EXPORTING, IMPORTING AND THE HETEROGENEITY OF CHINESE MANUFACTURING FIRMS

by

LIYUN ZHANG

A thesis submitted to

the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

Birmingham Business School

College of Social Sciences

The University of Birmingham

September 2014

Abstract

This thesis studies the internationalization strategies of Chinese manufacturing firms. Using rich information from a merged dataset with survey data and trade data of Chinese manufacturing firms over the period 2002 to 2006, we study the firms' exporting and importing behaviour and their relationships with firm heterogeneity. I first investigate firms' decision to export taking into account firms' import activities, unobserved heterogeneity and initial conditions. Sunk entry costs and firms' importing experience and characteristics including age, size, productivity, labour-force quality, financial health and ownership are found to be significant determinants of firms' export behaviour. Compared to the extensive studies on firm-level exports, the importing side has been largely neglected. I then examine the casual relationship between importing and firm productivity using propensity score matching with difference-in-differences techniques. More productive firms are found to self-select into the import markets and at the same time participation in the import markets improves the firms' productivity. Finally, I explore the links between the two sides of firms' international trade activities. Estimates across different specifications show that access to foreign inputs improves firms' export performances. Importing more and better quality inputs raises the number of export varieties, export value and export quality.

I would like to dedicate this thesis to

the memory of my father, Zhiming Zhang (1957-2001),

who emphasized the importance of education.

I hope you would have been proud.

I love you dearly.

Acknowledgements

The completion of this doctoral thesis would not have been possible without the help and support of the kind people around me. It gives me great pleasure to extent my heartfelt gratitude.

Above all, I would like to thank my family for their unconditional love and constant encouragement in all my pursuits at all my good and bad times. I know I always have my family to count on when times are tough. Most of all I would like to thank my husband Dr. Shan He. Shan has faith in me and my intellect and has always been supportive of my work. I truly thank him for standing by my side. Shan has shared the many uncertainties, challenges and sacrifices since our son Yang was born. I also thank Yang, who has grown into a wonderful 2.5 years old in spite of his mother spending so much time away from him working on this thesis. I am proud of him.

I would like to express my deepest thanks and appreciation to my lead supervisor, Professor Robert Elliott, for his support for all stages of completing the research and writing the thesis. I appreciate all his contributions of ideas, inspiration, guidance and encouragement to make my PhD experience productive. It is no exaggeration to say that I got constructive suggestions from every meeting and conversation with and email from Rob. Rob has always encouraged, or even pushed, me to think deeper with the results which I found is profoundly beneficial for the development of my research skills. I have learned so much from his enthusiasm, originality and insights in research.

I am grateful to my second supervisor, Mr Nicholas Horsewood, for his good advice on some econometric part of my thesis. My sincere gratitude also goes to Dr. Liza Jabbour for her guidance and expertise in one of the main chapters of the thesis. The early success of that chapter has given me much confidence. Also, I am indebted to Dr. Ying Zhou, one of my best friends, who has helped a lot with the search for the data used in the thesis and shared her research experience with me over the past few years.

My sincere thanks also go to the examiners of the thesis, Professor Farid Toubal and Professor Alessandra Guariglia, for their time and effort, in particular their constructive comments and suggestions which help me improve the quality of the thesis.

I must acknowledge as well the many friends and fellow doctoral candidates who have encouraged, enlightened, and entertained me throughout the journey. Especially I would like to thank Dr. Di Wang, Dr. Churn Min Chuang, Dr. Dejing Kong, Dr. Min Tan, Dr. Attasuda Lerskullawat, Dr. Le Ho An Chau, Dr. Pei Ling Liao, Dr. Kwangjo Jeong, Bing Liu, Rong Huang, Theresa Beyrle, Enrico Vanino, Charles Rahal, Marco Giansoldati and Mary Dawood.

Finally, I gratefully acknowledge the PhD scholarship for this research provided by Department of Economics, University of Birmingham.

Contents

List of Tables

List of Abbreviations

AME	Average marginal effects
ATT	Average treated effects on the treated
BEC	Broad economic category
CML	Conditional maximum likelihood
CIC	China industrial classification
DID	Difference-in-differences
FE	Fixed effects
GDP	Gross domestic product
GMM	General method of moments
HS	Harmonized Coding System
NBS	National Bureau of Statistics of China
RE	Random effects
OLS	Ordinary least square
PSM	Propensity score matching
R&D	Research and development
RMB	Renminbi, Chinese official currency
SOE	State-owned enterprise
TFP	Total factor productivity
US	United States of America
WDI	World development indicator
WTO	World Trade Organization

Chapter 1

Introduction

This chapter introduces the background and research topics of the thesis. Section 1.1 discusses the general background of the research. Three research questions are proposed in Section 1.2 and an outline of the thesis is then presented in Section 1.3 with a brief introduction of the subsequent chapters. Finally, the contributions of the research are summarized in Section 1.4.

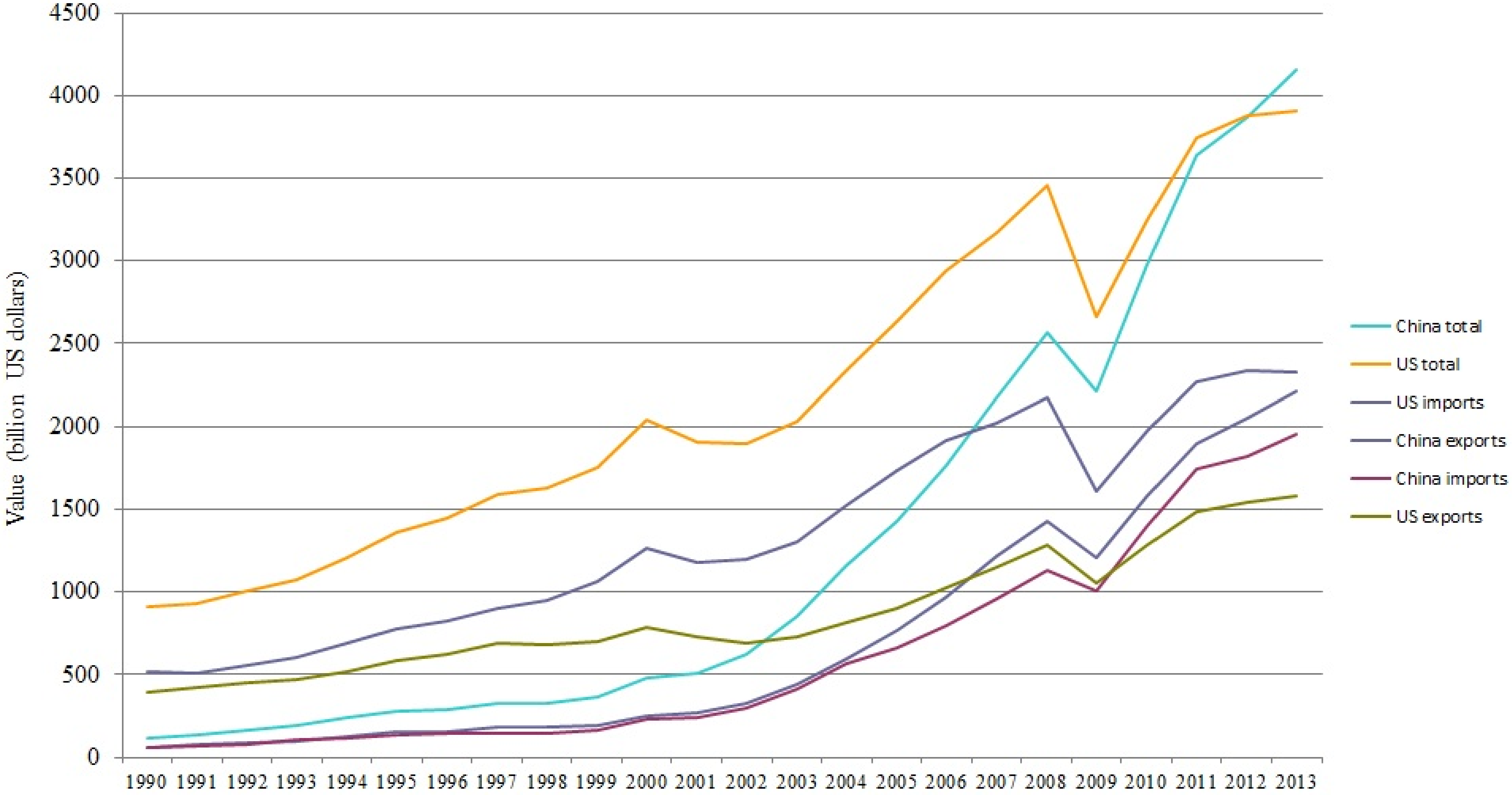
1.1 Background

"China is now No 1 in trade", proclaimed the front-page headline of *China Daily USA* on its January 13, 2014 issue (Li and He, 2014). Indeed, China's international trade in merchandise led the world in 2013 as its total value of exports and imports reached 4,159 million US dollars (USD), exceeding that of the US (3,909 million USD) for the first time (Statistics database, World Trade Organization (WTO)). Figure 1.1 plots the annual exports, imports and their total value of China and US for the period 1990 to 2013. Several features of Chinese and US international trade can be observed. Firstly, the US keeps a deficit of trade over the period while China maintains a surplus except 1993. Secondly, China's total trade value was about one eighth (12.7%) of that of the US in 1990 and the shares of China's total trade over that of the US reached over 50% in 2005 and 90% in 2010. Thirdly, China's trade with the world maintains a steady growth over the period except 2009 due to the hit from global economic crisis.

China's international trade has played an increasingly important role in its economic development over the past two decades and the trade with the world has been a major contributor of its gross domestic product (GDP). Table 1.1 reports the annual international trade value and GDP of China for the period 1990 to 2013. The share of China's exports and imports over its GDP increases from less than 30% in 1990 to over 65% in 2006. The shares drop during the global economic crisis period 2007 to 2009 and it stands around 50% from 2010.

Since international trade is most active in manufacturing industry, a good understanding of the dynamics of manufacturing firms' internationalization strategies is of high importance to a country whose economic development relies much on its international trade, such as China. I choose to study the exporting and importing behaviour of Chinese firms for this period for two reasons. Firstly, China entered the WTO in December 2011 after a long journey of 15 years' negotiation and China faced more opportunities as well as challenges immediately after that as it would enjoy the privileges and honour the promises at the same time as a WTO member. This is expected to have a great impact on Chinese manufacturing firms, no matter those that have been active in the international markets or those that are serving purely the domestic markets. Secondly, this period represents the beginning of the transition of Chinese firms from relatively low skilled assembly tasks to more high technology and high value added production. It is therefore useful to understand the roles of exports and imports in this transition process. Studies on costs and benefits for Chinese manufacturing firms to participate in international trade and the relationships between exporting, importing and firm heterogeneity are thus of interest to the economists and policy makers of China as well as its trade partner countries.

Figure 1.1: Exports and imports of China and US 1990-2013



Source: Statistics database, World Trade Organization (WTO)

Table 1.1: China's international trade and GDP 1990-2013

Year	1990	1991	1992	1993	1994	1995	1996	1997
Total trade	556.01	722.58	911.96	1,127.10	2,038.19	2,349.99	2,413.38	2,696.72
GDP	1,866.78	2,178.15	2,692.35	3,533.39	4,819.79	6,079.37	7,117.66	7,897.30
% (total trade/GDP)	29.78	33.17	33.87	31.90	42.29	38.66	33.91	34.15
Year	1998	1999	2000	2001	2002	2003	2004	2005
Total trade	2,684.97	2,989.62	3,927.32	4,218.36	5,137.82	7,048.35	9,553.91	11,692.18
GDP	8,440.23	8,967.71	9,921.46	10,965.52	12,033.27	13,582.28	15,987.83	18,493.74
% (total trade/GDP)	31.81	33.34	39.58	38.47	42.70	51.89	59.76	63.22
Year	2006	2007	2008	2009	2010	2011	2012	2013
Total trade	14,097.40	16,686.37	17,992.15	15,064.81	20,172.21	23,640.20	24,416.02	25,821.23
GDP	21,631.44	26,581.03	31,404.54	34,090.28	40,151.28	47,310.40	51,894.21	56,884.52
% (total trade/GDP)	65.17	62.78	57.29	44.19	50.24	49.97	47.05	45.39

Source: China Statistical Yearbook. Total trade value and GDP are in trillion RMB.

1.2 Research Questions

This thesis investigates the relationship between exporting, importing and heterogeneity of Chinese manufacturing firms. Three empirical studies are carried out in order to answer the following research questions.

The first research question emerges as the literature in international trade has documented that only a small proportion of firms engage in exporting. It is then of our interest with the question that whether the Chinese manufacturing firms have a similar pattern in exporting and if so, what factors determine Chinese manufacturing firms' decision to export? In order to answer the question, I test the existence of the sunk costs of entering the exports market and factors affecting the likelihood of firms export participation. Also as most of the existing literature on decision to export focus only one side of firms' international activities, i.e., exports and neglect the other side, imports, the first empirical study tries to fill the gap in the literature by exploring the extent to which being an importer affects the firms' probability of exporting.

Secondly, as extensive empirical studies indicate a positive relationship between exporting and firm productivity, and only three papers have looked at the relationship between firm imports and productivity, We are curious whether such finding on exporters can also be observed from importers. So the next research question is on the relationship between productivity and importing of Chinese manufacturing firms: whether more productive firms self-select into import markets, and (or) whether firms improve their productivity by learning from importing. It is the first study on this topic using Chinese firm-level data and findings will be compared with that from papers on other countries.

Finally, as documented in the international trade literature many firms that export also import at the same time, I would like to know whether there are any links between firms' exporting and importing activities. So the third research question of the research is to test whether the use of imported inputs affects firms' export performance such as export scopes and quality, and if so, to what extent.

1.3 Outline of the Thesis

The thesis is organized as follows. Chapter Two describes the data and key variables used for the empirical studies on the above mentioned research questions. The search for, and construction of the dataset has been a significant undertaking for the whole research programme, especially at the beginning of the research. Data sources, sample size and features of both datasets are described. Methods and criteria for the data construction are then presented with the method of estimation of total factor productivtity in the Appendix which is one of the key contribution of the research.

Chapter Three examines firms' export dynamics by examining the importance of sunk costs and firm heterogeneity in the decision to export. We emphasize the impact of firms' import activities, the other important side of international trade, on their entry to export markets. We also take into account the initial conditions problem when the lag of export status is included and possible correlation between the unobserved firm heterogeneity and other explanatory variables. Sunk costs and firms' importing experience and other heterogeneity including age, size, productivity, labour-force quality, ownership and location are found to be significant determinants of exporting behaviour.

In Chapter Four, I investigate the relationship between importing and firm productivity. Propensity score matching and matching difference-in-differences techniques are employed to examine the casual relationship between them. More productive firms are found to self-select into the import markets and at the same time participation in the imports market also improves the firms' productivity, indicating a learning effect from importing.

In Chapter Five, I try to establish the link between imported inputs and firms' export performance. Estimates from different specifications show that access to foreign inputs improves firms' export performances. Importing more inputs raise the scope of exports and better quality of imported inputs leads to quality upgrading of the exports.

Each of the above three chapters is a complete study when taken separately. Therefore, I deliberately keep a section on introduction to the data in the main text and the definition of variables in the Appendix in each of the empirical studies with details on the data and the construction of the key variables in Chapter Two.

Finally, Chapter Six concludes with a summary of the results for the empirical studies and a discussion of contributions, limitations of the thesis and potential future research.

1.4 Contributions of the Research

International trade has been a key component of economic growth to most of the economies, which is especially true for the emerging economies such as China. Although extensive research on international trade in country and industry level has been done, that on firm level has just began in recent years. Understanding the determinants of trade at firm level and the relationships between trading behaviour and firm characteristics are thus of policy makers' interests. The existing literature on firm-level international trade behaviour has been focusing on the development countries. This research tries to contribute to the literature by drawing the attention to the biggest developing country, China. The detailed contributions are as follows.

Firstly, in Chapter Three I provide evidence that firms' participation in imports market has a positive and significant effect on the propensity to export. Neglecting firms' imports activity overstates the size of sunk costs of exporting. Unobserved firm heterogeneity and the initial conditions problem should also be addressed properly when studying firms' exporting behaviour. This Chapter also highlights the potential role of government agencies in establishing and enhancing export promotion policies. Policies that reduce the sunk and fixed costs of firms' entering into foreign markets will be of great help to firms that are currently prevented from doing so because of the large entry costs.

Secondly, a bi-directional relationship between importing and productivity is found in Chapter Four when I test for Chinese manufacturing firms both the self-selection to importing and learning-from-importing hypotheses. Firms that source their inputs from high-income-economies and inputs with more skill and technology intensive contents exhibit stronger learning effects from importing. Policies on promoting international trade should consider the importing side as well.

Thirdly, the access to the imports market is found to have a positive and significant impact on firms' export performance in Chapter Five. Firms importing lager number of varieties and higher value of inputs expand their export scope and quality. Thus policies aiming for further trade liberalization will promote both the imports and exports of Chinese manufacturing firms.

Outcomes of this research are listed below:

(1). Firm Productivity and Importing: Evidence from Chinese Manufacturing Firms (with Robert Elliott and Liza Jabbour), *Canadian Journal of Economics* (revised and resubmitted).

(2). Dynamics of the Decision to Export and Import for Chinese Manufacturing Firms: Sunk Costs and Heterogeneity (with Robert Elliott and Nicholas Horsewood), to be submitted to *Review of International Economics*.

(3).Exporting and Firm Heterogeneity: An Empirical Analysis on Chinese Manufacturing, 4th International Forum for Contemporary Chinese Studies (IFCCS), Nottingham, UK, 2011.

(4). Decision to Export: Firm Heterogeneity and Initial Conditions Problem, 14th European Trade Study Group (ETSG) Annual Conference, Leuven, Belgium, 2012.

(5). Firm Productivity and Importing: Evidence from Chinese Manufacturing, 15th ETSG Annual Conference, Birmingham, UK, 2013.

(6). Dynamics of the Decision to Export for Chinese Manufacturing Firms: Sunk Costs and Heterogeneity, 16th ETSG Annual Conference, Munich, Germany, 2014.

Chapter 2

Data and Key Variables

This chapter describes the data and key variables used for the empirical studies in the thesis. Section 2.1 describes two sets of data, the firm-level production data and transaction-level trade data, including data sources, sample sizes and features. Section 2.1.3 presents the method and criteria for matching and cleaning both datasets and Section 2.2 introduces the key variables constructed for the empirical studies in the thesis.

2.1 Data

The data used in the thesis include two datasets: the firm-level production data and transaction-level trade data. The firm-level production data are from the Annual Survey of Industrial Enterprises provided by the National Bureau of Statistics of China (NBS) and the transaction-level trade data are provided by the Department of Customs Trade Statistics, the General Administration of Customs of China. We use data from both sources over the period from 2002 to 2006.

2.1.1 Firm-level survey data

The NBS data cover all state-owned industrial enterprises and non-state-owned industrial enterprises with annual sales of greater than 5 million *Renminbi* (RMB, the Chinese official currency).[1] According to the NBS industry classification, industrial enterprises refer to enterprises that operate in the mining and quarrying sector, manufacturing sector or in the production and supply of power, gas and water.

According to the Chinese law, it is mandatory for the enterprises which reach up the standards to report their basic information and relevant accounting data to local statistics bureaus according to the requirements set by the NBS every year. The NBS then assembles these enterprises' information from all provinces, municipal cities and autonomous regions every financial year. The NBS survey is the primary source for the construction of numerous aggregate statistics used in the China Statistical Yearbooks. The NBS survey data provide detailed information of each enterprise covered, including the firm's identification (tax code) and basic information such as year founded, location, ownership type, employment, China industrial classification (CIC) code as well as over 50 financial variables from the accounting statements, including current and fixed capital, current and fixed assets, liabilities, industrial value-added, annual sales, total income, profits, short and long term investment, current and accumulated depreciation,

[1] The official RMB exchange rate per US dollar between 2002 and 2004 was 8.277, 8.194 in 2005 and 7.973 in 2006 (World Development Indicators, World Bank). Hence, the threshold for inclusion in the dataset is equivalent to between USD600,000 and USD627,000 over the sample period.

total wage bill and R&D expenses.

2.1.2 Transaction-level trade data

Trade data record all import and export transactions with non-zero values that enter or exit through one of the Chinese customs. Each observation represents a shipment and contains detailed information on the time of the transaction (month and year), type of trade (import/ export), exporting/ importing firm identifier, ownership type, product traded (8-digit Harmonized Commodity Description and Coding System (HS) code and name), value, quantity, unit, destination country (of the exported commodities) / country of origin (of the imported products), type of trade (ordinary trade, processing trade, compensation trade, consignment, etc.) and finally mode of transport.

We now have a brief look at the trade data for 2006. The dataset includes a total sample size of 25,661,754 observations, each of which represents a shipment. There are 208,425 firms involved in the international trade with 121,835 firms participating in the importing, 171,205 in exporting and 84,615 firms in both. The total value of the Chinese international trade of 2006 is USD1756.85 billion, with USD968.51 billion for exports and USD788.34 billion for imports respectively.

2.1.3 Data Cleaning and Matching

Both the datasets provide rich information on firm production and firm international activities. The firm-level survey data include firms in mining, manufacturing and public utilities industries and transaction-level trade data record firms that have transactions across one of the Chinese customs. In order to explore the relationships between firm characteristics and their international trade activities, we match the survey data and the trade data over the period. Following Brandt et al. (2012), we first link firms from each year of the survey data using firm registration identification numbers and names and drop observations with duplicate firm IDs or names. To avoid complications involved with firm birth or death, we select firms that continuously operate for the whole sample period.[1] For the trade data, as the data are recorded in monthly frequency, we merge each of the five years' data into annual data by aggregating the observations in the firm-HS8-partner country level. Since the survey and trade data use completely different coding systems for the firm identifier, we cannot merge both datasets by firm IDs directly. Hence, matching is done by using names and registration places of the firms. We concentrate only on firms in the manufacturing sector (CIC13-43) as the manufacturing sector produces most tradable goods and the Chinese manufacturing sector is the most important sector for China's international trade.

By merging the firm-level survey data with customs data, we have selected 56,836 firms that appear throughout the five-year period in both datasets, with 284,180 observations in total. Furthermore, we exclude firms with incomplete records or negative values of key variables such as firm age, assets, real capital stock, number of employees, output, value-added and total wages. 16,354 observations are dropped for missing firm age, assets, real capital stock, number of employees, output or total wages. 20,701 observations are dropped for missing or negative values of firm age, number of employees, total wages, total assets and depreciation. Observations that are found with each of the following: total assets smaller than fixed assets; total

[1] Bernard and Jensen (2004), Girma et al. (2004), Arnold and Hussinger (2005), and Elliott and Zhou (2013) use a similar rule in their data construction.

assets smaller than current assets; total liability smaller than current liability; current depreciation smaller than accumulated depreciation, are regarded as abnormal. 1,716 observations are thus dropped. Finally 27,224 observations are dropped due to discontinuity over the period.[1] Our final sample is a balanced panel with 56,836 firms per year over the five year period 2002 to 2006, corresponding to 284,180 firm-year observations.

2.2 Key Variables

As shown above, both the survey and trade data are used in the subsequent chapters of the thesis, the key variables used are firm characteristic variables and trade variables. Firm characteristic variables are firm age, employment, wages of employees, total factor productivity (TFP), labor productivity, capital intensity, liquidity ratio, leverage ratio, ownership type, size, industrial and regional dummies. Nominal variables are deflated by 4-digit CIC industry deflators constructed by Brandt et al. (2012).[2] Firm ownership, size and regional dummies construction and TFP estimation are introduced as follows.

Firms are classified into different categories of ownership by the shares of capital from investors of the firms.[3] In this study, firms are categorized into five ownership types according to the share of capital of different investors. Firstly, firms are grouped into two main categories: domestic-funded and foreign-owned firms. Foreign-owned firms include two types of firms: Hong Kong, Macau or Taiwan-owned firms (*HMT*) and other foreign-owned firms (*FOREIGN*).

[1] As explained above, we construct a balanced panel. These observations are dropped since they are firms not appearing throughout the period.

[2] http://www.econ.kuleuven.be/public/n07057/China/.

[3] Official definition and classification of enterprises by registration status can be found on the NBS website: http://www.stats.gov.cn/statsinfo/auto2073/201310/t20131031_450535.html. See Kanamori and Zhao (2004, Appendix 1) for the English version of the definition and classification.

