable years, and it started exporting from the third observable year in this data set, the dummy variable *first-time export entry* is set to 1 for the third year and all other years get 0. For example, for a firm founded in 1998 and shown in the data since 2003, if the firm did not export during 2006 and 2007, but started to export in 2008, *first-time export entry* equals 1 in 2008, while it equals 0 in all other years in the data set. Since I do not know whether the firm exported or not before 2003, I simply assume the firm "forgets" its exporting experience in two years once it stops exporting. In other words, it is assumed that the experience of exporting has a diminishing impact on a firm's future exporting behavior. In addition, for the firms founded before 2003 and exported in 2003, which is the first observable year in our data set, I define *first-time export entry* as 1 in 2003 and 0 for rest of the years.

For those firms founded in or after 2003, there could be two cases in defining the variable of *first-time export entry*. Firstly, if the firm is a state-owned enterprise (SOE) or crosses the threshold of five million Chinese Yuan nominal revenue in its initial year, the firm is included and shows full historical observations in this data set if its revenue does not fall below five million Chinese Yuan in the following years. For this type of firms, its *first-time export entry* equals 1 in the year when the firm started exporting for the first time. Second, for those firms founded in or after 2003 but controlled by non-state shareholders or with less than five million Chinese Yuan nominal revenue in their initial years, the data set may not include their full export history. In this case, *first-time export entry* equals 1 for the firm's first-time export entry indicated in the data set by judgment arbitrarily. That is to say, I assume such firms did not export before they reach the scale of five million Chinese Yuan nominal revenue. For example, if a firm founded in 2002 enters the data set initially in 2005, and started exporting for the first time in the data set in 2007, only the *first-time export entry* in 2007 is equal to 1, while the rest of years are 0.

The dummy variable *first-time export exit* indicates that a firm stops exporting for the first time in the data set. Similar to the problems mentioned in defining the first-time export entry dummy, firms founded before 2003, or non-SOEs founded in or after 2003 but with less than five million Chinese Yuan nominal annual revenue may not have full history in export. Following the similar rules in defining *first-time export entry*, the *first-time export exit* is set to 1 for a firm in the following cases. For firms founded before 2003, *first-time export exit* is equal to 1 in the year when it stopped exporting and only if it had exported continuously in the two previous observable years. Only the first year of non-exporting will count and *first-time export exit* in the rest of years are set to 0.

For those firms founded in or after 2003, the SOEs and non-SOEs earning five million Chinese Yuan from the very initial year have full operating history in the data set. For firms with full operational history, their first non-exporting year right after an exporting year is set to 1 when defining *first-time export exit*. For other firms founded in or after 2003, the same rule mentioned above is employed, which simply assumes that these firms did not export before they entered the data set. Take a firm founded in 2003 as an example, if this firm is a state-owned enterprise or achieved five million Chinese Yuan nominal revenue in 2003, and started exporting in 2006 and became a non-exporter in 2007, *first-time export exit* is equal to 1 in 2007 for this firm, while the rest of the years' are 0. Similarly, if a specific firm was founded in 2002 and entered the data set in 2005 and shows full history after 2005, if further this firm started exporting from 2005 and became non-exporter in 2010, then *first-time export exit* is 1 in 2010, while the rest of years' *first-time export exit* are set to 0. However, if a firm never stops exporting or never starts exporting during the data set period from 2003 to 2011, *first-time export exit* is defined as 0 for all its observable years.

The reasons why two years export experience before *first-time export entry* or *first-time export exit* is used as the cut-off point are based on findings from the following research. Firstly, Arnold and Hussinger (2005) reveal that a typical firm may "forget" a fairly amount of trade experience in two years by using the German data for years 1992 to 2004, which verifies the assumption of diminishing impact of export experience. Hansson and Lundin (2004) apply Swedish data for years 1990 to 1999 and conclude similarly. Comparably, Bernard and Jensen (2004) only find strong connection between export experience within two years and export behavior using the US data. Secondly, with regard to the scope of observations, if the judging criteria is set to three continuous years of export or non-export instead of two years before a firm changes its export behavior, more than 60% of the available observations will be lost, which leaves a small sub-sample and limited firm coverage to this research. This may potentially introduce unexpected bias.

Appendix B. The Cases of Categorical Variables

Table B.1: Impact of R&D on the Introduction of New Product

	1	2	3	4	5
R&D (t-1)	0.0388***	0.0348***	0.0319***	0.0316***	0.0311***
	(0.0022)	(0.0021)	(0.0020)	(0.0232)	(0.0020)
R&D (t-2)	0.0283***	0.0262***	0.0238***	0.0233***	0.0228***
	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)
Training (t-1)		0.0142***	0.0078***	0.0068***	0.0053***
		(0.0015)	(0.0013)	(0.0013)	(0.0012)
Output (t-1)			0.0274***	0.0277***	0.0280***
			(0.0024)	(0.0023)	(0.0024)
Age				0.0011***	0.0009***
				(0.0001)	(0.0001)
Collective					−0.0364***
					(0.0063)
Private					−0.0354***
					(0.0072)
HK-Macao-TW					−0.0555***
					(0.0127)
Foreign					−0.0447***
					(0.0153)
Adj. R^2	0.3511	0.3564	0.3647	0.3669	0.3688
Obs	23,511	23,479	23,474	23,474	23,474

Notes: R&D, Job Training and Output are in the logarithmic form. New Product is discrete variable. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table B.2: Impact of R&D on the Introduction of New Product, Heterogeneity by Ownership