According to the standard of NBS, *HMT* firms are those more than 25% of whose registration capital is from Hong Kong, Macau or Taiwan investors and *FOREIGN* firms are those with more than 25% of capital from foreign other than Hong Kong, Macau or Taiwan investors. Domestic firms include those that are not categorized as *HMT* or *FOREIGN* firms. Domestic firms are further divided into three groups: state-owned enterprises (*SOE*), collectively-owned enterprises (*COLLECTIVE*) and private firms (*PRIVATE*). SOEs are state sole funded corporations and enterprises whose majority shares belong to the government or other SOE. Firms in which the state or SOE-owned share is less than 50 percent, as long as the state or SOE has controlling influence over management and operation are counted as SOEs. Collective firms are economic units where the assets are owned collectively, such as township and village enterprises. Private firms refer to profit-making economic units invested and established by natural persons, or controlled by natural persons using employed labour. Similar to SOEs, firms whose majority shares belong to collective (private) investors are categorized as *COLLECTIVE*(*PRIVATE*).

We follow the classification standard adopted by NBS and group our sample into three size categories: *LARGE*, *MEDIUM* and *SMALL*. Small firms are those with less than 300 employees or 30 million RMB sales or 40 million RMB total assets; medium-sized firms are those with between 300 and 2,000 employees and between 30 million and 300 million RMB sales and between 40 million and 400 million RMB total assets and large firms are those with more than 2,000 employees and 300 million RMB sales and 400 million RBM total assets. Due to the huge differences in infrastructure, economic development and education level between the eastern and western regions of China, we allocate each firm into one of three regions, *EAST*, *MIDDLE* and *WEST*, according to their location.

We calculate TFP using two methods: a traditional one by Levinsohn and Petrin (2003) and a recent one developed by De Loecker (2007). De Loecker (2007)'s TFP measurement is built on Olley and Pakes (1996) and includes firms' export status as an additional state variable to the production function as he argues that exporting firms face different market structures and factor prices and by introducing export in the estimation procedure, the decision to invest and to exit the market depends on export status.[1] We extent the method adding not only firms' export status, but also their import status to the production function.[2] The reason is that just as exporters, importing firms also face different market structure and factor prices from non-importing firms. Because ownership structure of a firm may influence its input decisions, ownership dummies are also included in the production function. Furthermore, since firms in different industries have different factor inputs and face different input prices, we estimate the production function for each 2-digit industry separately rather than doing this for the entire manufacturing sector.

Trade variables include export/ import status dummies, export/ import values, number of products traded, partner countries and varieties. Detailed definition of the variables used in each empirical study is presented in the subsequent chapters.

2.3 Overview of China's International Trade

Over the sample period, China is actively participating in the global value chain with a steady growth in both its exports and imports. However, its international trade is unevenly distributed

[1] Van Biesebroeck (2005) also includes export status in the production function.

[2] Kasahara and Rodrigue (2008) include import status and Bas and Strauss-Kahn (2014) use the number of imported inputs in their TFP estimation. include Studies that includes both export and import status for the TFP measurement are Amiti and Konings (2007) and Yu (2014).

across the industries, regions and firm ownership types.

Although international trade is most active in manufacturing sector, the trading activities are not equally distributed throughout the whole manufacturing sector over the sample period studied. China's international trade is most concentrated in textile industry and textile wearing apparel, footware and caps industry, followed by communication equipment, computer industry and electrical equipment industry. These industries are labour-intensive industries and China is making use of its advantage of abundant labour in importing materials, processing and exporting final goods. Least international trade activities are conducted in tobacco industry, followed by petroleum and nuclear fuel processing industry which are the industries protected by the state. Less than 1% of the international trade is generated in these industries.

Also, China's international trade is unbalanced between the East and West regions. Among the 30 provinces and municipalities, coastal provinces in the East region including Guangdong, Zhejiang, Jiangsu, Shanghai and Shandong are the most active provinces and municipalities. International trade in these provinces accounts about 75% of the total international trade of China. Underdeveloped provinces in the West region such as Shanxi, Xinjiang and Tibet are the least active provinces in international trade which contributes less than 0.2% of the country's international trade.

In terms of firm ownership types, foreign-owned firms are the biggest contributors of China's international trade, followed by HMT firms. International trade conducted by FOREIGN and HMT firms takes up about 68% of its total international trade. The reasons maybe that FOREIGN and HMT firms have the advantage of better access to foreign market information (on both suppliers and buyers), distribution network and more advanced technology from their parent companies. SOEs and collective firms are the least active firms in international trade, which contributes about 5% to the total. One interesting finding is that about 27% of China's international trade is carried out by the private firms which are becoming another mainstream of China's international trade.

In the subsequent chapters, both sides of China's international trade in manufacturing industry, i.e., exports and imports, are studied taking into consideration of firm characteristics, in particular TFP and ownership.

Appendix 2A

Total Factor Productivity Estimation

In this study we measure TFP by modifying the De Loecker (2007) approach. De Loecker (2007) extends Olley and Pakes (1996) framework by allowing market structure to be different for exporting firms by introducing export into the production function. We modify De Loecker (2007) method by introducing both export and import status into the underlying structural model. A number of recent papers have suggested that not only entry to the export markets, but also the imports market involves sunk costs and more productive firms usually self-select into importing and/or exporting (Das et al., 2007; Kasahara and Lapham, 2013; Van Biesebroeck, 2005), hence it is more appropriate to treat the decision to import or to export as endogenous when measuring firms' productivity.

As in the standard Olley and Pakes (1996) model, a firm is assumed to be risk-neutral and to maximize its expected value of both current and future profits. The production function is set up as the follows:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it} \tag{2A.1}$$

where i and t denote firm and time, y, k and l are the logs of output as measured by value added, capital input and labor input respectively, ω is the transmitted productivity shock which impacts the firm's decision rules and η is an i.i.d component which is uncorrelated with input choices.

At the beginning of every period, the firm makes the following decisions. First, it makes a discrete decision to continue its operation or exit by comparing the continuation value with a one-time sell-off value. Secondly, conditional on staying in operation, it decides the level of labor input (*l*) and investment (*i*). Capital is accumulated according to the law of motion $k_{it} = (1-\delta)k_{it} + i_{it}$ and it is assumed that investment in the current period becomes productive the next period.

The demand for investment is a function of the firm's capital *k*, productivity ω and two additional variables, export status EXP_{it} and import status IMP_{it}:

$$i_{it} = i_t(k_{it}, \omega_{it}, EXP_{it}, IMP_{it}) \tag{2A.2}$$

As investment is a control variable, it is costly to adjust and researchers often come across data with a substantial observations of zero investment. To circumvent the problem of firms with zero investment, Levinsohn and Petrin (2003) suggest a modification of the Olley and Pakes (1996) approach by using intermediate inputs (m), such as materials or energy usage, instead of investment, as a proxy variable to recover the unobserved firm productivity. Since intermediate inputs are not typically state variables and it is less costly to adjust intermediate inputs they may respond more fully to productivity shocks. Equation (2A.2) then becomes:

$$m_{it} = m_t(k_{it}, \omega_{it}, EXP_{it}, IMP_{it}) \tag{2A.3}$$

We can invert this demand function to obtain the productivity shock ω_{it} as given by:

$$\omega_{it} = \omega_t(k_{it}, m_{it}, EXP_{it}, IMP_{it}) \tag{2A.4}$$

Substituting ω_{it} with (2A.4) to the production function in (2A.1), we have:

$$y_{it} = \beta_0 + \beta_l l_{it} + \phi_t(k_{it}, m_{it}, EXP_{it}, IMP_{it}) + \eta_{it} \tag{2A.5}$$

where $\phi_t(k_{it}, m_{it}, EXP_{it}, IMP_{it}) = \beta_k k_{it} + \omega_t(k_{it}, m_{it}, EXP_{it}, IMP_{it})$.

In the first stage, OLS can be used to obtained a consistent estimate of β_l from Equation (2A.5) by substituting a third-order polynomial in the three variables, k_{it}, m_{it}, EXP_{it} and IMP_{it}, to approximate $\phi_t(.)$. Estimation is done industry by industry, adding the ownership and year dummies to capture the ownership and time effects.

In the second stage, the capital coefficient β_k is estimated as the follows.

To correct the selection bias, the survival decision depends on export and import status through the productivity shock and through the capital accumulation process. If we define the indicator function χ_{it} to be equal to one if firm i continues in operation at t and zero if it exits, then the survival probability is determined on the information set J at time t by:

$$
\begin{aligned}
& Pr\{\chi_{i(t+1)} = 1 | J_{it}\} \\
= \;& Pr\{\omega_{i(t+1)} \geq \underline{\omega}_{i(t+1)}(k_{i(t+1)}, EXP_{it}, IMP_{it}) | \underline{\omega}_{i(t+1)}(k_{i(t+1)}, EXP_{it}, IMP_{it}), \omega_{it}\} \\
= \;& \psi\{\underline{\omega}_{i(t+1)}(k_{i(t+1)}, EXP_{it}, IMP_{it}), \omega_{it}\} \\
= \;& \psi_{it}(k_{it}, m_{it}, EXP_{it}, IMP_{it}) \equiv P_{it} \qquad (2A.6)
\end{aligned}
$$

We assume that productivity follows a first order Markov process:

$$\omega_{i(t+1)} = E[\omega_{i(t+1)} | \omega_{it}, EXP_{it}, IMP_{it}, \chi_{i(t+1)} = 1] + \xi_{i(t+1)} \qquad (2A.7)$$

where $\xi_{i(t+1)}$ is the innovation in productivity for the next period which depends on current productivity, export and import status and survival in the next period.

Now consider the expectation of $y_{i(t+1)} - \beta_l l_{i(t+1)}$ conditional on information at t and survival:

$$
\begin{aligned}
& E[y_{i(t+1)} - \beta_l l_{i(t+1)} | k_{i(t+1)}, \chi_{i(t+1)} = 1] \\
= \;& \beta_0 + \beta_k k_{i(t+1)} + E[\omega_{i(t+1)} | \omega_{it}, EXP_{it}, IMP_{it}, \chi_{i(t+1)} = 1] \\
\equiv \;& \beta_k k_{i(t+1)} + g(\omega_{i(t+1)}, \underline{\omega}_{it}) \qquad (2A.8)
\end{aligned}
$$

Provided the density of $\omega_{i(t+1)}$ conditional on ω_{it} is positive in a region, $\omega_{i(t+1)}$ can be written as a function of P_{it} and ω_{it} from the survival equation in (2A.6). We then can write $g(.)$ as a function of P_{it} and ω_{it}.

Substituting P_{it} and ω_{it} into $g(.)$, we have the following from Equation (2A.1):

$$
\begin{aligned}
& y_{i(t+1)} - \beta_l l_{i(t+1)} \\
= \quad & \beta_0 + \beta_k k_{i(t+1)} + E[\omega_{i(t+1)} | \omega_{it}, IMP_{it}, \chi_{i(t+1)} = 1] + \xi_{i(t+1)} + \eta_{it} \\
= \quad & \beta_0 + \beta_k k_{i(t+1)} + g(P_{it}, \omega_{it}) + \xi_{i(t+1)} + \eta_{it}
\end{aligned}
\tag{2A.9}
$$

Using $\omega_{it} = \phi_t(k_{it}, m_{it}, EXP_{it}, IMP_{it}) - \beta_k k_{it}$ from (2A.5), we rewrite the first three terms of the right-hand side of Equation (2A.9) as a function of $\phi_t - \beta_k k_{it}$ and P_{it}:

$$
y_{i(t+1)} - \beta_l l_{i(t+1)} = \beta_0 + \beta_k k_{i(t+1)} + g(P_{it}, \phi_{it} - \beta_k k_{it}) + \xi_{i(t+1)} + \eta_{it} \tag{2A.10}
$$

A consistent estimate of β_k is obtained by running nonlinear least squares on Equation (2A.10) by substituting the coefficient on labor β_l obtained from the first stage, as well as the survival probability P_{it} estimated from Equation (2A.6). As in the first stage of the estimation procedure, the function $g(P_{it}, \phi_{it} - \beta_k k_{it})$ is approximated using a higher order polynomial expansion in P_{it} and $(\phi_{it} - \beta_k k_{it})$.

Chapter 3

Dynamics of Firms' Decision to Export: Sunk Costs and Heterogeneity

In this chapter I study the dynamics of firms' decision to export taking into account their importing activity together with firm heterogeneity and controlling for the initial conditions problem. These are all issues that the existing literature on firms' exporting behaviour often neglects. Specifically, I allow for possible correlations between unobserved heterogeneity and other variables by including the vector of means of time-varying observed firm characteristics to the estimated equation and address the initial conditions using Wooldridge (2005) method. Using a large dataset with rich information on firm production and international trade activities of Chinese manufacturing firms for 2002 to 2006, I build a dynamic random-effects Probit model and examine the determinants of firms' export behaviour. We find that the most important determinant of firms' decision to export is the sunk entry costs: firms that have exported in the previous period are 13.4% more likely to export in the current period. Past experience in imports market is also found to influence the propensity to export. Consistent with findings of previous studies, I find that firm size, financial health and ownership are significant factors of firms' decision to export. Medium-sized firms, more liquid firms and foreign-owned (including HMT) and private firms are more likely to export. Comparison of different specification shows that the size of estimated sunk entry costs of exporting is reduced by over 50% when firms' previous import activity, the correlation between firm heterogeneity and initial conditions are properly controlled for.

3.1 Introduction

In an increasingly globalized world a large literature has emerged to explore firms' exporting behaviour. An explanation of why some firms choose to export while others supply only the domestic market can be found by examining the characteristics of exporters and non-exporters. The existing literature has shown that exporters outperform non-exporters in terms of productivity, capital-intensity, labour-force quality, size and financial health (Albarran et al., 2013; Bernard and Jensen, 2004; Greenaway et al., 2007; Lawless, 2009; Van Biesebroeck, 2005). Central to this literature is the identification of factors that determine firms' decision to export. A number of studies have identified sunk-entry costs and firm heterogeneity as important determinants of exporting behaviour (Das et al., 2007; Greenaway and Kneller, 2008; Manez et al., 2008; Sinani and Hobdari, 2010).

One feature of the existing studies is that the majority have ignored firms' import performance and focused solely on export activity (see, for example, Alvarez and Lopez, 2005; Bellone et al., 2010; Bernard and Wagner, 2001; Das et al., 2007). This is despite the fact that a large number of exporters also import a wide range of intermediate products. Vogel and Wagner (2010) find that 72% of German exporters imported in 2005; Wang and Yu (2012) find that in China 64% of exporters imported between 2002 and 2006; Aristei et al. (2013) study exporting and importing activities for firms from 27 eastern Europe and central Asia (ECA) countries using the Business Environment and Enterprise Performance Survey (BEEPS) data and find that in 2008 almost 85% of exporter were also importers. Finally, Bas and Strauss-Kahn (2014) find that 86% of French exporters were also importers between 1996 and 2005. The existing literature on firm-level imports shows that access to foreign intermediate inputs improves firm performance via variety, quality and learning effects. Specifically, the use of imported inputs raises firm productivity (Amiti and Konings, 2007; Augier et al., 2013; Halpern et al., 2005; Kasahara and Rodrigue, 2008), increases the introduction of new products (Colantonea and Crino, 2014; Goldberg et al., 2010) and increases the probability of exporting and the range of countries exported to (Aristei et al., 2013; Bas, 2012; Bas and Strauss-Kahn, 2014; Fugazza and McLaren, 2014). If importing improves firm performance, ignoring import activity will upwardly bias the estimated premia associated with exporting. Moreover, established importers that are familiar with a particular foreign market may face lower sunk entry costs and fixed costs when they decide to export to the market as they have already had access to the market information and may be able to use their existing distribution channels for exports. Therefore, estimating firms' export market participation will overstate the size of sunk costs of exporting ignoring their import activity.

In this paper I revisit the literature on exporting and firm heterogeneity and make the following contribution. Firstly, as our dataset not only contains detailed information on firms' production and exporting activity, but also information on their import activities, we are able to examine the relationship between firms' exports and characteristics taking into account their import activities. To the best of our knowledge, only three papers, namely Muuls and Pisu (2009), Aristei et al. (2013) and Kasahara and Lapham (2013), have addressed the effects of firms' import market participation on the examination of their export decision. Using Belgium firm-level balance sheet data and customs data, Muuls and Pisu (2009) add lagged import status to the Probit model of export participation and find a positive and significant coefficient of lagged import status, indicating previous experience in importing inputs raises the probability of exporting in the current period. Aristei et al. (2013) estimate a bivariate Probit model of exporting and importing using firm-level survey data from 27 eastern Europe and central Asia countries and find that firms' importing activity increases the probability of exporting. However, when controlling for current productivity and innovation, the effect of past imports on current exports disappears. Kasahara and Lapham (2013) develop a theoretical model with heterogenous firms simultaneously choosing whether to export and import and find that plants that both export and import pay considerably lower sunk costs than only-exporters or only-importers because of cost complementaries using Chilean manufacturing plant data. The drawback of Kasahara and Lapham (2013) is that they do not consider the correlation between unobserved firm heterogeneity and explanatory variables. Aristei et al. (2013) also do not control for the initial conditions as discussed below.

Secondly, I address the initial conditions problem that plagues the majority of existing studies that use the lagged dependent variable as one of the explanatory variables. It is widely recognized that exporting involves large sunk-entry costs for example searching information on target markets, meeting legal requirements and establishing distribution networks (Bernard and Jensen, 2004; Manova, 2013; Roberts and Tybout, 1997; Vogel and Wagner, 2010). In the previous literature lagged export status has been used to capture the sunk costs associated with exporting. However, the majority of these studies have not addressed the initial conditions problem (Arnold and Hussinger, 2005; Bellone et al., 2010; Bernard and Wagner, 2001; Greenaway et al., 2007; Yi and Wang, 2012). This problem arises when the start of the sample does not coincide with the start of the firms' exporting process and hence the initial export status is correlated with unobserved firm heterogeneity. Assuming that the initial observations are pre-determined or exogenous leads to over estimation of the size of the sunk entry costs of exporting. A small number of papers have addressed the problem (Albarran et al., 2013; Campa, 2004; Das et al., 2007; Lawless, 2009; Manez et al., 2008; Requena-Silvente, 2005; Roberts and Tybout, 1997). However, these studies have either ignored firms' import activity, or assumed zero correlation between unobserved firm heterogeneity and other observable characteristics. In this chapter I employ the Wooldridge (2005) technique to control for the influence of the initial conditions.

Finally, I provide a comprehensive analysis of the export behaviour of Chinese firms. China has experienced a remarkable growth in exports and imports over the last ten years.[1] Studies that examine the export behaviour of Chinese firms include Yang and Mallick (2010), Lu et al. (2010), Sun and Hong (2011), Yi and Wang (2012) and Dai and Yu (2013). Again, these studies focus on exporting activity only. Research into the impact of importing on Chinese firms is limited.[2] Yi and Wang (2012) use a random-effects Probit model to study the decision to export of industrial firms in Zhejiang Province for the period 2001-2003. This chapter differs from theirs in several ways. We construct a large dataset with rich information on firms' char-

[1]Between 2000 and 2010 China's exports rose from USD249.2 billion to USD1,577.75 billion, and imports increased from USD225 billion to USD1,396.24 billion during this period (China Statistical Yearbook 2013).

[2]Examples are Manova and Zhang (2009), Feng et al. (2012) and Wang and Yu (2012).

acteristic and trade for the whole of China for 2002-2006 and I am able to control for firms' importing activity. Furthermore, I allow for the possible correlation between unobserved firm heterogeneity and the observed firm characteristics in modelling the exporting behaviour and the initial conditions problem.

To briefly summarise our main results I find that not only whether a firm exports but also whether it imports is important for explaining differences in firm performance. In the examination of firms' decision to export, different specifications of a dynamic random-effects Probit model show that firms' import activity is an important factor of their exporting behaviour and firms pay lower sunk costs of exporting if they have been active in the imports market previously. Our results show that sunk entry costs of exporting are overestimated by over 50% if firms' imports and initial conditions are not properly controlled for.

The remainder of this chapter is organized as follows. Section 3.2 reviews the literature on the determination of firms' export behaviour. Section 3.3 presents our model and discusses the related econometric issues and solutions. Section 3.4 describes the data and the variables. Empirical results and interpretation are presented in Section 3.5 and finally Section 3.7 concludes.

3.2 Literature Review

Following the work by Bernard and Jensen (2004), who demonstrate differences between exporters and non-exporters using firm-level data from the US manufacturing industry, a number of theoretical models were developed that suggest hysteresis in exporting may be due to the costs of entering foreign markets at the firm level. Entry costs can be categorized into upfront sunk costs (e.g., costs of accumulating information about demand in foreign markets, product customization and establishing distribution channels) and fixed costs (e.g., costs of advertising, transportation and customer services). Sunk costs are incurred before a firm enters a foreign market and cannot be recovered whether the firm remains or exits the market.

Roberts and Tybout (1997) find that for Colombian manufacturing plants there are substantial sunk costs associated with entering or exiting the export market and sunk costs are a significant source of export persistence. Bugamelli and Infante (2003) find that an exporting firm has a probability of exporting in the next year that is 70 percentage points greater than that of an otherwise identical non-exporting firm for Italy. Bernard and Jensen (2004) and Greenaway and Kneller (2004) find export experience in the previous period raises the probability of exporting for the US and UK manufacturing firms. In an attempt to quantify sunk-entry costs of exporting, Das et al. (2007) develop a dynamic structural model that characterizes firms' decisions about whether to export or not. They estimate that the average entry costs potential exporters face range from 344,000 1986 US dollars for basic chemical plants to 430,000 US dollars for leather producers, using Colombian manufacturing plant data between 1981-1991. Kasahara and Lapham (2013) estimate sunk cost of exporting for Chilean plants ranges from 363,000 1990 US dollars for wood products and 998,000 1990 US dollars for food products.

Apart from the sunk-entry costs, most empirical studies document firm heterogeneity, including productivity, size, capital intensity and skilled labor force and ownership, as an important determinant of firms' entry to export markets. Alvarez and Lopez (2008) find that within-industry heterogeneity, measured by differences in productivity, size, capital per worker and skills, has a significant effect on export market entry and exit rates for Chilean manufacturing plants for the years 1990-99. In a recent survey, Wagner (2012) reviews microeconometric studies since 2006 on international trade and firm performance and concludes that most of these papers reveal that exporters are larger and more productive than non-exporters and the more productive firms self-select into export markets. One exception is Greenaway et al. (2007) who find that TFP has no significant effect on the UK manufacturing firms' decision to export.

Firm ownership is also found to be a significant factor of export participation. Greenaway et al. (2007) find for UK manufacturing firms that foreign-owned firms are more likely to export over the period 1993-2003. Cole et al. (2010) use the annual survey of Thai manufacturing firms from 2001 to 2004 and find that the ownership structure of the firm has an important effect on the export participation, with foreign-owned firms having a higher probability of exporting than domestically owned firms. Kim and Park (2011) examine the relationship between ownership concentration and export performance of Korean manufacturing firms and find that higher ownership concentration rate increases the probability of firms entering foreign markets. However, Buck et al. (2010) find outside ownership has no discernible impact on enterprise exporting across medium and large industrial firms in Russia, Ukraine and Belarus using survey data over 1995-1997.

The impact of firm age on export behaviour is mixed. On the one hand, long-established firms may be more likely to look for foreign markets for further growth and expansion. On the other hand, some younger firms may be established with better technology and it is easier for them to enter export markets (born global) and since much of the domestic market is occupied by experienced firms, these younger firms aim to sell their products to foreign markets for profit. Roberts and Tybout (1997) show that as firm age increases, the probability of exporting is higher for Columbian firms. Kim and Park (2011) find the impact of business experience on entering export market diminishes as firms grow older for Korean manufacturing firms.

Firm financial situation is also regarded an important determinant of exporting. In the presence of sunk costs associated with exporting, firms with liquidity/credit constraints are prevented from selling their products abroad because of their limited access to external finances to cover the entry cost into foreign markets. Chaney (2013) develops a simple model of trade with liquidity constraints and heterogeneous firms and predicts that financial underdevelopment hinders exports. Manova (2013) also develops a heterogeneous-firm model with cross-country differences in financial development and cross-industry variation in financial vulnerability and shows that financial frictions impede firm selection into exporting as well as firm-level exports using a large panel of bilateral exports for 27 industries in 1985-1995 for the US. Empirical studies evidencing that less financial constrained firms are more likely to export include Bellone et al. (2010) for French firms, Minetti and Zhu (2011) and Caggese and Cunatc (2013) for Italian manufacturing firms and Muuls (2015) for Belgian manufacturing firms. However, Greenaway et al. (2007) do not find that better financial health *ex ante* leads to a higher probability of exporting, but that participation in export markets improves firms' financial health among UK manufacturing firms.

Other factors that may influence the firms' decision to export include corporate governance (Dixon et al., 2015; Filatotchev et al., 2001, 2007; Lu et al., 2009), exchange rate shocks (Alessandria and Choi, 2007; Campa, 2004; Chaney, 2013; Das et al., 2007; Manez et al., 2008), government incentives such as export subsidy programmes (Bernard and Jensen, 2004; Buck et al., 2010; Gorg et al., 2008), transport infrastructure (Albarran et al., 2013) and spillover effects from multinational enterprises, regional and industrial agglomeration (Aitken et al., 1997; Bugamelli and Infante, 2003; Greenaway and Kneller, 2008; Mayneris and Poncet, 2013; Sinani and Hobdari, 2010; Yi and Wang, 2012).

3.3 Identification and Estimation Strategy

3.3.1 The Model

Following Das et al. (2007), I assume that firms are rational profit-maximizers and are able to produce the profit-maximizing level of exports once they enter an export market. It is assumed that in the first year of exporting, firms have to incur costs of acquiring information on potential export markets, the cost of negotiation with the bureaucracy associated with customs and arranging finance and establishing distribution channels. After entering an export market, a firm is assumed to incur some fixed cost for each period in order to maintain a presence in that specific export market which includes, for example, transportation costs and cost of maintaining product standards. We define π_{it} as the gross profit of exporting for firm i at time t and Y_{it} a dummy variable which equals 1 if firm i exports at time t and 0 otherwise. Start-up sunk costs of exporting are assumed to be S and fixed costs C_{it}.[1] Hence the net current export profit, p_{it},

[1] As discussed in the previous sections, firms that import may face different sunk costs of exporting from non-imports due to their existing foreign network. It is more accurate to incorporate in the model two levels of sunk costs of exporting for importers and non-importers. We do not differentiate the sunk costs for these two groups of firms here since I will account for the effect of firms' import activity on their exporting decision by including firms' import activity in the main estimation equation later.

for firm i at time t can be written as:

$$p_{it} = \begin{cases} 0, & ifY_{it} = 0 \\ \pi_{it} - C_{it} - S, & ifY_{it} = 1\&Y_{i(t-1)} = 0 \\ \pi_{it} - C_{it}, & ifY_{it} = 1\&Y_{i(t-1)} = 1 \end{cases} \tag{3.1}$$

Every period, before making the decision to export, a firm considers the current realization on its gross profits and chooses the infinite-horizon decision $\{Y_{it}\}_{t=j}^{\infty}$ that maximizes the expected present value of payoffs V_{it} given by:

$$V_{it} = \max_{\{Y_{it}\}_{t=j}^{\infty}} E_t(\sum_{j=t}^{\infty} \delta^{j-t} p_{it} | \Omega_{it}) \tag{3.2}$$

where E_t is the expectation operator conditioned on the information available at time t, $\delta \in [0, 1)$ is the one-period discount rate and Ω_{it} is the firm-specific information set. Following the Bellman equation, Equation (3.2) is equal to:

$$V_{it} = \max_{\{Y_{it}\}} [p_{it} + \delta E_t(V_{i(t+1)})] \tag{3.3}$$

The exporting choice maximizes the bracketed term in (3.3) and the firm will export if:

$$p_{it}^{*} = p_{it} + \delta \Delta E_t(V_{i(t+1)}) \geq 0 \tag{3.4}$$

where $\Delta E_t(V_{i(t+1)}) = E_t(V_{i(t+1)}|Y_{it} = 1) - E_t(V_{i(t+1)}|Y_{it} = 0)$. Thus the decision to export is based on both terms in (3.4), the current net profit from exporting and the expected value of being able to export in the following period without incurring the sunk cost. Let:

$$\pi^*_{it} = \pi_{it} - C_{it} + \delta \Delta E_t(V_{i(t+1)}) \tag{3.5}$$

and combining Equation (3.5) with Equation (3.1), the export market participation rule can be written as:

$$Y_{it} = \begin{cases} 1, if \pi^*_{it} - S(1 - Y_{i(t-1)}) \geq 0 \\ 0, otherwise \end{cases} \tag{3.6}$$

We follow Roberts and Tybout (1997) and employ a reduced-form equation to approximate π^* in Equation (3.6) as a function of a set of observable firm and market characteristics, X_{it}, and noise, υ_{it}. Equation (3.6) can be expressed as:

$$Y_{it} = 1[X'_{it}\beta + \gamma Y_{i(t-1)} + \upsilon_{it} \geq 0] \tag{3.7}$$

where $1[.]$ is the indicator function. The probability of exporting, Pr, is given as:

$$Pr(Y_{it}|X_{it}, Y_{i(t-1)}, \upsilon_{it}) = \Phi[(X'_{it}\beta + \gamma Y_{i(t-1)} + \upsilon_{it})(2Y_{it} - 1)] \tag{3.8}$$

where Φ is the unit normal cumulative distribution function.

Given that changes in firm characteristics, such as size and productivity, may induce firms to switch their export status and exporting may in turn affect firm's characteristics, I follow the previous literature (Bernard and Jensen, 2004; Greenaway et al., 2007; Roberts and Tybout, 1997; Yi and Wang, 2012) and lag all independent variables by one year, making such factors pre-determined which helps to mitigate possible simultaneity concerns. Hence, I estimate the firm's decision to export by estimating the following:

$$Y_{it} = 1[X'_{i(t-1)}\beta + \gamma Y_{i(t-1)} + \upsilon_{it} \geq 0] \tag{3.9}$$

where X is a vector of observable firm and market characteristics, including firm age, productivity, labour quality (proxied by average wages of employees), size, financial status (liquidity and leverage ratios), ownership structure, region, industry and year dummies. Region (industry) dummies are included to control for time-invariant factors common to firms across regions (industries) and year dummies to account for business cycle effects and firm-invariant market factors such as changes in exchange rates or government policy. Lagged export status Y_{t-1} is included to account for sunk entry costs. If there are no sunk costs, then export participation does not depend on past exporting experience. If sunk entry costs are significant, the coefficient on lagged export status can be considered as a measure of the magnitude of these costs.

We expect firms' past export experience, productivity, labour quality, size and financial health have positive impact on the propensity of exporting of Chinese manufacturing firms. The reason is due to the sunk costs involved in exporting. Firms that have exported previously will be more likely to export in the current period because they do not need to incur the sunk entry costs again for market research, foreign regulation, etc. Larger sized and more productive firms are more able to afford such large sunk cost, and thus more likely to enter the export market. Firms with skilled labour force can produce better quality products and firms with better financial status can invest more in new products and technology. Foreign ownership including that from HMT is expected to positively influence the firms' decision to export due to the access to foreign market information and established distribution network by their parent companies. Firm age can be either positively or negatively correlated with their export participation depending on the importance of firm experience and the adventurism of the young firms.

The estimation of Equation (3.9) raises two main econometric issues: unobserved firm heterogeneity and the initial conditions problem. We discuss both issues in the following subsections respectively.

3.3.2 Unobserved firm heterogeneity

Although Equation (3.9) controls for a set of observed firm characteristics that are assumed to influence firms' decision to export, firms may still differ due to certain unobserved heterogeneity that affects the decision but is not influenced by the previous export experience. Examples might include technology, managerial ability or the existence of extensive foreign networks. If the unobserved firm heterogeneity exhibits persistence over time, ignoring it will lead to an overestimation of previous export experience as the lagged export status may appear to be the sole determinant of the future decision to export. For example, a firm that exported in an initial period due to the influence of a certain unobserved heterogeneity, such as management strategy focussed on internationalization, and continued to export for the same reason. If we employ standard econometric techniques, it is incorrect to assume that past export experience influenced the current decision when the current decision was caused by unobserved firm heterogeneity. The problem of such a conditional relationship between past and future export participation caused by the improper control of the unobserved firm heterogeneity is the spurious state dependence problem as discussed by Heckman (1981a,b).

Hence I assume that the error term υ_{it} in Equation (3.9) consists of two components, the unobserved firm-specific time-invariant effects, k_i, and an exogenous random disturbance, μ_{it}, which gives:

$$\upsilon_{it} = k_i + \mu_{it} \tag{3.10}$$

where k_i is normally distributed and independent on μ_{it} and μ_{it} is normally distributed and independent of the explanatory variables for each i and t. To control for unobserved firm heterogeneity, fixed-effects (FE) and random-effects (RE) approaches can be used. The unobserved effect k_i is treated as a parameter to be estimated in a fixed-effects framework while the random-effects framework treats it as a random variable. In a linear setting, a fixed-effects estimator using a transformation such as differencing or demeaning is sufficient to control for such unobserved effects. However, in the case of dynamic binary choice model with T fixed and $N \rightarrow \infty$, treating k_i as a parameter can potentially lead to seriously biased estimations of β and γ because of the incidental parameters problem (Heckman, 1981b; Wooldridge, 2010).

As Wooldridge (2010, p.286) points out, with a large number of random draws, to treat the unobserved effects as random draws always makes sense and is certainly appropriate from the neglected heterogeneity perspective. Hence, for our estimation I use a RE Probit model which is a popular model for binary outcomes with panel data. The standard RE Probit assumes that the unobserved heterogeneity is uncorrelated with all the other explanatory variables. This is a strong assumption and does not usually hold in reality. In our example, managerial expertise, part of the unobserved firm heterogeneity k_i that may affect firm's decision to export, is likely to be correlated with the quality of labour employed in a firm (proxied by employees' wages), which is an element of $X_{i(t-1)}$. Following Mundlak (1978) and Chamberlain (1984) I allow for possible correlation between k_i and $X_{i(t-1)}$ by:

$$k_i = \bar{X}_i'\delta + \epsilon_i \tag{3.11}$$

where $\bar{X}$ is a vector of means of all time-varying covariates in X and $\epsilon_i \sim$iid N(0, σ_ϵ^2), independent of X_{it} and μ_{it} for all i and t. Substituting Equation (3.11) into Equation (3.9), the estimating equation becomes:

$$Y_{it} = 1[\gamma Y_{i(t-1)} + X'_{i(t-1)}\beta + \bar{X}_i'\delta + \epsilon_i + \mu_{it} \geq 0] \tag{3.12}$$

3.3.3 The initial conditions problem

The estimation of Equation (3.12) requires an assumption on the relationship between the initial export participation, Y_{i1}, and the error ϵ_i. Studies of a dynamic binary choice model using a standard random-effects Probit assume that the initial conditions are pre-determined or exogenous, i.e., the initial observation values are independent of unobserved firm-specific effects. If the observation period for each firm starts from the very beginning of the process generating Y_{it}, the initial conditions can be assumed to be exogenous. However, as in most microeconometric studies with panel datasets, the start of our sample does not coincide with the start of the exporting process and the assumption of exogenous initial conditions is inappropriate since the unobserved time-invariant firm heterogeneity, such as foreign networks or managerial ability that affects current export participation, can influence the export behaviour in the initial period. Ignoring the endogeneity of the initial conditions will thus lead to a biased estimation, leading to an overestimation of the state dependence, i.e., the effect of past export experience on the current decision to export which would suggest that sunk costs are overstated.

Heckman (1981b) proposes a solution to the initial conditions problem. Specifically, he specifies a reduced-form equation for the initial period:

$$Y_{i1} = 1[X'_{i0}\beta_1 + \zeta_i \geq 0] \tag{3.13}$$

where ζ_i is correlated with ϵ_i, but uncorrelated with μ_{i1} for $t \geq 2$. ζ_i can be written as $\zeta_i = \theta\epsilon_i + \eta_{i1}$, with ϵ_i and η_{i1} being independent. Equation (3.13) is then specified as:

$$Y_{i1} = 1[X'_{i0}\beta_1 + \theta\epsilon_i + \eta_{i1} \geq 0] \tag{3.14}$$

Together with Equation (3.12), a complete model for firms' exporting behaviour can be estimated using the technique of maximum likelihood as in Roberts and Tybout (1997). However, the implementation of the Heckman estimator requires separate programming and is computationally cumbersome (Arulampalam and Stewart, 2009; Stewart, 2007).[1] A simpler, and more practical, method is proposed by Wooldridge (2005).

Wooldridge (2005) suggests a conditional maximum likelihood (CML) estimator that con-

[1] Mark Stewart has written a STATA command -redprob- for the Heckman estimator and I used a high performance computing cluster for estimations using this method although the regressions still failed to converge even after a very lengthy time period.

siders the distribution of the unobserved firm heterogeneity which is conditioned on the initial values and exogenous variables. The model for ϵ_i given Y_{i1} and X_i is expressed as:

$$\epsilon_i = a_0 + a_1 Y_{i1} + \bar{X}_i' \delta + \xi_i \tag{3.15}$$

where ξ is uncorrelated with the initial observation Y_1. Substituting Equation (3.15) into Equation (3.12) gives:

$$Y_{it} = 1[\gamma Y_{i(t-1)} + X'_{i(t-1)}\beta + a_0 + a_1 Y_{i1} + \bar{X}_i'\delta + \xi_i + \mu_{it} \geq 0] \tag{3.16}$$

where $\bar{X}_i$ has been incorporated previously by the Mundlak (1978) specification. The Wooldridge (2005) method leads to a simple and tractable likelihood function that has the same structure as a standard random-effects Probit model. We can estimate the parameters by expanding the list of explanatory variables to include Y_{i1} and $\bar{X}_i$. Since the Wooldridge estimator has the advantage of employing the standard RE Probit software, I use this method for our main estimation.[1]

As will be discussed shortly in the next section, the control variables in X include firm characteristics (such as age, wage, TFP, size, import activities, financial status and ownership type), regional, industry and year dummies. We expect younger, larger, more productive, more skill-intensive and financial healthier firms and firms with foreign ownership have a larger propensity to participate in the exports market.

[1] Arulampalam and Stewart (2009) propose a shortcut method for implementing the Heckman estimator using standard software and compare it with two other estimators suggested by Orme (1997) and Wooldridge (2005) and find that three estimators provide fairly similar results. However, they still find that the shortcut Heckman estimator is considerably more expensive in terms of computer time than the other two estimators.

3.4 Data and Descriptives

3.4.1 The Data

The data used in this chapter are drawn from two sources. The firm-level production data are from the Annual Survey of Industrial Enterprises provided by the National Bureau of Statistics of China (NBS) and the transaction-level trade data are provided by the Department of Customs Trade Statistics, the General Administration of Customs of China. We use data from both sources over the sample from 2002 to 2006.

The NBS data cover all state-owned industrial enterprises and non-state-owned industrial enterprises with annual sales of greater than 5 million Chinese Yuan (RMB).[1] According to the NBS industry classification, industrial enterprises refer to enterprises that operate in the mining and quarrying sector, manufacturing sector or in the production and supply of power, gas and water. The NBS survey is the primary source for the construction of numerous aggregate statistics used in the China Statistical Yearbooks. The NBS data include the firm's identification (tax code) and basic information such as year founded, location, ownership type, employment, China industrial classification (CIC) code as well as over 50 financial variables from the accounting statements, including capital, assets, liabilities, industrial value-added, sales, income, profits, investment, depreciation, wage bill and R&D expenses.

Trade data record all import and export transactions with non-zero values that enter or exit through Chinese customs. Each observation represents a shipment and contains detailed in-

[1]The official RMB exchange rate per US dollar between 2002 and 2004 was 8.277, 8.194 in 2005 and 7.973 in 2006 (World Development Indicators, World Bank). Hence, the threshold for inclusion in the dataset is equivalent to between USD600,000 and USD627,000 over the sample period.

formation on the time of the transaction (month and year), type of trade (import/ export), exporting/ importing firm identifier, ownership type, product traded (8-digit HS code and name), value, quantity, unit, destination country (of the exported commodities) / country of origin (of the imported products), type of trade (ordinary trade, processing trade, compensation trade, consignment, etc.) and finally mode of transport.

Please see Chapter 2 for detailed introduction to both datasets and construction of key variables such as TFP, firm size and ownership. Appendix 2A provides the method of the TFP estimation following De Loecker (2007).

In order to explore the relationships between firm characteristics and exporting activity, I match the NBS data and the trade data following Brandt et al. (2012). Detailed procedure on data cleaning is described in Chapter 2. I drop the intermediary firms (pure trading firms) as they do not have their own production but simply export or import for other domestic firms.[1] We also exclude processing trade firms as they are by definition importers and exporters.[2] We further drop all firms with less than 8 employees following Brandt et al. (2012) and Upward et al. (2010) as they fall under a different legal regime. Our final sample consists of 42,666 firms and corresponds to 213,330 firm-year observations for the period 2002 to 2006.

[1] Ahn et al. (2011) discuss the role of the Chinese intermediaries in facilitating international trade. We identify the intermediary firms based on Chinese characters that have the English-equivalent meaning of "importing", "exporting", and/or "trading company" in the firm name.

[2] See Manova and Yu (2012), Wang and Yu (2012) and Yu (2014) for discussions on Chinese processing trade firms.

3.4.2 Descriptives

We begin by looking at some simple summary statistics. See Table 3A.1 in Appendix 3A for detailed definitions of the variables used in the chapter. We first group firms into non-exporters and exporters and I further distinguish exporters by their participation in the imports market into only-exporters and importer-exporters. Table 3.1 provides a comparison of characteristics of firms from different groups. The majority of the Chinese manufacturing firms (73.68%) do not export. Of the remaining 26.32% that export, just over half import as well as export. In line with the stylized facts documented in the existing literature, Chinese exporters appear to be younger, larger, more productive, more capital-intensive, pay higher wages, more liquid and less leveraged than non-exporters. The importer-exporters firms appear to outperform export-only firms in terms of size, productivity, capital and skill intensity and financial health. Such premia show that apart from exporting, importing behaviour has a potentially important impact on firm performance.

To examine in more detail the differences among firms with different trade orientation, I test the premia for several performance measures by estimating the following:

$$P_{it} = \alpha_o + \alpha_1 EXPonly_{it} + \alpha_2 IMPonly_{it} + \alpha_3 EXPIMP_{it} + \lambda X_{it} + \varepsilon_{it} \tag{3.17}$$

where P_{it} represents firm performance, i.e., employment, wage, capital intensity, TFP and labour productivity respectively. The variables *EXPonly*, *IMPonly* and *EXPIMP* are dummy variables representing firms that export only, import only and both export and import, and firms that neither export nor import are omitted as the reference group; *X* is a vector of control vari-

Table 3.1: Summary statistics of key variables

	All firms	Non-exporters	Exporters	Only-exporters	Importer-exporters
	(1)	(2)	(3)	(4)	(5)
age	2.261	2.265	2.249	2.204	2.292
	(0.701)	(0.735)	(0.595)	(0.624)	(0.563)
employment	5.100	4.939	5.551	5.362	5.731
	(1.107)	(1.062)	(1.106)	(0.991)	(1.178)
capital intensity	5.005	4.957	5.138	4.792	5.470
	(1.034)	(1.015)	(1.073)	(0.954)	(1.077)
wage	2.422	2.356	2.608	2.417	2.791
	(0.588)	(0.573)	(0.590)	(0.497)	(0.613)
TFP_DL	6.538	6.427	6.850	6.540	7.146
	(1.080)	(1.055)	(1.085)	(0.938)	(1.133)
labor productivity	3.964	3.945	4.018	3.756	4.268
	(1.014)	(1.010)	(1.026)	(0.908)	(1.070)
liquidity	0.103	0.097	0.120	0.097	0.142
	(0.270)	(0.274)	(0.120)	(0.262)	(0.252)
leverage	0.966	0.987	0.909	0.972	0.848
	(2.652)	(2.350)	(3.354)	(4.751)	(0.634)
Observations	213,330	157,172	56,158	27,468	28,690
(% of total)	(100.00%)	(73.68%)	(26.32%)	(48.91%)	(51.09%)

Notes: Reported are the means of the variables with standard deviations in parentheses. All variables are expressed in natural logarithms except liquidity ratio. See Table 3A.1 for detailed definition of the variables.

ables, namely firm age, employment, wages, and ownership, region, industry and year dummies.[1] We estimate Equation (3.17) first using pooled OLS, and to control for unobserved firm heterogeneity which may lead to the performance differences, I also run fixed effects regressions.[2]

Table 3.2 presents the results with Panel A from pooled estimation and Panel B fixed effects. The coefficients of all trading status are significant at 0.01 level. The coefficients from fixed effects regressions are considerably smaller that those from the pooled ordinary least square (OLS) regressions, showing that unobserved firm heterogeneity is important for firm performance and is correlated with firms' international activities. Results show that firms that participate in international markets are larger, more productive, more capital- and skill-intensive than firms that serve domestic markets only. Firms that both import and export have highest premia, followed by only-importers and only-exporter have lowest premia. The results are consistent with Muuls and Pisu (2009) and Silva et al. (2013) who also find better performance of only-importers than only-exporters in terms of productivity, sales, wages and capital intensity.[3] We also test whether firms in each trade group differ from other groups and results of Wald tests for all performance measures. These results are reported in the lower half of both panels. We are interested in whether firms that both export and import perform similarly to those that export only. The hypotheses of $\alpha_{EXPonly} = \alpha_{EXPIMP}$ are rejected at 0.01 significance level for all performance indicators for both OLS and fixed-effects regressions. The results show that exporters that also import are significantly different from firms that export only. Hence, when

[1]When the dependent variable is employment (wage), employment (wage) is omitted from the control variables.

[2]We did Hausman tests on the choice between fixed effects and random effects and fixed effects model is favoured.

[3]Castellani et al. (2010), Vogel and Wagner (2010) and Haller (2012) also divide firms into four groups according to their trade orientation and find highest premia for firm that both import and export, but they find only-exporters outperform only-importers.

estimating export premia, neglecting firms' importing activity could overestimate the export premia.

3.5 Empirical Results

We present our empirical results in this section. Due to the non-linearity of the model, the coefficient estimates are not informative about the magnitude of the effects of the outcome variable. To facilitate interpretation, I report the average marginal effects (AME) in Table 3.3.[1] The coefficient estimates are presented in Table 3A.2 in the Appendix.

As a benchmark and following the majority of the existing literature, Column (1) provides a simple RE Probit estimation in which firm import activity is not considered and the unobserved heterogeneity is assumed to be independent of explanatory variables. Lagged export status is positive and significant at 0.01 level and firms that exported in the previous year are 30% more likely to export in the current period than those that did not. In line with the existing literature, sunk entry costs of exporting are found for Chinese manufacturing firms.

Other firm characteristics including age, labour force quality, productivity, liquidity, size and ownership are found to be significant determinants of their export participation decision. Younger, larger, more productive firms and firms with a skilled labour force are more likely to export. Firm age has a negative effect on the probability of exporting such that a one unit de-

[1] The average marginal effects of continuous variables are computed by $AME = \frac{1}{n}\sum_{i=1}^{n}\phi(X_i'\beta)\beta$ and dummy variables by $AME = \frac{1}{n}\sum_{i=1}^{n}[\Phi(X_i'\beta|X_j = 1) - \Phi(X_i'\beta|X_j = 0)]$, where $\Phi(.)$ is the standard normal cumulative distribution function and $\phi(.)$ is the standard normal density function.