	SOEs			non-SOEs		
	1	2	3	4	5	6
R&D (t-1)	0.0377***	0.0345***	0.0337***	0.0253***	0.0233***	0.0230***
	(0.0024)	(0.0023)	(0.0023)	(0.0035)	(0.0035)	(0.0035)
R&D (t-2)	0.0243***	0.0215***	0.0207***	0.0293***	0.0282***	0.0280***
	(0.0021)	(0.0021)	(0.0021)	(0.0045)	(0.0044)	(0.0044)
Training (t-1)	0.0146***	0.0070***	0.0043***	0.0128***	0.0088***	0.0083***
	(0.0018)	(0.0015)	(0.0015)	(0.0023)	(0.0023)	(0.0023)
Output (t-1)		0.0305***	0.0307***		0.0200***	0.0195***
		(0.0028)	(0.0028)		(0.0033)	(0.0033)
Age			0.0008***			0.0011**
			(0.0001)			(0.0004)
Ownership			*Y*			*Y*
Adj. R^2	0.3908	0.4001	0.4037	0.2332	0.2389	0.2400
Obs	16,410	16,406	16,406	6,971	6,970	6,970

Notes: R&D, Job Training and Output are in the logarithmic form. New Product is discrete variable. Columns (1), (2) and (3) report the results based on SOEs. Columns (4), (5) and (6) report the results based on non-SOEs. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table B.3: Impact of R&D on the Introduction of New Product in Industry 1

	All	SOEs	non-SOEs
R&D (t-1)	0.0318***	0.0322***	0.0274**
	(0.0059)	(0.0076)	(0.0114)
R&D (t-2)	0.0227***	0.0263***	0.0172
	(0.0065)	(0.0081)	(0.0168)
Training (t-1)	0.0024	0.0039	0.0076
	(0.0033)	(0.0042)	(0.0079)
Output (t-1)	0.0151**	0.0122	0.0085
	(0.0066)	(0.0087)	(0.0115)
Age	−0.0004	−0.0007	−0.0004
	(0.0005)	(0.0006)	(0.0009)
Adj. R^2	0.2306	0.2593	0.0521
Obs	2,184	1,533	558

Notes: R&D, Job Training and Output are in the logarithmic form. New Product is discrete variable. Ownership is controlled for across models. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table B.4: Impact of R&D on the Introduction of New Product in Industry 2

	All	SOEs	non-SOEs
R&D (t-1)	0.0420***	0.0468***	0.0320***
	(0.0080)	(0.0090)	(0.0120)
R&D (t-2)	0.0148**	0.0112	0.0088
	(0.0068)	(0.0090)	(0.0162)
Training (t-1)	0.0040	0.0052	0.0048
	(0.0035)	(0.0053)	(0.0048)
Output (t-1)	0.0257***	0.0319***	0.0147**
	(0.0068)	(0.0088)	(0.0061)
Age	0.0009*	0.0008	0.0014*
	(0.0005)	(0.0008)	(0.0007)
Adj. R^2	0.2446	0.2955	0.1102
Obs	3,158	1,834	1,243

Notes: R&D, Job Training and Output are in the logarithmic form. New Product is discrete variable. Ownership is controlled for across models. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table B.5: Impact of R&D on the Introduction of New Product in Industry 3

	All	SOEs	non-SOEs
R&D (t-1)	0.0198***	0.0203***	0.0195***
	(0.0026)	(0.0032)	(0.0047)
R&D (t-2)	0.0251***	0.0257***	0.0173***
	(0.0030)	(0.0036)	(0.0048)
Training (t-1)	0.0049***	0.00270	0.0111***
	(0.0018)	(0.0019)	(0.0028)
Output (t-1)	0.0226***	0.0268***	0.0125***
	(0.0035)	(0.0041)	(0.0048)
Age	0.0011***	0.0012***	0.0004
	(0.0002)	(0.0002)	(0.0007)
Adj. R^2	0.3014	0.3282	0.1862
Obs	9,554	6,810	2,647

Notes: R&D, Job Training and Output are in the logarithmic form. New Product is discrete variable. Ownership is controlled for across models. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Table B.6: Impact of R&D on the Introduction of New Product in Industry 4

	All	SOEs	non-SOEs
R&D (t-1)	0.0363***	0.0415***	0.0230***
	(0.0026)	(0.0029)	(0.0058)
R&D (t-2)	0.0229***	0.0185***	0.0360***
	(0.0033)	(0.0032)	(0.0079)
Training (t-1)	0.0078***	0.0075***	0.0107**
	(0.0023)	(0.0028)	(0.0051)
Output (t-1)	0.0373***	0.0368***	0.0323***
	(0.0039)	(0.0042)	(0.0066)
Age	0.0009***	0.0007***	0.0021**
	(0.0002)	(0.0003)	(0.0008)
Adj. R^2	0.4313	0.4677	0.3061
Obs	8,938	6,458	2,396

Notes: R&D, Job Training and Output are in the logarithmic form. New Product is discrete variable. Ownership is controlled for across models. All regressions include city fixed effects, 4-digit industry fixed effects, province-year fixed effects. Numbers in parentheses are standard errors, clustered at the 4-digit industry level; ***, **, and * represent statistical significance at the 1%, 5%, and 10% level, respectively.

Vita

Fan Duan was born in Yulin, Guangxi, China in 1986. He obtained his Bachelor's degree in Economics from the Guangxi University in 2008. In 2011, he completed his Master of Arts in Economics at Zhejiang University. After that, he entered the graduate school in the Department of Economics at Louisiana State University (LSU) and received his M.S. degree in Economics in 2015. During his graduate study at LSU, he worked as a research assistant and an instructor, teaching Principles of Macroeconomics. Fan Duan's primary research interest includes economic growth, regional economics, and international trade.