Table 3.2: Descriptive regressions: trade premia

	employment	wage	capital intensity	TFP	labour productivity
	(1)	(2)	(3)	(4)	(5)
Panel A: Pooled OLS					
EXPonly	0.418***	0.018***	0.045***	-0.045***	-0.042***
	(0.007)	(0.004)	(0.006)	(0.006)	(0.006)
IMPonly	0.653***	0.326***	0.598***	0.280***	0.393***
	(0.012)	(0.006)	(0.010)	(0.010)	(0.010)
EXPIMP	0.816***	0.272***	0.429***	0.186***	0.269***
	(0.008)	(0.004)	(0.007)	(0.006)	(0.007)
Observations	213,330	213,330	213,330	213,330	213,330
R-squared	0.188	0.216	0.340	0.406	0.260
F-tests					
$\alpha_{EXPonly} = \alpha_{IMPonly}$	355.24	2255.91	2759.64	974.49	1584.44
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\alpha_{EXPonly} = \alpha_{EXPIMP}$	2047.62	3079.36	2651.41	986	1622.38
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\alpha_{IMPonly} = \alpha_{EXPIMP}$	179.34	73.03	275.8	86.21	136.08
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Panel B: Fixed effects					
EXPonly	0.092***	0.018***	0.064***	0.015**	0.026***
	(0.004)	(0.005)	(0.004)	(0.008)	(0.008)
IMPonly	0.057***	0.021***	0.071***	0.034***	0.047***
	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)
EXPIMP	0.150***	0.036***	0.099***	0.050***	0.066***
	(0.005)	(0.006)	(0.005)	(0.009)	(0.009)
Observations	213,330	213,330	213,330	213,330	213,330
R-squared	0.070	0.097	0.371	0.147	0.160
F-tests					
$\alpha_{EXPonly} = \alpha_{IMPonly}$	28.74	0.16	1.32	2.78	3.65
	(0.000)	(0.694)	(0.251)	(0.095)	(0.056)
$\alpha_{EXPonly} = \alpha_{EXPIMP}$	169.81	13.1	66.92	21.8	28.71
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\alpha_{IMPonly} = \alpha_{EXPIMP}$	220.85	4.69	21.39	2.51	3.32
	(0.000)	(0.030)	(0.000)	(0.113)	(0.068)

Notes: All regressions include firm age, employment (except Column (1)), wage (except Column (2)), ownership, region, industry and year dummies. For both panels, the upper part reports the coefficients on firms' trade status with standard errors in parentheses and the lower part reports the F-statistics with p-values in parentheses. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

crease in the log of firm age leads to a 0.5 percentage points increase of probability of exporting. A one unit increase in the log of average wages of a firm increases its probability of exporting by 0.8 percentage points. Similarly, a one unit increase in TFP leas to a 0.5 percentage point increase in the probability of exporting. Firm liquidity has a negative and highly significant sign, indicating that more liquid firms are less likely to export. This result contradicts some previous studies (Bellone et al., 2010; Chaney, 2013). Such odd relationship is found due to the fact that firms' import activities are not taken into account in the estimation. We will address this shortly. Leverage ratio is not significant at usual level. Compared to small firms, medium- and large-sized firms have a higher propensity to export. State ownership has a negative impact on the firms' decision to export while foreign and HMT ownership increases the probability of exporting by 6% and 4.3% than SOEs respectively. This finding is not surprising as foreign-owned and HMT firms usually have more advanced technologies and established foreign networks. Privately owned firms also have a 2.6% higher probability of exporting than SOEs since they maybe more willing to experiment in foreign markets searching for profitable opportunities. For many SOEs, the decision is controlled by the government often for non-profit-maximizing reasons (e.g., the maintenance of a stable level of employment or implementation of specific economic policies in strategic sectors). Collective ownership is also found to be a negative determinant of exporting.

As found in the previous section firms that both import and export perform differently from firms that export only, I then include firms' import status as an additional explanatory variable in Column (2). Lagged export status remains positive and highly significant although the AME drops slightly to 0.290. The lagged import status is also positive and significant at 0.01 level such that the propensity of exporting in the current period increases by 3.4% if the firm was importing the previous year. Having experience in imports market means firms have better access to information in certain targeted markets that can be used to establish new distribution networks. Importing firms may also improve their productivity via learning from importing. Higher productivity in turn helps firms overcome the large sunk costs of exporting. Results of other variables are similar to those in Column (1), with slightly smaller coefficients of lagged variables *wage, TFP_DL, MEDIUM, LARGE, FOREIGN, HMT* and *PRIVATE*, since import activity, another important determinant of export decision, is controlled for in the estimation.

Applying standard RE Probit in specifications (1) and (2), firm unobserved heterogeneity is assumed to be uncorrelated with other explanatory variables. As discussed in Section 3.3, I then include the vector of means of time varying observable factors to control for possible correlation in Column (3). The AME of lagged export status drops slightly further to 0.278. Lagged import status is still significant at 0.01 level, but of negative sign, opposite to that found in Columns (1) and (2). Firm age is still a negative factor of export participation. Also in contrast to previous specifications, the coefficient of liquidity is now positive and significant at 5% level when import status is included in the estimation. However, wage and TFP are no long significant determinants. Similar results are found for firm size and ownership, except that large firms are no longer more likely to export than the small ones. Leverage is still found to have no significant impact on firms' decision to export. We now look at the results on the means of time-varying observable factors. Except for firm age and wage, all means of other variables have significant signs, indicating the correlation between firm unobserved heterogeneity and the observed characteristics.

Same as import status, I find opposite signs of lagged liquidity and its mean value, i.e., positive on lagged value and negative on mean value. Such finding is observed in a few studies on trade and other topics, such as the dynamic of unemployment and firms' investment decision. Egger and Url (2006) finds opposite signs for the coefficients on the current value and the average value of partner country's log population on log exports for Austria in the study of the impact of public export guarantees on trade using a Mundlak-type random effects GLS. Fugazza and McLaren (2014) also include means of time-varying dependent variables in a RE Probit model in the study of market access and survival of exports of Peruvian firms and find the coefficient on the current log value of sectoral imports has a positive effect on the probability of export survival while the average values have a negative effect. They point out that the coefficients on the current values reflect the immediate effect of the variable and those on the average value, the gradual effect. Studying the dynamics of unemployment using Wooldridge (2005) RE probit, Stewart (2007) finds opposite signs of education, marriage status with their lags and means. Drakos and Konstantinou (2013) also find opposite signs of coefficients on lags of employment and load to value-added and their means in the study on manufacturing firms' investment decision.

We interpret the results as that importing in the previous period decreases the probability of exporting in the current period by 5.6%, but being an importer on average for the sample period increases the probability of exporting by 14.1%. The reasons that firms that import in previous year are less likely to export in current period may be due the costs involved in importing as they may lag sufficient funds to invest on export markets when they shifted their investment to the import markets or the risks in international trade so they want to delay their decision to export immediately after their participation in importing. However, the over-time import activity plays a positive and more important role in the decision to export as being active in import markets firms have access to more varieties and better quality inputs and improve their performance gradually which in turn helps them break into the export market.

As for liquidity, the lagged has a positive impact on firms' export decision while its mean value has a negative impact. Firms with higher liquidity previous year are more likely to export in current year and the reason may be that entry to export markets incurs sunk costs and the liquid firms previous year will find it easier to pay such costs. But liquidity over the period has a negative impact on firms' export behaviour. The mean value of TFP (leverage) has a positive (negative) effect on firms' propensity of exporting at 5% level, meaning that over time more productive and less leveraged firms are more likely to participate in exports market. However the magnitudes of mean values of TFP, liquidity and leverage are far smaller than that of import status. So the most important determinants of Chinese manufacturing firms' export behaviour are the sunk costs and over time import experience.

In Column (4) I address the initial conditions problem by Wooldridge (2005) approach where the initial values of export status is included in the regression. Similar results are observed to Column (3) except the magnitudes of the AME are smaller for most variables. The largest change is the lagged export status from 0.278 in Column (3) to 0.134, a 50% drop. So the likelihood for firms that have export in the previous year to export in current year is 13.3%, rather than around 30% as found in previous specification. Slight changes are found for other factors. EXP_1, the exporting status in the initial period, is positive and significant at 0.01 level and firms that exported in the first period of the sample are 10.6% more likely to export in the current period. Our results show that without proper control for the initial condition overestimates the size of sunk entry costs of exporting by over 50%.

Table 3.3: Decision to export of Chinese manufacturing firms (AME)

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	0.300***	0.290***	0.278***	0.134***
	(0.002)	(0.002)	(0.001)	(0.005)
$IMP_{i(t-1)}$		0.034***	-0.056***	-0.028***
		(0.002)	(0.002)	(0.002)
$age_{i(t-1)}$	-0.005***	-0.006***	-0.010**	0.001
	(0.001)	(0.001)	(0.005)	(0.004)
$wage_{i(t-1)}$	0.008***	0.005***	-0.000	-0.001
	(0.001)	(0.001)	(0.002)	(0.001)
$TFP\_DL_{i(t-1)}$	0.005***	0.003***	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
$liquidity_{i(t-1)}$	-0.014***	-0.013***	0.009**	0.006**
	(0.002)	(0.002)	(0.004)	(0.003)
$leverage_{i(t-1)}$	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$MEDIUM_{i(t-1)}$	0.020***	0.017***	0.009***	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)
$LARGE_{i(t-1)}$	0.017***	0.011**	-0.005	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)
$FOREIGN_{i(t-1)}$	0.060***	0.049***	0.029***	0.025***
	(0.003)	(0.003)	(0.003)	(0.003)
$HMT_{i(t-1)}$	0.043***	0.037***	0.026***	0.023***
	(0.003)	(0.003)	(0.003)	(0.003)
$COLLECTIVE_{i(t-1)}$	-0.008**	-0.008**	-0.007**	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)
$PRIVATE_{i(t-1)}$	0.026***	0.025***	0.022***	0.023***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP_{i1}				0.106***
				(0.002)
$\overline{IMP_i}$			0.141***	0.110***
			(0.003)	(0.003)
$\overline{age_i}$			0.005	-0.008*
			(0.005)	(0.004)
$\overline{wage_i}$			-0.000	0.001
			(0.002)	(0.002)
$\overline{TFP\_DL_i}$			0.003**	0.002**
			(0.001)	(0.001)
$\overline{liquidity_i}$			-0.030***	-0.028***
			(0.005)	(0.004)
$\overline{liquidity_i}$			-0.001**	-0.001**
			(0.000)	(0.000)
Observations	170,664	170,664	170,664	170,664

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

3.6 Sensitivity Checks

We perform several sensitivity checks to verify our findings. Firstly, I use TFP calculated by the traditional method of Levinsohn and Petrin (2003) in the estimation of the probability of exporting and run the RE Probit regressions. The results of AME are presented in Table 3.4 with coefficients in Table 3A.3. Since TFP obtained by Levinsohn and Petrin (2003) and that by De Loecker (2007) are highly correlated (with a correlation of 0.994), the estimation results are almost the same as found in the previous section.

Secondly, I use the log value of imports of firms as an alternative measure of firms' import activity in our estimation and the results of AME are reported in Table 3.5 with coefficients in Table 3A.4. *impvalue* has the same sign but with slightly smaller magintude for each specification as that in Table 3.3.

Furthermore, I use log of employment as an alternative firm size in the estimation. Results of AME are reported in Table 3.6 with coefficients in Table 3A.5. The past export status is found to have a smaller impact on the propensity of current period exporting. Similar results are found as for the other dependent variabls as those presented above using alternative measurements of TFP, import activity and firm size.

Finally, I verify the results by removing observations with extreme values, i.e., the outliers. Variables in the regression are carefully examined to detect the extreme values and outliers are removed by dropping the 1st or (and) 99th percentile of the sample depending on the distribution of the variable. Different specifications of RE Probit are performed using the sample excluding outliers. The AME results are presented in Table 3.7 with coefficients in Table 3A.6. Results are broadly consistent with the previous ones with a slightly larger size of sunk entry costs and smaller magnitude of the impact of export status in the initial period.

3.7 Conclusions

In this chapter I investigate firm exporting behaviour by examining how sunk costs, firm heterogeneity and the initial conditions influence the decision to export. An important feature of this chapter is that I take into account firms' import activities. Using a large dataset of Chinese manufacturing firms with rich information on firm production and international trade over the period 2002 to 2006, I find that exporting incurs large sunk costs and that past export experience plays an important role in the current exporting participation decision. We also find that previous year import status decrease the propensity of exporting, but over period import activity increases the propensity. Size and ownership are also found to be significant factor affecting firms' decision to export. Firms of medium size, higher liquidity and foreign (including HMT) and private ownership are more likely to participate in exports market. Comparison across different specifications shows that the sunk costs of exporting are overstated when using simple RE Probit model without controlling for firms' import behavior, correlation between unobserved heterogeneity and explanatory variables and endogenous initial conditions. We correct for these concerns by applying a Mundlak-Wooldridge approach, i.e., introduction of a vector of means of observed firm heterogeneity and taking account of initial conditions, and I find that the size of sunk costs of exporting is reduced by over 50%.

Our results suggest a number of implications for the development of Chinese manufactur-

Table 3.4: Decision to export of Chinese manufacturing firms (AME): alternative TFP estimation

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	0.300***	0.290***	0.278***	0.134***
	(0.002)	(0.002)	(0.001)	(0.005)
$IMP_{i(t-1)}$		0.034***	-0.056***	-0.028***
		(0.002)	(0.002)	(0.002)
$age_{i(t-1)}$	-0.005***	-0.006***	-0.010**	0.001
	(0.001)	(0.001)	(0.005)	(0.004)
$wage_{i(t-1)}$	0.008***	0.005***	-0.000	-0.001
	(0.001)	(0.001)	(0.002)	(0.001)
$TFP\_LP_{i(t-1)}$	0.005***	0.004***	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
$liquidity_{i(t-1)}$	-0.014***	-0.013***	0.009**	0.006**
	(0.002)	(0.002)	(0.004)	(0.003)
$leverage_{i(t-1)}$	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$MEDIUM_{i(t-1)}$	0.019***	0.016***	0.009***	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)
$LARGE_{i(t-1)}$	0.016***	0.010**	-0.006	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)
$FOREIGN_{i(t-1)}$	0.060***	0.049***	0.029***	0.025***
	(0.003)	(0.003)	(0.003)	(0.003)
$HMT_{i(t-1)}$	0.043***	0.037***	0.026***	0.023***
	(0.003)	(0.003)	(0.003)	(0.003)
$COLLECTIVE_{i(t-1)}$	-0.008**	-0.008**	-0.007**	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)
$PRIVATE_{i(t-1)}$	0.025***	0.025***	0.022***	0.023***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP_{i1}				0.106***
				(0.002)
$\overline{IMP_i}$			0.141***	0.109***
			(0.003)	(0.003)
$\overline{age_i}$			0.005	-0.008*
			(0.005)	(0.004)
$\overline{wage_i}$			-0.001	0.001
			(0.002)	(0.002)
$\overline{TFP\_LP_i}$			0.003***	0.003**
			(0.001)	(0.001)
$\overline{liquidity_i}$			-0.030***	-0.029***
			(0.005)	(0.004)
$\overline{leverage_i}$			-0.001**	-0.001**
			(0.000)	(0.000)
Observations	170,664	170,664	170,664	170,664

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

Table 3.5: Decision to export of Chinese manufacturing firms (AME): alternative measure of import activity

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	0.300***	0.292***	0.284***	0.136***
	(0.002)	(0.002)	(0.002)	(0.006)
$impvalue_{i(t-1)}$		0.002***	-0.005***	-0.002***
		(0.000)	(0.000)	(0.000)
$age_{i(t-1)}$	-0.005***	-0.006***	-0.011**	-0.000
	(0.001)	(0.001)	(0.005)	(0.004)
$wage_{i(t-1)}$	0.008***	0.004***	0.000	-0.001
	(0.001)	(0.001)	(0.002)	(0.001)
$TFP\_DL_{i(t-1)}$	0.005***	0.003***	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
$liquidity_{i(t-1)}$	-0.014***	-0.013***	0.009**	0.006*
	(0.002)	(0.002)	(0.004)	(0.003)
$leverage_{i(t-1)}$	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$MEDIUM_{i(t-1)}$	0.020***	0.017***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)
$LARGE_{i(t-1)}$	0.017***	0.010**	-0.005	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)
$FOREIGN_{i(t-1)}$	0.060***	0.050***	0.034***	0.030***
	(0.003)	(0.003)	(0.003)	(0.003)
$HMT_{i(t-1)}$	0.043***	0.038***	0.030***	0.026***
	(0.003)	(0.003)	(0.003)	(0.003)
$COLLECTIVE_{i(t-1)}$	-0.008**	-0.008**	-0.008**	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)
$PRIVATE_{i(t-1)}$	0.026***	0.025***	0.022***	0.023***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP_{i1}				0.109***
				(0.002)
$\overline{impvalue_i}$			0.010***	0.007***
			(0.000)	(0.000)
$\overline{age_i}$			0.007	-0.007
			(0.005)	(0.004)
$\overline{wage_i}$			-0.000	0.001
			(0.002)	(0.002)
$\overline{TFP_DL_i}$_dl			0.003**	0.002**
			(0.001)	(0.001)
$\overline{liquidity_i}$			-0.030***	-0.029***
			(0.005)	(0.004)
$\overline{leverage_i}$			-0.001**	-0.001**
			(0.000)	(0.000)

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

Table 3.6: Decision to export of Chinese manufacturing firms (AME): alternative measure of firm size

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	0.297***	0.287***	0.275***	0.130***
	(0.002)	(0.002)	(0.001)	(0.005)
$IMP_{i(t-1)}$		0.033***	-0.055***	-0.026***
		(0.002)	(0.002)	(0.002)
$age_{i(t-1)}$	-0.007***	-0.008***	-0.008	0.001
	(0.001)	(0.001)	(0.005)	(0.004)
$wage_{i(t-1)}$	0.012***	0.008***	-0.001	-0.000
	(0.001)	(0.001)	(0.002)	(0.001)
$TFP\_DL_{i(t-1)}$	0.001	-0.001	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
$employment_{i(t-1)}$	0.013***	0.012***	-0.003	0.003**
	(0.001)	(0.001)	(0.002)	(0.002)
$liquidity_{i(t-1)}$	-0.008***	-0.008***	0.009**	0.006*
	(0.002)	(0.002)	(0.004)	(0.003)
$leverage_{i(t-1)}$	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$FOREIGN_{i(t-1)}$	0.063***	0.052***	0.034***	0.030***
	(0.003)	(0.003)	(0.003)	(0.003)
$HMT_{i(t-1)}$	0.045***	0.039***	0.029***	0.026***
	(0.003)	(0.003)	(0.003)	(0.003)
$COLLECTIVE_{i(t-1)}$	-0.003	-0.003	-0.000	0.002
	(0.003)	(0.003)	(0.003)	(0.003)
$PRIVATE_{i(t-1)}$	0.030***	0.029***	0.028***	0.028***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP_{i1}				0.106***
				(0.002)
$\overline{IMP_i}$			0.137***	0.106***
			(0.003)	(0.003)
$\overline{age_i}$			0.000	-0.011***
			(0.005)	(0.004)
$\overline{wage_i}$			0.005**	0.006***
			(0.002)	(0.002)
$\overline{TFP\_DL_i}$			-0.004***	-0.004***
			(0.001)	(0.001)
$\overline{employment_i}$			0.015***	0.009***
			(0.002)	(0.002)
$\overline{liquidity_i}$			-0.021***	-0.020***
			(0.005)	(0.004)
$\overline{leverage_i}$			-0.001*	-0.001*
			(0.000)	(0.000)
Observations	170,664	170,664	170,664	170,664

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

Table 3.7: Decision to export of Chinese manufacturing firms (AME): outliers excluded

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	0.300***	0.291***	0.279***	0.158***
	(0.002)	(0.002)	(0.002)	(0.006)
$IMP_{i(t-1)}$		0.034***	-0.055***	-0.034***
		(0.002)	(0.003)	(0.002)
$age_{i(t-1)}$	-0.005***	-0.006***	-0.010*	-0.001
	(0.001)	(0.001)	(0.005)	(0.005)
$wage_{i(t-1)}$	0.008***	0.005***	-0.000	-0.001
	(0.001)	(0.001)	(0.002)	(0.001)
$TFP\_DL_{i(t-1)}$	0.005***	0.004***	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
$liquidity_{i(t-1)}$	-0.015***	-0.014***	0.011***	0.008**
	(0.002)	(0.002)	(0.004)	(0.003)
$leveragei(t-1)$	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$MEDIUMi(t-1)$	0.019***	0.016***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
$LARGEi(t-1)$	0.016***	0.010*	-0.006	-0.006
	(0.005)	(0.005)	(0.005)	(0.005)
$FOREIGNi(t-1)$	0.061***	0.050***	0.030***	0.026***
	(0.004)	(0.004)	(0.004)	(0.004)
$HMTi(t-1)$	0.044***	0.038***	0.027***	0.025***
	(0.004)	(0.004)	(0.004)	(0.004)
$COLLECTIVEi(t-1)$	-0.006	-0.005	-0.005	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)
$PRIVATEi(t-1)$	0.027***	0.026***	0.024***	0.025***
	(0.003)	(0.003)	(0.003)	(0.003)
EXP_{i1}				0.095***
				(0.003)
$\overline{IMP_i}$			0.141***	0.116***
			(0.003)	(0.003)
$\overline{age_i}$			0.004	-0.006
			(0.006)	(0.005)
$\overline{wage_i}$			0.000	0.002
			(0.002)	(0.002)
$\overline{TFP\_DL_i}$			0.003**	0.003**
			(0.001)	(0.001)
$\overline{liquidity_i}$			-0.034***	-0.032***
			(0.005)	(0.004)
$\overline{leverage_i}$			-0.000	-0.000
			(0.000)	(0.001)
Observations	157,901	157,901	157,901	157,901

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

ing firms as well as for policy makers. Since exporting incurs large sunk entry costs, policies aiming to assist firms in overcoming barriers to exporting, such as providing information about potential markets, should be more effective in promoting more export entrants than the current policies of export subsidies and export tax rebates which promote the export volume expansion of existing exporters. Although a large literature has shown that importing has a negative effect on employment and other aspects of the economy, this chapter shows that in developing countries like China, firms that import over a period time can learn from importing and then become future exporters. Policies on importing may be adjusted to some extent, i.e., imports of better quality inputs and embedded with more advanced technology should be encouraged.

Appendix 3A

Table 3A.1: Definition of variables

Variable	Definition
EXP	a binary variable which equals 1 if a firm reports positive exports and 0 otherwise
EXPonly	a binary variable which equals 1 if a firm reports positive exports and zero imports, 0 otherwise
IMP	a binary variable which equals 1 if a firm reports positive imports and 0 otherwise
EXPonly	a binary variable which equals 1 if a firm reports positive imports and zero exports, 0 otherwise
EXPIMP	a binary variable which equals 1 if a firm reports both positive exports and imports, 0 otherwise
impvalue	log of value of a firm's imports
age	log of a firm's age
employment	log of number of employees
wage	log of wage bill divided by the number of employees of a firm
TFP_DL	total factor productivity of a firm obtained by the method of De Loecker (2007)
TFP_LP	total factor productivity of a firm obtained by the method of Levinsohn and Petrin (2003)
labour productivity	log of value-added divided by the number of employees of a firm
capital intensity	log of total assets divided by the number of employees of a firm
liquidity	ratio of current assets minus current liabilities over total assets of a firm
leverage	ratio of total liabilities over total assets of a firm
SOE	a dummy variable which equals 1 if a firm is state-owned and 0 otherwise
COLLECTIVE	a dummy variable which equals 1 if a firm is collectively-owned and 0 otherwise
PRIVATE	a dummy variable which equals 1 if a firm is private-owned and 0 otherwise
FOREIGN	a dummy variable which equals 1 if a firm with over 25% of its capital from foreign investors and 0 otherwise
HTM	a dummy variable which equals 1 if a firm with over 25% of its capital from Hong Kong, Taiwan or Macau investors and 0 otherwise
SMALL	a size dummy variable, which equals 1 if a firm is a small firm and 0 otherwise.
MEDIUM	a size dummy variable, which equals 1 if a firm is a medium firm and 0 otherwise.
LARGE	a size dummy variable, which equals 1 if a firm is a large firm and 0 otherwise.
EAST	a region dummy which equals 1 if a firm is located in the East of China and 0 otherwise.
CENTRAL	a region dummy which equals 1 if a firm is located in Central area of China and 0 otherwise.
WEST	a region dummy which equals 1 if a firm is located in the West of China and 0 otherwise.

Table 3A.2: Decision to export of Chinese manufacturing firms

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	2.997***	2.917***	2.890***	2.034***
	(0.012)	(0.012)	(0.013)	(0.034)
$IMP_{i(t-1)}$		0.338***	-0.582***	-0.420***
		(0.016)	(0.025)	(0.030)
$age_{i(t-1)}$	-0.053***	-0.056***	-0.107**	0.008
	(0.008)	(0.009)	(0.052)	(0.062)
$wage_{i(t-1)}$	0.079***	0.049***	-0.001	-0.010
	(0.010)	(0.010)	(0.016)	(0.019)
$TFP\_DL_{i(t-1)}$	0.046***	0.033***	-0.010	0.005
	(0.006)	(0.006)	(0.011)	(0.013)
$liquidity_{i(t-1)}$	-0.135***	-0.132***	0.094**	0.094**
	(0.021)	(0.021)	(0.040)	(0.048)
$leverage_{i(t-1)}$	-0.001	-0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
$MEDIUM_{i(t-1)}$	0.198***	0.169***	0.098***	0.151***
	(0.015)	(0.016)	(0.017)	(0.023)
$LARGE_{i(t-1)}$	0.169***	0.108**	-0.050	-0.053
	(0.042)	(0.043)	(0.045)	(0.063)
$FOREIGN_{i(t-1)}$	0.604***	0.493***	0.301***	0.381***
	(0.033)	(0.033)	(0.035)	(0.051)
$HMT_{i(t-1)}$	0.432***	0.374***	0.274***	0.356***
	(0.032)	(0.033)	(0.034)	(0.049)
$COLLECTIVE_{i(t-1)}$	-0.081**	-0.080**	-0.076**	-0.070
	(0.035)	(0.035)	(0.036)	(0.051)
$PRIVATE_{i(t-1)}$	0.255***	0.250***	0.231***	0.346***
	(0.031)	(0.031)	(0.032)	(0.046)
EXP_{i1}				1.608***
				(0.069)
$\overline{IMP_i}$			1.466***	1.664***
			(0.031)	(0.045)
$\overline{age_i}$			0.054	-0.117*
			(0.056)	(0.067)
$\overline{wage_i}$			-0.004	0.021
			(0.022)	(0.029)
$\overline{TFP\_DL_i}$			0.032**	0.036**
			(0.013)	(0.017)
$\overline{liquidity_i}$			-0.310***	-0.432***
			(0.048)	(0.062)
$\overline{leverage_i}$			-0.009**	-0.015**
			(0.004)	(0.007)
Constant	-2.398***	-2.211***	-1.952***	-2.440***
	(0.066)	(0.067)	(0.073)	(0.110)
Observations	170,664	170,664	170,664	170,664
log likelihood	-32,665	-32,430	-31,269	-30,627

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

Table 3A.3: Decision to export of Chinese manufacturing firms: alternative TFP estimation

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	2.996***	2.917***	2.890***	2.033***
	(0.012)	(0.012)	(0.013)	(0.034)
$IMP_{i(t-1)}$		0.337***	-0.582***	-0.420***
		(0.016)	(0.025)	(0.030)
$age_{i(t-1)}$	-0.053***	-0.057***	-0.107**	0.008
	(0.008)	(0.009)	(0.052)	(0.062)
$wage_{i(t-1)}$	0.078***	0.048***	-0.001	-0.010
	(0.010)	(0.010)	(0.016)	(0.019)
$TFP\_LP_{i(t-1)}$	0.049***	0.036***	-0.009	0.005
	(0.006)	(0.006)	(0.011)	(0.013)
$liquidity_{i(t-1)}$	-0.136***	-0.133***	0.094**	0.094**
	(0.021)	(0.021)	(0.040)	(0.048)
$leverage_{i(t-1)}$	-0.001	-0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
$MEDIUM_{i(t-1)}$	0.194***	0.165***	0.093***	0.145***
	(0.016)	(0.016)	(0.017)	(0.023)
$LARGE_{i(t-1)}$	0.159***	0.099**	-0.061	-0.067
	(0.042)	(0.043)	(0.045)	(0.064)
$FOREIGN_{i(t-1)}$	0.603***	0.493***	0.300***	0.380***
	(0.033)	(0.033)	(0.035)	(0.051)
$HMT_{i(t-1)}$	0.432***	0.373***	0.273***	0.356***
	(0.032)	(0.033)	(0.034)	(0.049)
$COLLECTIVE_{i(t-1)}$	-0.082**	-0.080**	-0.077**	-0.072
	(0.035)	(0.035)	(0.036)	(0.051)
$PRIVATE_{i(t-1)}$	0.255***	0.249***	0.229***	0.344***
	(0.031)	(0.031)	(0.032)	(0.046)
EXP_{i1}				1.608***
				(0.069)
$\overline{IMP_i}$			1.464***	1.662***
			(0.031)	(0.045)
$\overline{age_i}$			0.054	-0.117*
			(0.056)	(0.067)
$\overline{wage_i}$			-0.006	0.019
			(0.022)	(0.029)
$\overline{TFP\_LP_i}$			0.035***	0.041**
			(0.013)	(0.017)
$\overline{liquidity_i}$			-0.311***	-0.433***
			(0.048)	(0.062)
$\overline{leverage_i}$			-0.009**	-0.015**
			(0.004)	(0.007)
Constant	-2.413***	-2.226***	-1.972***	-2.467***
	(0.066)	(0.067)	(0.073)	(0.110)
Observations	170,664	170,664	170,664	170,664
log likelihood	-32,661	-32,427	-31,267	-30,625

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

Table 3A.4: Decision to export of Chinese manufacturing firms: alternative measure of import activity

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	2.997***	2.931***	2.930***	2.058***
	(0.012)	(0.012)	(0.012)	(0.034)
$impvalue_{i(t-1)}$		0.023***	-0.050***	-0.036***
		(0.001)	(0.002)	(0.002)
$age_{i(t-1)}$	-0.053***	-0.055***	-0.117**	-0.003
	(0.008)	(0.009)	(0.052)	(0.062)
$wage_{i(t-1)}$	0.079***	0.045***	0.001	-0.009
	(0.010)	(0.010)	(0.016)	(0.019)
$TFP\_DL_{i(t-1)}$	0.046***	0.031***	-0.009	0.005
	(0.006)	(0.006)	(0.011)	(0.013)
$liquidity_{i(t-1)}$	-0.135***	-0.132***	0.090**	0.092*
	(0.021)	(0.021)	(0.040)	(0.047)
$leverage_{i(t-1)}$	-0.001	-0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
$MEDIUM_{i(t-1)}$	0.198***	0.171***	0.115***	0.171***
	(0.015)	(0.016)	(0.016)	(0.023)
$LARGE_{i(t-1)}$	0.169***	0.099**	-0.048	-0.050
	(0.042)	(0.043)	(0.045)	(0.064)
$FOREIGN_{i(t-1)}$	0.604***	0.500***	0.352***	0.448***
	(0.033)	(0.033)	(0.034)	(0.050)
$HMT_{i(t-1)}$	0.432***	0.379***	0.305***	0.399***
	(0.032)	(0.032)	(0.033)	(0.049)
$COLLECTIVE_{i(t-1)}$	-0.081**	-0.083**	-0.079**	-0.075
	(0.035)	(0.035)	(0.035)	(0.051)
$PRIVATE_{i(t-1)}$	0.255***	0.247***	0.231***	0.347***
	(0.031)	(0.031)	(0.032)	(0.046)
EXP_{i1}				1.645***
				(0.072)
$\overline{impvalue_i}$			0.104***	0.111***
			(0.002)	(0.003)
$\overline{age_i}$			0.068	-0.102
			(0.056)	(0.067)
$\overline{wage_i}$			-0.005	0.022
			(0.022)	(0.029)
$\overline{TFP\_DL_i}$			0.028**	0.033**
			(0.013)	(0.017)
$\overline{liquidity_i}$			-0.311***	-0.436***
			(0.048)	(0.062)
$\overline{leverage_i}$			-0.009**	-0.016**
			(0.004)	(0.007)
Constant	-2.398***	-2.185***	-1.940***	-2.431***
	(0.066)	(0.067)	(0.073)	(0.110)
Observations	170,664	170,664	170,664	170,664
log likelihood	-32,665	-32,469	-31,526	-30,858

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

Table 3A.5: Decision to export of Chinese manufacturing firms (AME): alternative measure of firm size

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	2.980***	2.903***	2.875***	2.005***
	(0.012)	(0.012)	(0.013)	(0.033)
$IMP_{i(t-1)}$		0.330***	-0.569***	-0.407***
		(0.016)	(0.025)	(0.030)
$age_{i(t-1)}$	-0.075***	-0.079***	-0.081	0.023
	(0.009)	(0.009)	(0.053)	(0.063)
$wage_{i(t-1)}$	0.116***	0.084***	-0.011	-0.005
	(0.010)	(0.011)	(0.017)	(0.020)
$TFP\_DL_{i(t-1)}$	0.008	-0.006	-0.007	0.002
	(0.006)	(0.007)	(0.011)	(0.013)
$employment_{i(t-1)}$	0.133***	0.124***	-0.032	0.048**
	(0.006)	(0.006)	(0.020)	(0.024)
$liquidity_{i(t-1)}$	-0.082***	-0.078***	0.090**	0.093*
	(0.021)	(0.021)	(0.040)	(0.048)
$leverage_{i(t-1)}$	-0.001	-0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.003)
$FOREIGN_{i(t-1)}$	0.634***	0.530***	0.356***	0.461***
	(0.033)	(0.033)	(0.035)	(0.051)
$HMT_{i(t-1)}$	0.448***	0.394***	0.307***	0.407***
	(0.032)	(0.033)	(0.034)	(0.050)
$COLLECTIVE_{i(t-1)}$	-0.031	-0.026	-0.000	0.036
	(0.035)	(0.035)	(0.036)	(0.052)
$PRIVATE_{i(t-1)}$	0.297***	0.294***	0.293***	0.437***
	(0.031)	(0.031)	(0.032)	(0.047)
EXP_{i1}				1.633***
				(0.068)
$\overline{IMP_i}$			1.431***	1.635***
			(0.031)	(0.045)
$\overline{age_i}$			0.000	-0.176***
			(0.056)	(0.068)
$\overline{wage_i}$			0.056**	0.091***
			(0.023)	(0.030)
$\overline{TFP\_DL_i}$			-0.040***	-0.063***
			(0.014)	(0.018)
$\overline{employment_i}$			0.152***	0.136***
			(0.021)	(0.026)
$\overline{liquidity_i}$			-0.221***	-0.304***
			(0.048)	(0.063)
$\overline{leverage_i}$			-0.008*	-0.013*
			(0.004)	(0.007)
Constant	-2.924***	-2.685***	-2.269***	-2.954***
	(0.067)	(0.068)	(0.072)	(0.114)
Observations	170,664	170,664	170,664	170,664
log likelihood	-32,512	-32,286	-31,145	-30,499

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

Table 3A.6: Decision to export of Chinese manufacturing firms: outliers excluded

	(1)	(2)	(3)	(4)
$EXP_{i(t-1)}$	3.016***	2.937***	2.909***	2.203***
	(0.012)	(0.013)	(0.013)	(0.033)
$IMP_{i(t-1)}$		0.342***	-0.579***	-0.467***
		(0.016)	(0.026)	(0.030)
$age_{i(t-1)}$	-0.055***	-0.059***	-0.100*	-0.018
	(0.009)	(0.009)	(0.055)	(0.063)
$wage_{i(t-1)}$	0.081***	0.051***	-0.002	-0.011
	(0.011)	(0.011)	(0.018)	(0.020)
$TFP\_DL_{i(t-1)}$	0.049***	0.035***	-0.009	0.003
	(0.007)	(0.007)	(0.012)	(0.013)
$liquidity_{i(t-1)}$	-0.146***	-0.143***	0.110***	0.112**
	(0.023)	(0.023)	(0.042)	(0.049)
$leverage_{i(t-1)}$	-0.000	-0.000	0.001	0.001
	(0.002)	(0.002)	(0.003)	(0.003)
$MEDIUM_{i(t-1)}$	0.192***	0.163***	0.089***	0.126***
	(0.016)	(0.016)	(0.017)	(0.022)
$LARGE_{i(t-1)}$	0.162***	0.097*	-0.058	-0.084
	(0.052)	(0.053)	(0.055)	(0.071)
$FOREIGN_{i(t-1)}$	0.617***	0.505***	0.312***	0.366***
	(0.036)	(0.037)	(0.038)	(0.051)
$HMT_{i(t-1)}$	0.445***	0.384***	0.284***	0.342***
	(0.036)	(0.036)	(0.037)	(0.050)
$COLLECTIVE_{i(t-1)}$	-0.056	-0.054	-0.051	-0.035
	(0.038)	(0.038)	(0.039)	(0.052)
$PRIVATE_{i(t-1)}$	0.273***	0.267***	0.246***	0.346***
	(0.034)	(0.035)	(0.036)	(0.048)
EXP_{i1}				1.321***
				(0.065)
$\overline{IMP_i}$			1.468***	1.611***
			(0.033)	(0.044)
$\overline{age_i}$			0.045	-0.082
			(0.059)	(0.068)
$\overline{wage_i}$			0.004	0.028
			(0.024)	(0.030)
$\overline{TFP\_DL_i}$			0.033**	0.037**
			(0.014)	(0.017)
$\overline{liquidity_i}$			-0.350***	-0.447***
			(0.051)	(0.062)
$\overline{leverage_i}$			-0.004	-0.006
			(0.005)	(0.007)
Constant	-2.414***	-2.221***	-1.978***	-2.351***
	(0.071)	(0.072)	(0.080)	(0.109)
Observations	157,901	157,901	157,901	157,901
log likelihood	-30102	-29882	-28813	-28275

Notes: Standard errors in parentheses. All specifications include region, industry and year dummies. ***, ** and * denote significance at 0.01, 0.5 and 0.1 respectively.

Chapter 4

Firm Productivity and Importing

The relationship between firm performance and international trade in a period of rapid globalisation is of interest to academics and policymakers. In this chapter we investigate various aspects of the relationship between firm productivity and importing for a large sample of Chinese firms between 2002 and 2006. Measuring total factor productivity treating imports as endogenous following De Loecker (2007) we test the learning-by-doing and self-selection hypotheses by random effects Probit and propensity score matching with difference-in-differences (PSM-DID) techniques, taking into account differences in firm size, ownership structure, origin of imports, skill and technology content and number of varieties of imported inputs. Our results show evidence of a bi-directional causal relationship between importing and productivity. Although importing firms tend to be more productive before entering the import market, once they start importing firms appear to experience significant productivity gains for up to two years following entry. We also find that except for collective firms, all other ownership types of firm experience a positive learning effect. Small and medium sized firms also experience larger and longer lasting productivity gains. Finally, we find that import starters that source their inputs from high income countries, source high skill and technology content goods and importing more varieties experiencing the higher productivity gains.

4.1 Introduction

In an increasingly globalized world, the relationship between firm performance and the extent to which a firm is internationalized is subject to ever greater scrutiny from academics and policy makers. In recent years, the general trend has been for firms to export to and import from an ever increasing number of countries and in ever greater volumes. The internationalization of firms is especially important for developing and newly industrialising countries that continue to pursue an export-led growth strategy and remain dependent to a large extent on exports for future growth and employment. A recent example is China where the common perception is that the rapid growth over the last two decades has been driven by exports to the West. However, less well documented are the benefits to Chinese firms from process of importing raw materials and intermediate inputs from abroad.[1]

The motivation for this chapter is to investigate how engagement with global trade (importing and exporting) impacts firm performance and how this is influenced by firm size, ownership and where a firm sources its intermediate inputs. Central to our analysis is an investigation of the causal relationship between productivity and importing. That is to say, whether more productive firms decide to import (self-selection hypothesis) or whether those firms that start

[1]Between 2000 and 2010 China's trade balance increased from USD24.1 billion to over USD181.51 billion. Imports during this period increased from USD225 billion to USD1,396.24 billion while exports rose from USD249.2 billion to USD1,577.75 billion over the same period (China Statistical Yearbook 2013).

importing increase their productivity (learning-by-doing hypothesis). Our data consist of a large sample of Chinese manufacturing firms between 2002 and 2006. We chose to study China and the period immediately after China's 2001 entry into the WTO for two reasons. First, because of China's increasingly important role in the global economy and second, because this period represents the beginning of the transition of Chinese firms from relatively low skilled assembly tasks to more high technology, high value added production. It is therefore useful to understand the role of imports in this transition process.

The existing literature has tended to concentrate on the determinants of export participation categorizing firms into exporters and non-exporters (Baldwin and Gu, 2004; Lileeva and Trefler, 2010; Roberts and Tybout, 1997; Ruane and Sutherland, 2005) and on the causal relationship between exporting and productivity (Arnold and Hussinger, 2005; Girma et al., 2004; Wagner, 2007). Studies specific to China have also tended to concentrate on exporting and include Kraay (1999), Du and Girma (2007),Yang and Mallick (2010), Lu et al. (2010), Sun and Hong (2011) and Yi and Wang (2012). A feature of these papers is that they have tended to ignore importing even though we know that a large number of exporters also import. Likewise, a significant number of non-exporters source significant levels of raw materials and intermediate inputs from abroad. Therefore, if importing promotes exporting, ignoring import activity will upwardly bias the estimated premium associated with exporting.

Perhaps surprisingly, research on importing is still fairly limited certainly compared to the number of studies on exporting. Previous research on firm importing behavior includes Castellani et al. (2010); Halpern et al. (2005); Kasahara and Lapham (2013); Kugler and Verhoogen (2009); MacGarvie (2006). The characteristics of importers tend to be similar to those in the exporting literature which is to say that importers are larger (generally larger than exporters), more productive and more capital-intensive than non-importers. In terms of productivity, a positive and significant productivity differential between importing and non-importing firms has been found by Halpern et al. (2005) and Andersson et al. (2008). Research into the impact of importing on Chinese firms is more limited with the exception of working papers by Manova and Zhang (2009) and Feng et al. (2012). Finally, in a related literature, a small number of studies have examined the performance of both exporting and importing firms (Aristei et al., 2013; Bernard et al., 2005, 2007; Kasahara and Rodrigue, 2008; Muuls and Pisu, 2009) and find that two-way traders outperform non-traders and one-way traders (firms that only export or import) in terms of productivity and size.

Studies that investigate the direction of causality between productivity and importing, to the best of our knowledge, are limited to McCann (2009), Vogel and Wagner (2010) and Augier et al. (2013). Most recently Augier et al. (2013) study Spanish firms and find an insignificant productivity effect of switching to importing, although when firms both import and have a large share of skilled labour there is some evidence of a learning effect from the use of imported intermediates suggesting an absorptive capacity effect. Vogel and Wagner (2010) use German manufacturing firm data and find some evidence for self-selection of more productive firms into importing but no support for the learning effect of importing on productivity. Finally, McCann (2009) uses the Irish Census of Industrial Production to study productivity gains from international trade and finds that becoming an exporter significantly increases total factor productivity (TFP) but for firms that become importers there is no such effect.

The contribution of this chapter is three-fold. First, to the best of our knowledge, this is the first paper to investigate the causal relationship between importing and firm productivity for Chinese manufacturing firms. Our approach is to merge Chinese industrial enterprise survey data with transaction-level customs data to provide a uniquely rich dataset from which to analyse the international activity of Chinese manufacturing firms. The comprehensive nature of the data means we are able to examine various aspects of the import performance relationship by estimating the impact of firm size, ownership structure and the source of intermediate products on firm productivity. Second, we measure productivity using a relatively new measure of total factor productivity (TFP) using a modified algorithm by De Loecker (2007) where firm import status is endogenous and introduced to all stages of the estimation procedure to generate our measure of productivity. This allows us to control for simultaneity and selection bias and allows for different market structures, demand conditions and factor markets for importing and non-importing firms. The introduction of import status as an additional state variable in the production function means the import status has a dynamic effect on the evolution of productivity.[1] Third, our methodological approach is to combine propensity score matching with difference-in-differences (PSM-DID) techniques to examine the relationship between firm productivity and importing behaviour using narrowly defined 2-digit industries in the matching mechanism. Such an approach means that we can control for unobserved firm level heterogeneity more effectively. It is only because our data includes such a large number of firms that we have sufficient sample size to match firms at levels of detail to allow us address questions previously overlooked.

Before we describe our methodology in detail we briefly rehearse the arguments for the self-selection hypothesis and learning-by-doing hypothesis, usually discussed from an exporting

[1] A similar approach has been employed by Van Biesebroeck (2005) and Kasahara and Rodrigue (2008).

perspective, in an importing context. The self-selection argument is that firms that want to start importing have to incur sunk costs which include the costs associated with the search for information on possible inputs (for example, to find out which foreign firms in which countries can supply the required inputs), learning to navigate often complex customs procedures and having to understand tax and trade credit regulations (Andersson et al., 2008; Castellani et al., 2010). Additional fixed costs may include quality inspection costs and those transport costs incurred by importers (Andersson et al., 2008; Kasahara and Lapham, 2013). The self-selection hypothesis assumes that only highly productive firms can afford to incur these costs and the additional risks associated with trade and hence profit from importing. One may therefore expect firms to improve their productivity before entering into the process of importing.

In contrast, the learning-by-doing hypothesis argues that it is the very act of engaging in import activity with firms able to access to products or inputs at lower prices, getting access to a broader range of inputs, or finding inputs of better quality than are available in the domestic market that helps to drive productivity growth (Halpern et al., 2005; Muuls and Pisu, 2009). It is hypothesised that importing firms are able to extract knowledge and learn about the technology embodied in the imported inputs which may eventually contribute to improved production efficiencies at home (Andersson et al., 2008; Castellani et al., 2010). The learning-by-importing hypothesis also argues that access to foreign markets is a source of international technology transfer as firms can adopt advanced manufacturing technologies from their trading partners and engage in more product innovation at lower costs both of which boost firm-level productivity (Blalock and Veloso, 2007; Goldberg et al., 2009; Topalova and Khandelwal, 2011). The other benefit to importing is that new foreign inputs increase the ability of firms to manufacture new varieties or improve the quality of existing varieties. [1]

Our main finding is of a bi-directional causal relationship between firm productivity and importing for Chinese manufacturing firms. That is to say, more productive firms self-select into the import market and this process of importing also further enhances firm productivity. Our PSM-DID results are broadly consistent and robust to various sensitivity checks. In contrast to McCann (2009), Vogel and Wagner (2010) and Augier et al. (2013) we find that Chinese firms appear to have the capacity to absorb new technologies and production techniques from the West to learn from importing and subsequently improve productivity. Our results also show that except for collective firms, all new import starters of other ownership types demonstrate a learning effect. The strongest learning effect is generally exhibited by the SOE importers, followed by Hong Kong, Macao and Taiwan (HMT)-owned importers, private importers and foreign-owned importers. We also find that small and medium sized firms experience higher, and in some cases longer lasting, productivity growth than larger new importers. Furthermore, we find that new importers who import from high income countries experience greater productivity gains than those that import from lower income countries. Finally new importers who import medium and high skill and technology intensive products, have lower TFP level initially and importing more varieties of inputs display stronger learning effects.

The remainder of this chapter is organized as follows. Section 4.2 presents our identification strategy and estimation methodology. Section 4.3 describes the data and the construction of our dataset, followed by Section 4.4 which reports and discusses our results. Section 4.5 presents

[1] In a related literature, recent empirical studies have found that imported intermediate inputs and/or a decline in input tariffs are associated with significant productivity gains, quality upgrading and increased exports (Acharya and Keller, 2009; Amiti and Khandelwal, 2013; Amiti and Konings, 2007; Bas and Strauss-Kahn, 2014; Goldberg et al., 2010; Kasahara and Rodrigue, 2008).

various robustness checks and Section 4.6 concludes.

4.2 Empirical Methodology

4.2.1 Identification Strategy

Our methodological approach is to employ propensity score matching with difference-in-differences (PSM-DID) techniques to identify the direction and magnitude of any causal effects of importing on firm productivity.

Our primary motivation is to discover whether there is any change in a firm's productivity growth following entry into an import market. We define y_{it} as firm *i*'s total factor productivity (TFP) at time *t* and $Y_{i(t+s)}$ as the productivity *s* period(s) later $(s \geq 0)$. The causal effect of importing on productivity of firm *i* at *t*+*s* can be identified by looking at the difference:

$$y^1_{i(t+s)} - y^0_{i(t+s)} \tag{4.1}$$

where the superscripts denote import behaviour which is equal to one if a firm imports at *t* and 0 otherwise. Hence $y^0_{i(t+s)}$ represents the productivity of firm *i* at period *t*+*s* *had* it not participated in import markets since time *t*.

The fundamental evaluation problem is that only one of the two outcomes of (1) is observable. For example, if $y^1_{i(t+s)}$, is observed for firm *i*, then $y^0_{i(t+s)}$, the counter-factual outcome, is not observed. This means a direct estimation of the individual treatment effect is not possible. Hence, we need to calculate the population average treatment effect (ATE) which is the difference in the expected outcomes of participants and non-participants where:

$$ATE = E[y^1_{i(t+s)} - y^0_{i(t+s)}] \tag{4.2}$$

In order to identify differences in firm productivity after a firm begins to import we need to identify what we term import starters. Hence, we define a dummy variable $START_{it}$ which is equal to one if firm *i* begins to import at time *t* and 0 otherwise. The average productivity effect that new importers would have experienced if they had not previously imported, i.e., the average treatment effect on the treated (ATT), is given by:

$$ATT = E[y^1_{i(t+s)} - y^0_{i(t+s)} | START_{it} = 1]$$
$$= E[y^1_{i(t+s)} | START_{it} = 1] - E[y^0_{i(t+s)} | START_{it} = 1] \tag{4.3}$$

Likewise, the counter-factual which is the average productivity of new importers, $E[y^0_{i(t+s)} | START_{it} = 1]$, is not observed. However, as Heckman et al. (1998) point out, the average productivity of an appropriate control group of non import starters, i.e., $E[y^0_{i(t+s)} | START_{it} = 0]$, can be used as a substitute. Hence, equation (3) can be rewritten as:

$$ATT = E[y^1_{i(t+s)} | START_{it} = 1] - E[y^0_{i(t+s)} | START_{it} = 0] \tag{4.4}$$

To select a valid control group we employ a matching approach. The purpose of matching is to pair each new importers with a firm that has not never entered the imports market on the basis of a set of observable characteristics. First, we estimate the probability of a firm entering the imports market (or the propensity score) by estimating the following Probit model:

$$Pr(START_{it} = 1) = \phi(age_{i(t-1)}, wage_{i(t-1)}, TFP_{i(t-1)},$$
$$pregrowth_{it}, EXP_{i(t-1)}, size_{i(t-1)}, ownership_{i(t-1)}, D_r, D_j, D_t) \qquad (4.5)$$

where *Pr* denotes the predicted probability of firm *i* starting to import at time t, and $\phi(.)$ is the normal cumulative distribution function. Firm characteristics such as age, average employee wages, TFP, past productivity growth rate ($pregrowth$), export status, size and ownership are included in the estimation of the propensity score. Taking into account the past productivity growth rate is important if it is autocorrelated over time (Girma et al., 2007). Without control for it, we would mistakenly attribute a causal effect to importing on post entry productivity growth as it could be that firms that start importing were already on a permanently different growth trajectory and the switch just picks up that. Full set of region dummies (D_r), industry dummies (D_j) and year dummies (D_t) are also included to capture location, industry and time effects. All time-variant explanatory variables are lagged by one year in order to mitigate simultaneity concerns.

With the estimated propensity score, the next step is to check that propensity score is balanced across treated and control groups. Following Imbens (2004), De Loecker (2007) and Garrido et al. (2014), we split the sample in *k* equally spaced intervals of the propensity score and test within each interval whether the mean propensity score is equivalent in treatment and comparison groups. If not equivalent, split up the interval into smaller blocks and test again and continue this until equality holds for every interval. After the propensity score is balanced within blocks across the treated and control groups, we check for the balance of each observed covariates within blocks of the propensity score. If the balance test is rejected, covariates included in the propensity score estimation can be modified, such as recategorizing variables or including higher order terms or splines of variables.

After creating a balanced propensity score, we start matching the import starters with a group of non-importing firms in a way that the estimated propensity score of a non-importing firm is as close as possible to that of a new importer. Several matching algorithms have been developed, e.g., nearest neighbour matching, calliper and radius matching, kernel matching and stratification matching.[1] Our matching is based on and we impose the common support condition by dropping the importing starters whose propensity scores are higher than the maximum or lower than the minimum of those persistent non-importers. In kernel matching, each treated firm is given a weight of one and control firms are weighted by the distance in propensity score from the treated firm within a range, i.e., bandwidth, of the propensity score. Kernel matching maximizes precision as more information is used than other matching algorithms by retaining the sample size as only observations outside the range of common support are discarded (Garrido et al., 2014). The choice of bandwidth is important which leads to a tradeoff between bias and variance (Garrido et al., 2014; Silverman, 1998). High bandwidth values yield a smoother estimated density function, therefore leading to a better fit and a decreasing variance between the estimated and the true underlying density function. However, bias arises from selecting a wide bandwidth as underlying features may be smoothed away (Caliendo and Kopeinig, 2008).

[1]Although all matching estimators construct the differences between the outcome of a treated individual and outcomes of units from the control group, they vary from each other in terms of how the neighbourhood for the treated individual is defined, how the weights are assigned to these neighbours and how the common support problem is handled, and each of them has its own advantages and disadvantages. See Caliendo and Kopeinig (2008) for discussion on these issues and practical guidance for the implementation of propensity score matching.

Given that we have a large sample as will be shown in the next section and aiming for an unbiased estimate, we choose a small bandwidth of 0.01.

Rather than matching across the entire manufacturing sector (Girma et al., 2004; Vogel and Wagner, 2010), our matching is performed separately for each 2-digit industry of the manufacturing sector within each year.[1] In this way we create control groups within narrowly defined industries of the same year. This is important as firms in different industries face different technological and market conditions and the marginal effects of such variables on the propensity to enter the import market of these firms may differ substantially between different industries. Similarly if matching is not done within each year, an import starter in the treatment year can be matched with a control firm in *any* year or even worse, itself after its entry.

Having constructed the control group of firms (C) that are similar to the treatment firms (T) by propensity score matching, we use a difference-in-differences (DID) methodology to estimate the causal effect of importing on productivity.

A DID estimator first measures the difference in productivity before and after entry in to the import market for importing firms conditioned on past performance and a set of dummy variables. However, such differences in productivity cannot exclusively be attributed to importing behaviour as post-entry productivity growth might be caused by factors that are contemporaneous with entry into the import market. The second step is to difference the difference obtained for the import starters with the corresponding difference for non-importing firms. Since DID estimates the difference before treatment it removes the effects of common shocks and hence provides a more accurate estimate.

[1] We create a bin for each industry-year category and add the estimated propensity score to 10 times of the bin number, creating large wedges in propensity scores between bins to force the matching to be within bins.

As Blundell and Costa Dias (2000, p.438) point out, a non-parametric approach that combines propensity score matching with difference-in-differences has the potential "...to improve the quality of non-experimental evaluation results significantly". Hence, we combine PSM with DID such that the selection on unobservable determinants can be allowed when the determinants lie on separable firm and /or time-specific components of the error-term. Hence, imbalances in the distribution of covariates between the treated and control groups account for varying unobserved effects influencing importing and productivity. Our PSM-DID estimator based on a sample of matched firms is therefore given by:

$$ATT^{PSM-DID} = \frac{1}{N_T} \sum_{i \in T} [\Delta y_{i(t+s)} - \sum_{j \in C} w_{ij} \Delta y_{j(t+s)}] \tag{4.6}$$

where N_T is the number of treated units (firms that start to import) in the common support, Δy_i is the difference between the average productivity before and after the entry into the imports market of firm i from the treatment group, Δy_j is the before and after difference of firm j from the control group and w_{ij} is the weight placed on the control firm j in the construction of the estimated expected counterfactual outcome for treated firm i, determined by the propensity score matching algorithm, $\sum_{j \in C} w_{ij} = 1$. As matching is always performed at time t when a firm starts importing, $\Delta y_{i(t+s)}$ presents the productivity growth s periods after the decision to start importing compared to the year before the entry to imports market.

4.2.2 Assessing the Propensity Score Matching Quality

In this chapter we perform several balancing tests suggested in the literature to assess the quality of our propensity score matching (e.g., Austin, 2009; Caliendo and Kopeinig, 2008; Rosenbaum and Rubin, 1985; Smith and Todd, 2005). We first compare the situation before and after the matching and check if any differences in means of the observable characteristics for firms from treatment and control groups remain after conditioning on the propensity score. Differences between both groups are expected before matching, but these differences should be reduced significantly after matching. A formal two-sample *t*-test between the treated and control groups for each variable is performed to ensure that no significant bias exists.[1]

The second test is to examine the standardized difference (SD) (or standardized bias) for all variables used in the PSM. The lower the standard difference, the more balanced the treated and control groups will be in terms of the variable under consideration. Standardized differences for comparing means between groups are computed as follows. For continuous variables, the standardized difference is defined as:

$$SD = 100\frac{\bar{X}_T - \bar{X}_C}{\sqrt{\frac{s_T^2+s_C^2}{2}}} \tag{4.7}$$

where $\bar{X}_T$ and $\bar{X}_C$ denote the sample mean of the variable X in treated and control groups, respectively, while s_T^2 and s_C^2 denote the sample variance of the variable in treated and control groups, respectively.

For dichotomous variables, the standardized difference is defined as:

$$SD = 100\frac{\hat{P}_T - \hat{P}_C}{\sqrt{\frac{\hat{P}_T*(1-\hat{P}_C)+\hat{P}_C*(1-\hat{P}_C)}{2}}} \tag{4.8}$$

[1]Caution needs to be paid if using *t*-tests to check balance of covariates. Because the goal of matching is to ensure balance within a sample, the larger population from which the sample was drawn is not of concern. Moreover, *t*-tests are affected by sample size and might not be statistically significant even in the presence of covariate imbalance (Austin, 2009; Garrido et al., 2014; Ho et al., 2007).

where $\hat{P}_T$ and $\hat{P}_C$ denote the mean of the dichotomous variable P in treated and comparison groups, respectively.

Unlike t-tests, the standardized difference is not influenced by sample size. Thus, the use of the standard difference can be used to compare balance in measured variables between treated and the control in the matched sample with that in the unmatched sample (Austin, 2009). There are no formal criteria in the literature for when a standardized bias is too large. Rosenbaum and Rubin (1985) suggest that a value of 20% of standardized difference is large. We follow Augier et al. (2013) and Garrido et al. (2014) and use the same criteria.

Also as Sianesi (2004) suggests we reestimate the propensity score on the matched sample and compare the pseudo-R^2s before and after matching. The pseudo-R^2 indicates how well the variables X explain the participation probability. After matching there should be no systematic differences in the distribution of covariates between both groups and therefore the pseudo-R^2 should be fairly low. Furthermore, we also perform a likelihood ratio test on the joint significance of all variables in the Probit model. The test should not be rejected before matching and should be rejected after matching.

4.3 Data and Descriptives

4.3.1 Data

The data used in this chapter are drawn from two sources. The firm-level production data are from the Annual Survey of Industrial Enterprises provided by the National Bureau of Statistics of China (NBS) and the transaction-level trade data are provided by the Department of Customs Trade Statistics, the General Administration of Customs of China. We use data from both sources for the years 2002 to 2006.

The NBS survey data includes the firm's identification (tax code) and basic information such as year founded, location, ownership type, employment, China industrial classification (CIC) code and principle products.[1] This chapter focuses exclusively on the manufacturing sector.

Trade data record all import and export transactions with non-zero values that enter or exit through Chinese customs. Each observation represents a shipment and contains detailed information on the time of the transaction (month and year), type of trade (import/ export), exporting/importing firm identifier, ownership type, product traded (8-digit HS code and name), value, quantity, unit, destination country (of the exported commodities) / country of origin (of the imported products), type of trade (ordinary trade, processing trade, compensation trade, consignment, etc.) and finally mode of transport.

To exploit the information on firm production and international trade, the next step is to match the NBS survey data with the trade data. Please see Chapter 2 for detailed description of the methodology of merging the two datasets as well as construction of key variables such as TFP, firm size and ownership.

[1]The data also provides information on more than 50 financial variables from the accounting statements, including capital, assets, liabilities, creditors equity, gross output, industrial value-added, sales, income, profits, investment, value of exports, current / accumulated depreciation, the wage bill and R&D expenses.

4.3.2 The Internationalisation of Chinese Firms

In this section we describe the characteristics of Chinese firms and the extent to which they engage in international trade. The firm heterogeneity and exporting literature has shown that, broadly speaking, exporters are larger, more productive, more capital- and skill-intensive and pay higher wages than non-exporters. These studies have tended to focus only on firm exporting status and categorized firms into two mutually exclusive groups, i.e., exporters and non-exporters, ignoring any firm import activity. However, a large number of exporters also import at the same time. Similarly, there will be a number of non-exporters who also import. Firms that export and import at the same time may perform rather differently from those that only export and ignoring import activities may lead to an upward bias in the estimated export premia. Likewise, ignoring exports may bias the impact of importing on productivity. In order to get a better picture of the international activities of the Chinese manufacturing firms, we divide our sample into four categories: exporter-only (firms that export but do not import), importer-only (firms that import but do not export), two-way traders (firms that both export and import) and non-traders (firms that neither export nor import).

First, we examine the number of firms in each category and their average output. Table 4.1 documents the participation of Chinese manufacturing firms in international trade. On average, between 2000 and 2006, about 70% of firms are classified as non-traders. Of the other 30%, more than half (around 17%) import and export and of the rest 7% export only and 6% import only.[1] When we consider average output, two-way traders are the largest, followed by importers, exporters and finally non-traders whose output is one-quarter that of two-way traders.

[1] Although we are looking at SOEs and relatively large firms with output greater than 5 million RMB we still find similar participation rates to those found in Andersson et al. (2008) for Swedish manufacturing firms, Muuls and Pisu (2009) for Belgium firms and Vogel and Wagner (2010) for German manufacturing firms.

Hence, although non-traders are about 70% of the total sample, their output accounts just about 40% of total output. In contrast, two-way traders make up just 17% of total firms yet contribute to more than 40% of total output. Similar patterns are observed for individual years.

Table 4.1: Export and import participation of Chinese manufacturing firms

	Trade status							
Year	Exporters-Only		Importers-Only		Two-way traders		Non-traders	
	No. of firms (% of total)	Avg output (% of total)	No. of firms (% of total)	Avg output (% of total)	No. of firms (% of total)	Avg output (% of total)	No. of firms (% of total)	Avg output (% of total)
2002	2,438	72.62	2,768	127.03	6,907	198.77	31,524	54.59
	5.59	4.89	6.34	9.71	15.83	37.90	72.24	47.51
2003	2,858	73.40	2,663	145.34	7,382	244.12	30,734	65.86
	6.55	4.74	6.10	8.75	16.92	40.74	70.43	45.76
2004	3,302	73.85	2,712	163.67	8,015	277.18	29,608	70.94
	7.57	4.87	6.21	8.86	18.37	44.35	67.85	41.93
2005	3,680	98.04	2,580	195.79	8,072	312.43	29,305	81.92
	8.43	6.23	5.91	8.73	18.50	43.57	67.16	41.47
2006	4,012	93.01	2,511	238.18	7,762	356.05	29,352	93.66
	9.19	5.75	5.75	9.22	17.79	42.62	67.26	42.40
2002-2006	16,290	83.77	13,234	172.71	38,138	280.09	150,523	73.05
	7.47	5.39	6.07	9.02	17.48	42.18	68.99	43.41

Notes: Only-exporters refer to firms that export but do not import. Only importers refer to firms that import but do not export. Two-way traders are firms that both export and import while non-traders are those that neither export nor import. Average output is measured in millions of Yuan and deflated in 2000 prices.

The previous literature highlights the importance of ownership structure in studies of firm heterogeneity and international trade. Table 4.2 presents the number and share of firms and output of each ownership type (SOE, collective, private, foreign- and HMT-owned) in the domestic and international markets. The final two columns of Table 2 show that just over a half of the firms in our sample (50.57%) are private domestic firms, 15% are foreign-owned and another 15% are HMT-owned firms. Private firms also have the largest total output, followed by foreign-firms who produce more than twice the amount produced by HMT and state-owned firms. The vast majority of domestic firms serve the domestic market only with only a small percentage (15%) involved in international trade. Similarly, 88% (10,365 out of 11,760) of state-owned firms and 94% (26,628 out of 28,317) of collectively-owned firms are non-traders. In contrast, for foreign-owned firms, 10% export only, 15% import only and 45% import and export. Among the non-traders, stated-owned firms have the highest average output per firm, followed by foreign-owned firms, HMT firms, private firms and finally collectively-owned firms. In general, in terms of output, firms that trade tend to be larger than non-traders. State-owned importers have the largest average output followed by state-owned two-way traders and then private and foreign two-way traders.

In Table 4.3 we compare the characteristics of firms in each trade group. The main characteristics include firm size (in terms of employment and total sales), productivity (labour productivity and TFP estimated by De Loecker (2007) and Levinsohn and Petrin (2003)), capital intensity and average wages. Three observations stand out. First, firms that participate in international trade tend to be larger, more productive, more capital intensive and pay higher wages than non-traders. Second, among trading firms, two-way traders also tend to be larger, more productive, capital intensive and pay higher average wages than one-way traders. Third, differentiating between importers and exporters, we find that importer only firms outperform exporter only firms. These findings are in line with Castellani et al. (2010) and Vogel and Wagner (2010) although Castellani et al. (2010) finds that importer only firms are the most capital intensive of the four groups while we find importer only firms come in a close second to the two-way traders but are still more capital intensive than exporters-only and non-traders. p-values of t-tests for the equality of means of each characteristic between each pair of ownership types are also reported in the bottom half of the table. All differences except one (capital intensity between non-traders and only-exporters) are significant at 1% level.

Table 4.2: Trading status and output of Chinese manufacturing firms by ownership 2002-2006

	Trade status									
Ownership	Non-Traders		Exporters-Only		Importers-Only		Two-way traders		Total	
	No. of observations	Avg. output	No. of observations	Avg. output	No. of observations	Avg. output	No. of observations	Avg. output	No. of observations	
SOE	10,365	191.77	447	238.46	259	866.12	689	724.52	11,760	2
	(6.89)	(18.08)	(2.75)	(7.81)	(1.96)	(9.82)	(1.81)	(4.67)	(5.39)	
COLLECTIVE	26,628	43.16	940	134.98	190	184.48	559	222.46	28,317	1
	(17.60)	(10.45)	(5.77)	(9.30)	(1.44)	(1.53)	(1.47)	(1.16)	(12.98)	
PRIVATE	93,662	66.5	8,257	94.75	2,304	296.43	6,065	353.83	110,288	9
	(51.54)	(56.66)	(50.71)	(57.34)	(17.42)	(29.88)	(15.92)	(20.10)	(50.57)	
FOREIGN	7,756	87.76	3163	52.28	5,156	137.98	17,416	334.58	33,491	7
	(5.15)	(6.19)	(19.43)	(12.12)	(38.99)	(31.13)	(45.72)	(54.57)	(15.36)	
HMT	12,078	78.42	3,475	52.75	5,316	118.82	13,365	155.81	34,234	3
	(8.03)	(8.62)	(21.34)	(13.43)	(40.20)	(27.64)	(35.08)	(17.95)	(15.70)	
Total	150,489	10,993,307	16,282	1,364,491	13,225	2,285,423	38,094	10,678,974	218,090	2
	(69.00)	(43.41)	(7.47)	(5.39)	(6.06)	(9.03)	(17.47)	(42.17)	(100.00)	

Notes: Reported are the numbers of firms and average output of each category of trading status for each ownership group with percentages of total in parentheses. Non-traders are firms that neither export nor import while two-way traders are those that both export and import. Only-exporters refer to firms that export but do not import and only-importers refer to firms that import but do not export. Average output is measured in millions of Yuan and deflated in 2000 prices.

Table 4.3: Trade status and firm characteristics for Chinese manufacturing 2002-2006

Trade status	Firm characteristics							
	No. of obs (% of total)	Employment	Total sales	Labour productivity	TFP_DL	TFP_LP	Capital intensity	Average wages
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Non-Traders (G1)	150,489	4.877	10.044	3.817	6.294	6.350	3.530	2.274
	(69.00)	(1.03)	(1.18)	(0.99)	(1.02)	(1.04)	(1.21)	(0.55)
Only-exporters (G2)	16,282	5.250	10.462	3.787	6.504	6.555	3.542	2.374
	(7.47)	(1.03)	(1.12)	(0.92)	(0.96)	(0.98)	(1.16)	(0.50)
Only-importers (G3)	13,225	5.548	10.840	3.961	6.779	6.807	3.825	2.648
	(6.06)	(1.15)	(1.33)	(1.20)	(1.12)	(1.15)	(1.50)	(0.63)
Two-way traders (G4)	38,094	5.711	11.108	4.005	6.929	6.967	3.964	2.678
	(17.47)	(1.17)	(1.39)	(1.12)	(1.16)	(1.18)	(1.29)	(0.61)
All firms	218,090	5.091	10.310	3.857	6.450	6.500	3.625	2.375
	(100.00)	(1.12)	(1.30)	(1.02)	(1.08)	(1.09)	(1.25)	(0.59)
t-tests for the equality of means for different groups								
G1=G2 (*p*-value)		0.000	0.000	0.0002	0.000	0.000	0.2497	0.000
G1=G3 (*p*-value)		0.000	0.000	0.000	0.000	0.000	0.000	0.000
G1=G4 (*p*-value)		0.000	0.000	0.000	0.000	0.000	0.000	0.000
G2=G3 (*p*-value)		0.000	0.000	0.000	0.000	0.000	0.000	0.000
G2=G4 (*p*-value)		0.000	0.000	0.000	0.000	0.000	0.000	0.000
G3=G4 (*p*-value)		0.000	0.000	0.0002	0.000	0.000	0.000	0.000

Notes: Column (1) gives the numbers of observations of each trade group and their shares to the total respectively. Columns (2)-(8) provide the means and standard deviations (in parentheses) of corresponding firm characteristics for each group of firms. All indicators of firm characteristics are in logs. Labour productivity refers to value-added per employee, TFP_DL and TFP_LP are total factor productivity estimated by De Loecker (2007) and Levinsohn and Petrin (2003) respectively. Capital intensity is capital per employee and average wage is wage per employee. The bottom panel of the table reports the *p*-values of *t*-tests for the equality of means for each pair of the groups.

4.4 Empirical Results

The descriptive evidence suggests that productivity for Chinese manufacturing firms differs by the type of participation in international markets. However, the existence of productivity differentials may be due to other factors related to firm productivity. Hence, the next step in our empirical investigation is to estimate the extent of any productivity premia controlling for the productivity related factors. The productivity premia for traders are estimated by regressing *TFP* on a set of dummy variables including trade status (non-traders are omitted as the reference group). The estimating equation is given by:

$$TFP_{it} = \alpha + \beta_1 EXPonly_{it} + \beta_2 IMPonly_{it} + \beta_3 EXP/IMP_{it} + \gamma X_{it} + \varepsilon_{it} \tag{4.9}$$

where *i* and *t* represent firm and year respectively, *EXPonly* (*IMPonly*) is a dummy variable that equals 1 if a firm exports (imports) but does not import (export), and 0 otherwise; *EXP/IMP* is a dummy for a firm that both exports and imports and 0 otherwise; *X* is a vector of control variables (including firm size, ownership, firm age, employees' average wages, industry, region and year dummies). Finally, ε is the error term.

Table 4.4 presents the productivity premia for Chinese firms between 2002 and 2006 for TFP measured by both our De Loecker method and the more traditional Levinsohn and Petrin (2003) method. The coefficients on the three trade status dummies are positive and highly significant (at the 1% level) and show that compared to non-traders, firms that trade enjoy higher productivity levels controlling for a set of firm characteristics. However, the degree of the productivity premia differs across the three groups. Two-way traders enjoy the largest productivity premia, followed by firms that only import. Firms engaged in exporting only have the smallest premia. Firms that both import and export are 24.86 percent (100[exp(0.222)-1]) more productive than non-traders. The productivity advantage of firms that import only over non-traders is 16.88 percent (100[exp(0.156)-1]), double that of exporters only who have a premium over non-traders of 8.11 percent (100[exp(0.078)-1]). Our results are consistent with the descriptive evidence in Table 4.3 and are in line with findings from Muuls and Pisu (2009) for Belgian firms, Vogel and Wagner (2010) for German manufacturing firms, Castellani et al. (2010) for the Italian manufacturing industry and Kasahara and Lapham (2013) for Chilean manufacturing plants. Importing clearly plays an important role in the generation of productivity premia for firms engaged in international trade and the productivity premia of exporters over non-exporters is not as large as has been claimed in the previous literature where importing has not been specifically accounted for in the analysis.

4.4.1 Random-effects Probit Results

We apply a Random-effects Probit model to test the self-selection of importing for Chinese manufacturing firms. The probability of starting to import is regressed on a set of firm characteristics and the results are presented in Table 4.5 with Column (1) the coefficients and Column (2) the average marginal effects (AME).[1] Past export experience is found to have a highly significant effect (1% level) on firms' decision to enter to imports market. Firm that exported in

[1]The average marginal effects of continuous variables are computed by $AME = \frac{1}{n}\sum_{i=1}^{n}\phi(X_i^{'}\beta)\beta$ and dummy variables by $AME = \frac{1}{n}\sum_{i=1}^{n}[\Phi(X_i^{'}\beta|X_j = 1) - \Phi(X_i^{'}\beta|X_j = 0)]$, where $\Phi(.)$ is the standard normal cumulative distribution function and $\phi(.)$ is the standard normal density function.

Table 4.4: Productivity premia for Chinese firms 2002-2006

Dependent variable: TFP*it*		
	TFP_De Loecker	TFP_LP
EXPonly	0.078***	0.084***
	(0.008)	(0.008)
IMPonly	0.156***	0.158***
	(0.009)	(0.009)
EXP/IMP	0.222***	0.224***
	(0.007)	(0.007)
SMALL	-0.474***	-0.482***
	(0.006)	(0.006)
LARGE	0.346***	0.367***
	(0.012)	(0.012)
FOREIGN	0.438***	0.426***
	(0.015)	(0.016)
HMT	0.370***	0.359***
	(0.015)	(0.015)
COLLECTIVE	0.299***	0.290***
	(0.014)	(0.014)
PRIVATE	0.343***	0.335***
	(0.013)	(0.014)
age	0.069***	0.073***
	(0.005)	(0.005)
wage	0.283***	0.281***
	(0.003)	(0.003)
Constant	5.806***	5.813***
	(0.025)	(0.025)

Notes: Standard errors in parentheses. Region, indusrty and year dummies included. *** indicates significance levels at 0.01.

previous year are found to have a 3% higher likelihood to start importing in the current period. Age is negatively correlated with import entry, indicating the importance of firm experience. TFP, labour force skills (proxied by average wages of employees) and firm size are important determinants of entry to imports market. Larger and more productive firms and those with more skilled labour are more likely to start to import. Ownership is also am important factor. Compared to SOEs (the reference group), firms with foreign ownership (including HMT) and private firm are more likely to start to import. The main variable of our interest is TFP. After controlling for other firm characteristics, firms that start to import are more productive prior to the decision. As time-varying independent variables are lagged one year, these results show a causal effect of importing and TFP of Chinese manufacturing firms: more productive firms self-select to import.

4.4.2 Propensity Score Matching with Difference-in-differences Results

As described in Section 2, matching is performed after the procedures that ensure the propensity score is balance and covariates are balance within blocks of propensity score across treatment and control groups. Three matching estimators are applied and after matching we test the reliability of the matching using several methods. See Table 4A.2 and 4A.3 for tests of assessing the quality of matching. Statistics from standardized differences, *t*-tests, mean and median biases, various between unmatched and matched sample show The quality of our PSM are satisfied. After assuring the matching to be of good quality we can now make a comparison between the treated and control groups. Our PSM-DID results are presented in Table 4.6. The PSM-DID results provide the causal effect of importing on TFP and the ATTs can be interpreted as per-

Table 4.5: Self-selection to importing for Chinese manufacturing firms

Independent variable: STARTit		
	(1)	(2)
	xtprobit	AME
EXP	0.432***	0.026***
	(0.015)	(0.001)
age	-0.035***	-0.002***
	(0.011)	(0.001)
wage	0.040***	0.002***
	(0.012)	(0.001)
TFP	0.044***	0.003***
	(0.008)	(0.000)
MEDIUM	0.168***	0.010***
	(0.018)	(0.001)
LARGE	0.211***	0.013***
	(0.049)	(0.003)
FOREIGN	0.255***	0.015***
	(0.043)	(0.003)
HMT	0.369***	0.022***
	(0.042)	(0.003)
COLLECTIVE	-0.139***	-0.008***
	(0.047)	(0.003)
PRIVATE	0.170***	0.010***
	(0.041)	(0.002)
EAST	0.110***	0.007***
	(0.026)	(0.002)
WEST	-0.026	-0.002
	(0.040)	(0.002)
Constant	-2.876***	
	(0.082)	
Observations	171,428	171,428
Number of firms	42,857	42,857
log likelihood	-20,090	

Notes: Standard errors in parentheses. Industry and year dummies included. All time varying variables are lagged one year. *** indicates significance levels at 0.01.

centage changes in TFP. Guassian Kernel matching with bandwidth of 0.01 is firstly performed in panel (a). The ATTs for the entry year and up to two years after starting to import are all positive and statistically significant (at the 1% level), suggesting a positive effect of importing on productivity.[1] Our results suggest that import market entrants have 10.9% higher TFP growth than matched non-importers in the year of entry. New importers also appear to experience steady and increasing productivity growth for the first two years after entry. In the year after entry, new importers have a 13.5% higher TFP growth which increases to 17.1% two years after. These results suggest a strong learning by importing effect for Chinese firms.

Results form two other matching estimators are also presented in Table 4.6 with one-to-one matching in (b) and nearest neighbours matching with the number of neighbours of 5 in (c). Results are similar to those obtained from Kernel matching. Since Kernel mathing uses all the observations within the commom support and thus maximizes precision, we use Kernel estimator for the rest of the matching estimation.

We now investigate the impact of ownership structure on the learning by importing results. Compared to domestic firms, a much higher proportion of foreign- and HMT-owned firms import. Table 5.3 also indicates that firms with foreign capital and private firms are more likely to import than state-owned firms. The results are presented in Table 4.7. Except for collective firms, all other four types of firms show the learning effect of importing on TFP. However, the extent of the learning effect varies among different ownership categories. For foreign-owned firms, import starters have a 11.1% higher TFP growth than the group of matched non-importing firms in the year of entry. The TFP growth rate for these new entrants drops to 10.7% and then

[1] We do not look at the ATT for the pre-entry year as the good quality of matching indicates no significant difference between the matched import starters and non-importers before the treatment. We are not able to look at the effects on TFP of importing three or more years after the entry due to the relatively short period of our sample.

Table 4.6: Importing and productivity for Chinese manufacturing firms: PSM-DID estimates

	s=0	s=1	s=2
Outcome variable: year-to-year productivity growth rate			
a) Guassian Kernel matching			
ATT	0.109***	0.135***	0.171***
	(0.013)	(0.018)	(0.024)
N (control)	93,523	62,174	30,808
N (treated)	3,034	2,265	1,429
b) One-to-one matching			
ATT	0.129***	0.167***	0.176***
	(0.019)	(0.024)	(0.033)
N (control)	93,523	62,174	30,808
N (treated)	3,034	2,265	1,429
c) Nearest neighbours matching			
ATT	0.118***	0.138***	0.166***
	(0.014)	(0.019)	(0.025)
N (control)	93,523	62,174	30,808
N (treated)	3,034	2,265	1,429

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

increases to 19.1% two years after entry. A similar but bigger learning effect is found for HMT new importers with the TFP growth rate for 16.2%, 18.4% and 21.1% for entry year, one year and two years after the entry.

Private new importers also show a steady learning effect from importing. These new entrants have a 10.4% higher TFP growth compared to matched non-importers and their TFP continues to grow at a higher speed after the entry with 13.8% one year after entry and 18.1% two years after. Interesting finding is observed for SOE new importers. Only a small number of SOEs start to import, and at the year of entry to imports market, these new import entrants do not experience a significant TFP growth. However, after the entry these SOE new entrants enjoy a big increase of their TFP growth with 26.4% and 43.7% for first year and second year after the entry. For collective firms, new importers have a rate of TFP growth that is 4.7% higher than that of matched non-importers in the year of entry. The new entrants have smaller TFP growth one and two years after the entry compared to the matched control group. However, the ATTs for the collective firms are never significant at usual level.

Table 4.7: PSM-DID estimates by ownership

	s=0	s=1	s=2
a) FOREIGN			
ATT	0.111***	0.107***	0.191***
	(0.034)	(0.041)	(0.051)
N (control)	5,753	3,908	1,978
N (treated)	728	596	418
b) HMT			
ATT	0.162***	0.184***	0.211***
	(0.030)	(0.037)	(0.050)
N (control)	9,960	6,659	3,321
N (treated)	801	630	389
c) PRIVATE			
ATT	0.104***	0.138***	0.181***
	(0.018)	(0.027)	(0.038)
N (control)	57,574	37,847	18,675
N (treated)	1,327	918	539
d) SOE			
ATT	0.119	0.264**	0.437**
	(0.092)	(0.027)	(0.175)
N (control)	3,270	1,914	1,346
N (treated)	78	52	44
e) COLLECTIVE			
ATT	0.047	-0.054	-0.053
	(0.077)	(0.111)	(0.163)
N (control)	9,970	7,218	3,226
N (treated)	97	68	38

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

Our results indicate that new importers are more productive and that productivity keeps increasing when they remain importers compared to the matched control group. Across the different ownership groups, state-owned new importers display the strongest learning effects from importing, followed by private firms, HMT and foreign-owned firms. Such a phenomenon can be explained as follows. Foreign-owned firms in China are on average more productive and hire more skilled employees and most importantly, have access to a larger number of varieties in terms of inputs and advanced technologies through their parent companies. The existence of an overseas network means they do not pay the sunk entry that the domestic firms have to incur when they decide to import. In contrast, for domestic firms, only the most productive of those with strong state support in the case of SOEs can afford the high entry costs. Firms that start to import and stay active in the import market are those which can use imported advanced technologies or learn from their trading partners in order to compete. During this period, those with the most to gain were the traditional SOEs where the technology gap would have been the largest. The reason why SOE new importers experience the larger TFP growth than foreign-owned and private firms may also be due to the increased competition from the exposure to international markets. It has been shown theoretically and empirically that the firm productivity is closely linked to the level of competition in the international market (Bloch and McDonald, 2001; Galdon-Sanchez and Schmitz, 2002; Levinsohn, 1993; MacDonald, 1994). Thanks to the substantial level of support from the government, the SOEs are not exposed to the international competition as much as foreign and private firms. However, when these SOEs enter the import market, they are exposed to international competition which is new to them and such competition drives them to improve their productivity rapidly. It is not surprising that we find these SOE new importers have the highest TFP gains.

We now investigate whether the learning effect differs by firms size. The results are presented in Table 4.8. The ATTs for medium and small firms are all positive and significant at 1% level, suggesting a strong and positive effect of importing on productivity for medium and small sized new importers. A 10.2% higher TFP growth is observed for small new importers at the entry and 10.7 % the year following the entry and 13% two years after. The ATTs for medium firms are 7.5%, 16.7 % and 17.8 % at entry year, one year and two years after the entry.

However, the ATTs for large firms are not significant at usual levels in the year of entry and up to two years after entry. The reason may be that compared to medium and small firms, large-sized firms usually have the ability to invest in new products and more advanced technology and improve their productivity. So importing may not be the only or most important way for them to survive and improve. Importing is thus not found to have a significant effect on these large entrants compared to the matched control firms.

Table 4.8: PSM-DID estimates by size

	s=0	s=1	s=2
a) SMALL			
ATT	0.102***	0.107***	0.130***
	(0.017)	(0.022)	(0.030)
N (control)	81,658	54,493	27,102
N (treated)	2,013	1,499	915
b) MEDIUM			
ATT	0.075***	0.167***	0.178***
	(0.024)	(0.033)	(0.045)
N (control)	11,234	7,285	3,523
N (treated)	923	688	453
c) LARGE			
ATT	0.001	0.107	0.089
	(0.092)	(0.120)	(0.178)
N (control)	442	316	176
N (treated)	92	72	54

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

We move on to investigate whether the origin of imports has an impact on productivity. If Chinese firms import most of their inputs from high income economies one might expect that importers will show higher productivity gains as they are able to learn from exposure to sellers in high income markets. However, these inputs are likely to be more expensive. We follow the classification by the World Bank and categorise the new importers into two groups based on the origin of their imports.[1] One group consists of those firms that import from high in-

[1] For a list of countries classified as high income see http://data.worldbank.org/about/country-classifications/country-and-lending-groups#High_income

come economies (HI) and the other group with firms that import only from non-high income economies (non-HI). During this period 86% of new importers imported from high income economies and 14% from non-high income economies. The PSM-DID results are presented in Table 4.9. Firms that start importing experience productivity gains at the entry and after regardless from the origins of their imports. However, firms that start to import from high income economies display stronger learning effects compared to those starting to import from non-high income economies.

Table 4.9: PSM-DID estimates by origins of imports

	s=0	s=1	s=2
a) High-income economies			
ATT	0.120***	0.141***	0.176***
	(0.016)	(0.022)	(0.030)
N (control)	93,568	62,222	30,856
N (treated)	2,000	1,501	984
b) Non-high-income economies			
ATT	0.082***	0.103***	0.117***
	(0.024)	(0.030)	(0.040)
N (control)	91,582	61,522	30,563
N (treated)	859	630	370

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

Finally, we test whether the effect of importing on TFP growth depends on the skill and technology content of inputs imported. We classify the firms into two groups based on the skill and technology intensity of their imported goods.[1] The PSM-DID estimates are presented in Table 4.10. ATTs for both groups are found positive and highly significant for all three periods examined (all at 1% significance level, except 10% level for firms importing low skill and technology intensive products at one year after the entry.) However, the magnitude of the

[1]The classification of HS 6-digit products for skill and technology intensity is available from United Nations Conference on Trade and Development (UNCTAD) Database. See http://www.unctad.info/en/Trade-Analysis-Branch/Data-And-Statistics/Other-Databases/

effect of learning from importing for the two groups are different. For firms that import high and medium skill and technology intensive products, the TFP growth is 11.9% higher for the import starters at the year of their entry compared to the control firms. These new importers maintain a steady TFP growth after the entry with 15.5% and 17% higher TFP growth rate for the first and second year compared to the matched non-importers. For firms that import only low skill and technology intensive products, new importers have a 8.1% higher TFP growth than the control group at the year of the entry to imports market. One year after their entry, these new importers still have a 5.1% higher TFP growth than the matched non-importers, but only one-third of the growth rate of the new importers who source high and medium skill and technology intensive products. Two years after the entry, new entrants importing only low skill and technology intensive products have a 12.3% higher TFP growth than the control group, about 5% lower than firms that import high and medium skill and technology intensive products. Though all new importers experience productivity growth in the year of their entry to imports market, firms that import more skill and technology intensity products are found to learn more from importing.

Table 4.10: PSM-DID estimates by technology intensity of imports

	s=0	s=1	s=2
a) High and medium technology intensive products			
ATT	0.119***	0.155***	0.170***
	(0.017)	(0.022)	(0.030)
N (control)	92,449	62,152	30,826
N (treated)	1,878	1,425	936
b) Low technology intensive products			
ATT	0.081***	0.051*	0.123***
	(0.023)	(0.031)	(0.039)
N (control)	90,948	60,360	29,431
N (treated)	919	661	392

Notes: Standard errors in parentheses. *** and * indicate significance levels at 0.01 and 0.1.

4.5 Robustness Checks

We present the results of several sensitivity checks in this section. Firstly, we test the treatment effect of importing on productivity using an alternative TFP measure by Levinshon and Petrin (2003).[1] The results are presented in Table 4.11 and the ATTs are

Table 4.11: PSM-DID estimates with TFP measure by Levinshon and Petrin (2003)

	s=0	s=1	s=2
ATT	0.111***	0.138***	0.177***
	(0.014)	(0.018)	(0.025)
N (control)	93,522	62,173	30,807
N (treated)	3,034	2,265	1,429

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

Secondly, following Silverman (1998) and Heckman et al. (1997), we try an alternative bandwidth of 0.06 in the PSM-DID estimation and the results are presented in Table 4.12. We compare the results with those previously obtained using a bandwidth of 0.01 (Panel (a) in Table 4.6) and find that the ATTs are slightly higher with bandwidth 0.01, but the standard errors are very similar for both choices of bandwidth.

Table 4.12: Gaussian Kernel matching with bandwidth 0.06

	s=0	s=1	s=2
ATT	0.085***	0.099***	0.125***
	(0.013)	(0.017)	(0.022)
N (control)	93,523	62,174	30,808
N (treated)	3,034	2,265	1,429

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

Thirdly, we examine whether the impact of importing on productivity varies for firms with different levels of TFP. The results are presented in Table 4.13. Panels a) and b) examine firms

[1] The correlation between TFP estimated by De Loecker (2007) and that by Levinsohn and Petrin (2003) is 0.994.

with TFP at the sample mean or above in the initial year and those below. New importers from both groups learn from participation in imports market as ATTs are positive and significant (at 1% level). New importers of firms whose TFP are below the sample mean have 10.7% higher TFP growth than the matched non-importer at the year of entry and keep a high TFP growth after the entry (14.7% and 19% higher for their first and second year of import participation). New importers whose TFP is at sample mean or above show a 12.2% higher TFP growth at the year they start importing and their TFP growth rate is 13.3% and 15.9% higher than the control group. So although firms that have a TFP below the mean initially have a slight smaller productivity growth rate than those who have mean or above TFP at the entry year, these new importers display a stronger learning effect from importing afterwards as their TFP growth rates are higher at first and second year after the entry. Panels c) and d) compare firms of the first and fourth quartiles of TFP of the sample in the initial year and those below. Similar findings are observed as those from Panels a) and b).

Also, we test the impact of variety of imported inputs on the learning from importing effect. Results are presented in Table 4.14. New importers are found to display higher TFP growth at entry year and two years afterwards regardless the number of varieties imported. The new entrants who import only one variety show a 6.5% higher TFP growth at the entry year and 8.6% and 9.3% for first and second year after entry than the control group. However, starters that import two or more varieties have much higher TFP growth rates since their entry with 13.1%, 13.4% and 16.1% higher than the matched non-importers at entry year, first and second year after the entry.

Furthermore, as import entrants may exit from the imports market at some point of the

Table 4.13: PSM-DID estimates by TFP level

	s=0	s=1	s=2
a) Below mean TFP			
ATT	0.107***	0.147***	0.190***
	(0.021)	(0.029)	(0.038)
N (control)	55,910	37,168	18,442
N (treated)	1,276	918	543
b) Mean and above TFP			
ATT	0.122***	0.133***	0.159***
	(0.018)	(0.023)	(0.032)
N (control)	37,303	24,783	12,226
N (treated)	1,758	1,347	885
c) 1st quartile of TFP			
ATT	0.106***	0.131***	0.185***
	(0.035)	(0.050)	(0.065)
N (control)	26,650	17,582	8,655
N (treated)	27,174	371	218
d) 4th quartile of TFP			
ATT	0.135***	0.119***	0.140***
	(0.026)	(0.033)	(0.042)
N (control)	17,200	11,395	5,670
N (treated)	1,008	789	538

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

Table 4.14: PSM-DID estimates by the variety of imported inputs

	s=0	s=1	s=2
a) one variety			
ATT	0.065***	0.086***	0.093***
	(0.018)	(0.024)	(0.035)
N (control)	93,674	62,485	31,033
N (treated)	1,448	1,023	570
b) two varieties or more			
ATT	0.131***	0.134***	0.161***
	(0.020)	(0.026)	(0.033)
N (control)	93,768	62,337	30,960
N (treated)	1,411	1,018	784

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

period, we test the learning effect excluding the exiters and also test whether these import exiters ever learn from importing. Results for excluding exiters are presented in Table 4.15 and exiters only in Table 4.16. For firms that start importing during our sample and continue being active in imports market, they have a 11% higher TFP growth rate than the matched non-importers at the year they start importing. These firms continue to show a big increase of TFP growth after the entry with 19.7% and 24.8% for first and second year after their entry. Compared with the results obtained for the whole sample which include firms that start importing and stop during the sample (see results from Table 5.6 in the previous section), the new importers that continue importing after the entry (surviving import starters) experience higher TFP growth since they start importing and up to two years after their entry.

When we look at firms that start importing and stop during the sample, the new importers still have about 10% higher TFP growth than the matched non-importers since their entry to the imports market. Compared with the results for the whole sample (Table 5.6) and sample with surviving import starters only (Table 5.13), these import starter-exiters show the smallest TFP growth rates compared to the control group for the three years since they start importing.

Table 4.15: PSM-DID estimates excluding import exiters

	s=0	s=1	s=2
ATT	0.110***	0.197***	0.248***
	(0.018)	(0.028)	(0.037)
N (control)	92,772	61,420	30,054
N (treated)	1,668	899	514

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

Finally, as the decision to import is likely to relate to the decision to export and thus all or part of the productivity growth of the firms may be due to the learning from exporting rather than from importing. We perform our PSM-DID for importers only by dropping observations

Table 4.16: PSM-DID estimates for import starter-exiters

	s=0	s=1	s=2
ATT	0.105***	0.008***	0.104***
	(0.019)	(0.022)	(0.029)
N (control)	61,960	61,960	30,804
N (treated)	1,366	1,366	915

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

that export during the sample period. Results are presented in Table 4.17. Focusing on the sample without exporters, the new importers are still found to have a 9% higher TFP growth than the matched non-importers at the year of entry and 10.8% and 13.1% higher for the first and second year after the entry. Leaning-from-importing effects are confirmed.

Table 4.17: PSM-DID estimates for non-exporters

	s=0	s=1	s=2
ATT	0.009***	0.108***	0.131***
	(0.031)	(0.042)	(0.051)
N (control)	77,537	51,168	25,703
N (treated)	622	414	253

Notes: Standard errors in parentheses and *** indicates significance levels at 0.01.

4.6 Conclusions

In this chapter we have examined various aspects of the relationship between firm performance and international trade concentrating on the little explored relationship between importing and productivity. This chapter presents an empirical analysis of the causal effects between productivity and importing using propensity score matching method and matching difference-in-differences approaches. Using Chinese manufacturing firm-level data we observe bi-directional causality between importing and productivity. Generally speaking, more productive firms self-select into the imports market and after import-market entry on average firms experience productivity gains. However, these gains are not evenly distributed.

Our results show that compared to the matched non-importing firms, except for collective firms, all new import starters of other ownership types demonstrate a learning effect. The strongest learning effect is generally exhibited by the SOE importers, followed by Hong Kong, Macao and Taiwan (HMT)-owned importers, private importers and foreign-owned importers. We also find that small and medium sized import starter experience higher productivity growth than the control group and in comparison to the large new importers. Furthermore, we find that import starters that source their inputs from high income economies have larger productivity gains than those that start to import only from non-high income economies. Furthermore, we find that new importers who import from high income countries experience greater productivity gains than those that import from lower income countries. Finally new importers who import medium and high skill and technology intensive products, have lower TFP level initially and importing more varieties of inputs display stronger learning effects.

Our finding of strong supporting evidence of learning-by-importing for Chinese manufacturers, in contrast to the results of Vogel and Wagner (2010) for German manufacturers and Augier et al. (2013) for Spanish firms, may be because developed countries have long been exposed to foreign competition and had access to global import markets. However, in China's case, joining the WTO in 2001 marked a step change in the international opportunities available to Chinese firms. At this time, firms from China were more likely to be some distance from the technological frontier meaning Chinese firms would have been exposed to considerable learning opportunities from the use of superior inputs because of existing gaps in technology and product quality between them and potential new trading partners.

Our results have potentially important policy implications. First, because Chinese manufacturing firms appear to benefit from importing, China could derive additional productivity gains from further trade liberalization and might want to consider promoting trade liberalisation alongside continued export promotion. Recent senior trade delegations from China to the developed countries and the subsequent highly publicised deals make it clear that China is interested in importing high quality intermediate inputs and attracting further foreign direct investment. Our result that indigenous firms and small- and medium-sized firms exhibit high productivity growth as they learn from importing means that government support to help these firms break into the imports markets and overcome potentially large sunk costs could be beneficial. Examples include lowering barriers to importing by providing more information on sources of intermediate inputs, lowering tariffs and arranging for such firms to attend overseas trade fairs and similar events and possible support mechanisms that policy makers could employ.

Appendix 4A

Table 4A.1: Definition of variables

Variable	Definition
EXP	a binary variable which equals 1 if a firm had positive exports and 0 otherwise
EXPonly	a binary variable which equals 1 if a firm had positive exports and no positive imports and 0 otherwise
IMP	a binary variable which equals 1 if a firm had positive imports and 0 otherwise
IMPonly	a binary variable which equals 1 if a firm had positive imports and no positive exports and 0 otherwise
EXP/IMP	a binary variable which equals 1 if a firm had both positive exports and positive imports and 0 otherwise
START	a binary variable which equals 1 if a firm starts importing and 0 otherwise
age	log of a firm's age: the report year minus the founded year of a firm
wage	log of average wage of employees of a firm (ratio of total wage bill to the number of employees)
employment	log of number of employees
TFP	total factor productivity of a firm estimated by the method of De Leocker(2010)
TFP_LP	total factor productivity of a firm obtained from estimation of the semi-parametric approach of Levinsohn and Petrin (2003)
pregrowth	difference of TFP between the year when a firm starts importing and one year before the entry
SOE	a dummy which equals 1 if a firm is state-owned and 0 otherwise
COLLECTIVE	a dummy which equals 1 if a firm is collectively-owned and 0 otherwise
PRIVATE	a dummy which equals 1 if a firm is private-owned and 0 otherwise
FOREIGN	a dummy which equals 1 if a firm with over 25% share of capital from foreign investors and 0 otherwise
HMT	a dummy which equals 1 if a firm with over 25% share of capital from Hong Kong, Taiwan or Macao investors and 0 otherwise
SMALL	a size dummy which equals 1 if a firm is a small firm and 0 otherwise.
MEDIUM	a size dummy which equals 1 if a firm is a medium firm and 0 otherwise.
LARGE	a size dummy which equals 1 if a firm is a large firm and 0 otherwise.
EAST	a region dummy which equals 1 if a firm is located in the East of China and 0 otherwise.
CENTRAL	a region dummy which equals 1 if a firm is located in Central area of China and 0 otherwise.
WEST	a region dummy which equals 1 if a firm is located in the West of China and 0 otherwise.

Table 4A.2: Balancing test from Guassian Kernel matching (I)

Variables	Sample	Mean		%bias	% bias reduction	t-test	
		Treated	Control			t-stat	p>ltl
EXP	Unmatched	0.54348	0.10300	106.7		95.42	0.000
	Matched	0.51945	0.49818	5.2	95.2	1.66	0.098
age	Unmatched	2.23260	2.26410	-4.7		-2.93	0.003
	Matched	2.30490	2.30740	-0.4	92.2	-0.17	0.867
wage	Unmatched	2.51520	2.28360	42		28.53	0.000
	Matched	2.52760	2.50580	4	90.6	1.58	0.115
TFP	Unmatched	6.72950	6.31000	41.5		28.79	0.000
	Matched	6.79170	6.74860	4.3	89.7	1.64	0.100
pregrowth	Unmatched	0.14409	0.10481	5.7		3.08	0.002
	Matched	0.14409	0.13141	1.8	67.7	0.74	0.460
MEDIUM	Unmatched	0.26651	0.10813	41.5		33.99	0.000
	Matched	0.27554	0.26493	2.8	93.3	0.93	0.352
LARGE	Unmatched	0.02529	0.00637	15.2		15.36	0.000
	Matched	0.02999	0.02790	1.7	88.9	0.49	0.626
FOREIGN	Unmatched	0.25857	0.05871	56.9		55.27	0.000
	Matched	0.22808	0.20636	6.2	89.1	2.05	0.040
HMT	Unmatched	0.28365	0.11151	44.3		36.43	0.000
	Matched	0.27324	0.27876	-1.4	96.8	-0.48	0.630
COLLECTIVE	Unmatched	0.04202	0.17560	-43.9		-24.16	0.000
	Matched	0.04120	0.04746	-2.1	95.3	-1.19	0.236
PRIVATE	Unmatched	0.38921	0.58947	-40.9		-27.64	0.000
	Matched	0.43045	0.43725	-1.4	96.6	-0.53	0.594

Notes: Reported are the means of variables for treated and control firms for unmatched and matched sample together with the bias and standardized differences (% bias reduction) and t-tests in matched sample compared to those in unmatched sample.

Table 4A.3: Balancing test from Guassian Kernel matching (II)

Sample	Psuedo R2	LR chi2	p>chi2	Mean Bias	Median Bias	B	R	%Var
Unmatched	0.204	5518.09	0.000	17.5	8.5	150.1*	1.12	50
Matched	0.002	14.88	1.000	0.7	0	9.9	0.91	0

Notes: Rubins' B is the absolute standardized difference of the means of the linear index of the propensity score in the treated and (matched) non-treated group and Rubin's R is the ratio of treated to (matched) non-treated variances of the propensity score index. Rubin (2001) recommends that B be less than 25 and that R be between 0.5 and 2 for the samples to be considered sufficiently balanced. An asterisk is displayed next to B and R values that fall outside those limits.

Chapter 5

The Links Between Imported Inputs and Exports

The literature on firm heterogeneity and international trade has documented that a large number of firms that export also engage in importing activities. However, questions such as why firms engage in both importing and exporting and how and to what extent imported inputs contribute to firms' export performance have not been widely investigated. This chapter examines the link between imported inputs and firms' export performance, in particular, trying to answer two related questions: first, whether an increased use of foreign inputs generates more exports and second, whether these inputs results in an increase in the quality of exports. Using a large dataset with rich information on firm production and international trade activities of Chinese manufacturing firms for the period 2002 to 2006, I find that importing more inputs raises both the extensive and intensive margins of firms' exports, including the number of varieties, values and quality of exports. Results also show that differentiated imported inputs have a larger impact on firms' export performance than homogenous inputs. In addition, the effect is stronger for inputs imported from high income economies than those sourced from other countries. These results suggest that Chinese firms improve their export performance via variety complementarity, technology absorption and quality transformation from the use of foreign inputs. We address the potential endogeneity of the imported inputs using a general method of moments (GMM) estimator. Several robustness checks are performed and similar results are observed across different specifications.

5.1 Introduction

In an increasingly globalized world, the relationship between firm performance and the extent to which a firm is internationalized is subject to ever greater scrutiny from academics and policy makers. The internationalization of firms is especially important for developing and newly industrialising countries that continue to pursue an export led growth strategy and remain dependent to a large extent on exports for future growth and employment. A large number of studies have been done on different aspects of exports at firm level in recent years, ranging from the entry and exit into and out of export markets, export performance and firm heterogeneity, leaning by exporting to quality upgrading of exports. Examples include Roberts and Tybout (1997), Baldwin and Gu (2004), Bernard and Jensen (2004), Girma et al. (2004), Arnold and Hussinger (2005), De Loecker (2007), Bustos (2011), Manova and Zhang (2012), Berman et al. (2012) and Lawless and Whelan (2014).

Compared to the extensive research on exports, imports, the other side of international activities have received far less attention. The existing literature on firm imports includes Halpern et al. (2005), MacGarvie (2006), Kugler and Verhoogen (2009), Castellani et al. (2010), Blalock and Veloso (2007), Vogel and Wagner (2010), Kasahara and Lapham (2013), Augier et al. (2013), Chen and Ma (2012) and Cadot et al. (2014). Although there can be a negative impact of importing, such as fierce competition and job losses, evidence has been found that importing also contributes positively in several ways. First, firms can lower their production costs and increase product scope with access to new cheaper varieties of inputs via importing. Second, with more advanced or better quality imported inputs, firms improve productivity through technology transfer or product quality upgrading (Acharya and Keller, 2009; Amiti and Khandelwal, 2013; Goldberg et al., 2010; Kasahara and Rodrigue, 2008; Lo Turco and Maggioni, 2014).

There are a small number of papers that study both sides of international trade activities, such as Bernard et al. (2005), McCann (2009), Manova and Zhang (2009), Muuls and Pisu (2009), Aristei et al. (2013), Wang and Yu (2012), Seker (2012), Castellani et al. (2010), Kasahara and Lapham (2013) and Lo Turco and Maggioni (2014). Evidence has been documented that most of the firms that export also engage in importing activities. Empirical analyses have found that these two-way traders are larger, more productive, more capital- and skill-intensive than firms that only export or only import. However, the investigation into questions such as why firms engage in two-way trade, how and to what extent the use of imported inputs relates to firms' export performance has been fairly limited. The several existing studies that try to establish the links between firm-level imported inputs and export performance are Feng et al. (2012) for China, Lo Turco and Maggioni (2013) for Italy, Bas and Strauss-Kahn (2014) for France and Parra and Martinez-Zarzoso (2014) for Egypt.

This chapter examines the role of imported inputs on firms export performance using Chinese manufacturing firm production and international trade data for the period of 2002 to 2006. With rich information on firms' characteristics from survey data and both export and import activities from customs data, I test whether importing more ranges and better quality of inputs improves firms' export performance, on both extensive and intensive margins. We find positive and significant effects of imported inputs on export scope across different specifications and larger impacts are found for differentiated inputs and inputs imported from high-income economies.

This chapter contributes to the literature in the following ways. First, it is one of the first few papers that explore the links between firm-level imports and exports. Similar to Lo Turco and Maggioni (2013) and Bas and Strauss-Kahn (2014), who find a positive connection between imported inputs and exports for firms from developed countries, I provide evidence that such a link exists among firms from a developing country. Secondly, unlike the majority of studies on Chinese firm-level international trade focusing on the benefits of exports, I look at the benefits of imports, in particular on the export scope. It is a common perception that the rapid growth of China has been driven by its massive growth of exports over the last two decades. While Chinese exports rose from USD249.2 billion to USD1,577.75 billion between 2000 and 2010, imports increased from USD225 billion to USD1,396.24 billion during this period (China Statistical Yearbook 2013). The benefits of imports to the Chinese economy should not be ignored and the link between imports and exports is not yet clearly identified. Feng et al. (2012) also study the benefits of imported inputs on exports for Chinese firms. This chapter differs from theirs in that apart from the effects of an increase of imported inputs on the volume of firms' exports and the number of exported products, I also look at the quality of export varieties. We provide a more comprehensive understanding of the roles of imported inputs in enhancing exports.

The remainder of this chapter is organized as follows. Section 5.2 describes the data and the construction of our dataset. Section 5.3 presents our econometric analysis and empirical results for the relationship between imported inputs and firm export performance, in particular the number of varieties, value and quality of exprots. Section 5.4 addresses the potential endogeneity issues by GMM techniques for different specifications. Section 5.5 provides several robustness checks including alternative estimation method, alternative product quality proxies and large sample data. Finally Section 5.6 concludes.

5.2 Data and Descriptives

5.2.1 The Data

The data used in this chapter are drawn from two sources. The firm-level production data are from the Annual Survey of Industrial Enterprises provided by the National Bureau of Statistics of China (NBS) and the transaction-level trade data are provided by the Department of Customs Trade Statistics, the General Administration of Customs of China. We use data from both sources for the years 2002 to 2006.

The NBS data cover all state-owned industrial enterprises and non-state-owned industrial enterprises with annual sales of greater than 5 million Chinese Yuan (RMB).[1] According to the

[1] The official USD and RMB exchange rate between 2002 and 2004 was 8.277, 8.194 in 2005 and 7.973 in 2006 (World Development Indicators, World Bank). Hence, the threshold for inclusion in the dataset is equivalent to between USD600,000 and USD627,000.

NBS industry classification, industrial enterprises refer to enterprises that operate in the mining and quarrying sector, manufacturing sector or in the production and supply of power, gas and water. The NBS survey is the primary source for the construction of numerous aggregate statistics used in the China Statistical Yearbooks. The NBS data includes the firm's identification (tax code) and basic information such as year founded, location, ownership type, employment, China industrial classification (CIC) code and principle products. The data also provide information on more than 50 financial variables from the accounting statements, including capital, assets, liabilities, creditors equity, gross output, industrial value-added, sales, income, profits, investment, value of exports, current / accumulated depreciation, the wage bill and R&D expenses. As firms in manufacturing industries are most active in international trade , in this chapter I focus on firms in the manufacturing sector only (CIC13-43).

Trade data record all import and export transactions that enter or exit through Chinese customs. Each record represents a shipment and contains detailed information, including time of the transaction (month and year), type of trade (import/ export), exporting/importing firm identifier, ownership type, commodities traded classified by 8-digit Harmonized Commodity Description and Coding System (HS), value, quantity, unit, destination country (of the exported commodities) / country of origin (of the imported products), type of trade (ordinary trade or processing trade) and finally mode of transport. In this study, I focus only on ordinary trade for the reason that firms engaged in processing trade are by definition importers and exporters as the imported goods have to be exported after processed and assembly. The effects of imported inputs on firms' export performance cannot be accurately estimated by looking at processing trade firms.[1] The customs data on imports records include transactions of all types of products,

[1]See Wang and Yu (2012) and Yu (2014) for the study of Chinese processing trade.

including capital goods, intermediate goods and final goods. The dataset has no explicit specification for the types of products imported. In order to capture the effect of imported inputs, I follow the Broad Economic Categories (BEC) by the United Nations and drop all observations with final goods imported. United Nation Statistics Division classifies each HS 6-digit into one of the BEC categories. Since the imports data are recorded at the HS 8-digit level, I first aggregate the data into 6-digit level and classify the imported products into three BEC categories using the correspondence table.[1]

Our next step is to match the NBS data with the trade data. We restrict our analysis to those firms that participate in the survey for the whole period.[2] Since the NBS and trade data use different coding systems for the firm identifier, I cannot merge these datasets by firm identification codes alone. Hence, I do the matching using a number of common variables in both datasets, e.g., firm name, registration place and year of the establishment.[3] Following Brandt et al. (2012), I exclude observations with incomplete records or negative values of key variables such as firm age, assets, real capital stock, number of employees, output, value-added and total wages. We also drop abnormal observations if any of the following are found to be negative: net of total assets and fixed assets; net of total assets and current assets; net of total liability and current liability; net of current depreciation and accumulated depreciation.

To measure firm-level total factor productivity (TFP) I follow De Loecker (2007)'s approach. Detailed mehtod on TFP estimation is presented in Appendix 2A in Chapter 2.

[1]The commodities traded during 2002 and 2006 are coded in HS 2002 version. Correspondence between BEC and HS 2002 can be found at: https://unstats.un.org/unsd/cr/registry/regso.asp?Ci=12&Lg=1.

[2]Girma et al. (2004) and Arnold and Hussinger (2005) use a similar rule in their data construction.

[3]Brandt et al. (2012) use the same survey data for the period 1998 to 2007. Our matching method is similar to theirs except that they first link firms over time with IDs and then match firms that might have changed their IDs as a result of restructuring, merger or acquisition using other information such as firm's name, address, industry, etc. They point out that only 4% of all matches are constructed using information of the firm other than IDs. As I use both IDs and names for matching, the fraction I exclude is small.

Tariff data are also used for the analysis in this chapter to address the endogeneity issue. The tariff rates at 6-digit HS level are obtained from the Integrated Database, World Trade Organization (IDB, WTO).[1] We apply the tariff rates to each corresponding HS 6-digit products in the customs imports data. Industry-level input tariff is widely used in studies of trade liberalization in the literature (Amiti and Konings, 2007; Feng et al., 2012; Goldberg et al., 2010; Topalova and Khandelwal, 2011). The drawback of using industry-level input tariff for firm-level studies is the measure is usually highly aggregated and cannot capture the difference across firms in the same industry. Following Bas and Strauss-Kahn (2014), I construct a firm-level import tariff relying on the tariff rate of each HS 6-digit product firms import. Since firms change the mix of their imported inputs over time and imports are negatively correlated with tariff rates, I measure the firm-level tariff as the simple average of all imported inputs of the firms.[2]

As will be discussed shortly in the next section, our empirical study will comprises two parts, (a) the impact of imported inputs on firm-level export performance (number of exported varieties and value of exports); (b) transaction-level link between the quality of exports and the imported inputs. The sample used in first part analysis consists 243,080 observations and the second part 639,274 observations.

5.2.2 Descriptives

We present summary statistics in Table 5.1 (see Table 5A.1 for the definition of the variables in Appendix A.). For our main variables of interest in terms of export scope, I calculate the num-

[1]Tariff data at the standard codes of the Harmonized System (HS) for all WTO members are available at: http://tariffdata.wto.org/reportersandproducts.aspx.

[2]Lileeva and Trefler (2010), Fan et al. (2015) and Yu (2014) also use firm-level tariffs for their empirical analysis.

ber of products exported, number of destination country traded with and number of varieties exported. As standard in the international trade literature, an HS 6-digit code can be considered to correspond to a product and the number of countries refers to the partner countries a firm trades with. Following Broda and Weinstein (2006), Arkolakis et al. (2008), Khandelwal (2010), Chen and Ma (2012) and Martin and Mejean (2014), I define a variety as an HS product exported to (imported from) a country. The intuition is that an HS product imported from US is considered to have different technology or labour input from that imported from Mexico. In this way, they are counted as different varieties. On average, Chinese firms export 7 HS 6-digit products to some 8 countries over the period. In terms of the number of varieties, the average is 21 and the maximum number is 1,805. For the imports, the average number of varieties is 22 and the maximum is 1,302.

Chinese firms import all kinds of products from a number of countries. To get more understanding of the structure of the imports, I divide the imported inputs into groups depending on product differentiation and origins of country. We first divide imported products into homogenous and differentiated goods following Rauch (1999)'s method. Rauch (1999) classifies internationally traded commodities into three groups: those traded on organized exchanges, those traded as having a reference price, and other commodities that could not be priced by either of these means. Commodities in the first two groups are classified as homogenous products while the other group are differentiated products. Rauch (1999) proposes both conservative and liberal classifications and I follow the liberal classification in this study.[1] As the customs

[1]The classification of SITC Rev. 2 commodities into the three groups can be found at: http://www.macalester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/TradeData.html#Rauch. We use the HS and SITC conversion tables and classify the HS 6-digit products into the three categories. The HS and SITC conversion tables are available at the UN web: http://unstats.un.org/unsd/trade/conversions/HS%20Correlation%20and%20Conversion%20tables.htm.

data provide information on the origin of the countries for the imported products, I also divide the imported products into two groups: products imported from high-income economies and those from other countries according to the classification of the World Bank.[1] The bottom part of Table 5.1 shows that Chinese firms import many homogenous products (about 19 varieties on average) and few differentiated products (approximately 3 varieties), and about 16 varieties from high income economies and less than 7 from other countries. In terms of values, these firms import more homogenous products than differentiated products and import more from high income economies than from other countries.

In Table 5.2 I break down the number of varieties exported and imported into several groups. About 10% of the firms export only 1 variety, almost 50% export less than 10 varieties and over 90% export less than 50 varieties. There are 6.5% of firms exporting 50 to 99 varieties and less than 3% export 100 varieties or more. On the import side, more than 20% of firms import just 1 variety, double that of exports. However, there are over 12% of firms importing 50 or more varieties, more than that of exports. Although the majority of firms export or import less than 50 varieties, some firms export and import a large number of varieties. For example, in 2006, 3M China Ltd exported 286 varieties and imported 386 varieties and Huawei Technologies Co. Ltd., the Chinese telecommunication giant, exported 1,226 varieties and import 509 varieties. In the next section, I examine the links between imported inputs and export performance of the Chinese manufacturing firms.

[1] The list of high income economies can be found at: http://data.worldbank.org/about/country-and-lending-groups#High_income.

Table 5.1: Summary statistics of key variables

Variables	Obs	Mean	Std. Dev.	Min	Max
age	31,139	2.287	0.565	0	5.050
employment	31139	5.709	1.179	1.386	11.790
wage	31136	2.808	0.622	-2.389	10.184
TFP	30,678	7.126	1.157	-0.899	12.828
No. of export products	31,139	7.363	10.031	1	282
No. of countries exported to	31,139	8.516	10.372	1	155
No. of export varieties	31,139	20.530	40.161	1	1,805
Export value	31,139	15.229	2.514	2.070	24.330
No. of import varieties	31,139	22.200	47.713	1	1,302
No. of import varieties: differentiated	31,139	3.427	8.622	0	286
No. of import varieties: homogenous	31,139	18.773	43.026	0	1,222
No. of Import varieties: high-income countries	31,139	15.532	37.230	0	1,257
No. of import varieties: non high-income countries	31,139	6.669	18.346	0	507
Import value	31,139	13.044	3.266	1.942	23.229
Import value: differentiated	31,139	7.112	6.493	0	22.437
Import value: homogenous	31,139	11.680	4.630	0	23.224
Import value: high-income countries	31,139	10.588	5.715	0	23.124
Import value: non high-income countries	31,139	8.000	6.516	0	22.120

Notes: Please see detailed definition of the variables in the Appdendix.

Table 5.2: Export and import varieties of Chinese manufacturing firms

No. of varieties	Exports		Imports	
	No. of firms	% of total	No. of firms	% of total
1	2,977	9.56	6,497	20.86
2-9	12,553	40.31	12,542	40.28
10-49	12,746	40.92	8,300	26.65
50-99	2,032	6.53	2,245	7.21
100 or over	831	2.67	1,555	4.99

5.3 Econometric Analysis

We now explore the links between the use of imported inputs and firms' export performance. The first part of our analysis examines whether more imported inputs increase the scope of firms' exports. We aggregate firms' export and import transactions in each year and link both sides of international activities with other firm characteristics. We expect a positive impact of increased use of imported inputs on firms' export scope. The reason is that the use of imported inputs on the one hand will improve firms' productivity with more advanced technology embedded, and on the other hand lower firms' production costs with cheaper foreign inputs. Firms will raise their profit thanks to higher productivity and lower input costs. Hence they can produce and export more varieties with their improved technology and ability to bear the export fixed costs. The second part of the analysis tests whether the use of imported inputs improves the quality of export varieties. We use detailed data on firms' export transactions at firm-HS6-country-year level, linking firm-level imports information and other firm-level characteristics. A positive impact of imported inputs on the quality of export varieties is also expected since the use of imported inputs raises the quality of products and firms export better quality varieties aiming for larger foreign market share.

5.3.1 Imported inputs and scopes of exports

We first examine the impacts of imported input on firms' extensive margins of exports by estimating the following equation:

$$EXPno_{it} = \alpha + \beta IMP_{it} + \gamma X_{it} + \eta_i + \delta_t + \varepsilon_{it} \tag{5.1}$$

where i and t index firm and year respectively; $EXPno$ represents number of export varieties; IMP represents number of import varieties ($IMPno$), or (and) logarithm of import value ($IMPvalue$); X is a vector of control variables of firm characteristics, including (log) firm age, (log) average wages of employees, TFP, size and ownership dummies; η is the firm fixed effect, δ is the year fixed effect and ε is the error term. Please refer to Table 5A.1 for detailed definition of variables. We expect the coefficient on $IMPno$ or $IMPvalue$ to have a positive sign.

As the dependent variable is discrete and non-negative counts of firms' export varieties, estimation of Equation (5.1) with a simple linear estimator, e.g., ordinary least square (OLS), is not appropriate. The reason is that in a simple linear setting the independent variable can assume any real value and predict negative or non-integer values for the outcome variable while independent variable always has non-negative discrete values of the count data.

Poisson and negative binomial methods can be used for estimation of count data as both distributions are discrete probability distributions. The most important property of Poisson distribution is equi-dispersion of the count data, i.e., equality of mean and variance (Cameron and Trivedi, 2005, p.668). However, the equidispersion property is violated in our case as in most applied studies: the variance of the number of export varieties (1,612.87) far exceeds the mean (20.53). Over-dispersion exists because of many factors and the most common one is heterogeneity (Cameron and Trivedi, 2010, p.569). In our case, the large variance in the number of export varieties is generated by firm heterogeneity such as strategies on whether to export, what and how many products and where to export vary much across firms. Such over-dispersion of the distribution of the dependent variable means a Poisson estimator is not appropriate.[1] Negative binomial estimator is designed to explicitly handle the over-dispersion of count data, since the negative binomial distribution has an extra parameter than the Poisson, the dispersion parameter can be used to adjust the variance independently of the mean and improve the efficiency in estimation(Cameron and Trivedi, 2010, p.641). Hence I choose the negative binomial estimator.

Unobserved firm fixed effects are included in Equation (5.1) and random effects (RE) or fixed effects (FE) models can be used to control for such effect. RE models treat the unobserved effects as random variables while FE models treat them as parameters. Hausman tests favour fixed effects estimation, hence I run fixed effects negative binomial regressions of Equation (5.1). The FE negative binomial estimator requires that there be at lest two periods of data, leading to 2,671 observations dropped because of only one observation in these firms.

We now look at the results. Since the estimates are obtained by non-linear regressions, the coefficients cannot be interpreted as the percentage changes directly. To facilitate interpretation I calculate the average marginal effects (AME) of the number of imported varieties and values and present them in Table 5.3. The marginal effect (ME) of a regressor as specified above denoted as Z_j is computed as $\partial E(y_{it}|Z_{it})/\partial X_{j,it} = \beta_j * E(y_{it}|Z_{it}) = \beta_j exp(Z'\beta)$. AME is

[1] We have also tried Poisson regressions, and tested the goodness of fit. The associated p-value under the assumption of a Poisson distribution is $Prob > \chi^2(30, 663) = 0.0000$, which implies the number of export varieties does not have a Poisson distribution and Poisson estimator is inappropriate here.

the average ME at each Z.[1] The coefficients corresponding to specifications of Table 5.3 are presented in Table 5A.2 in the Appendix.

After controlling for firm other characteristics e.g., age, wage, TFP, size and ownership, and year fixed effects, the results in Columns (1) to (3) show that the import varieties and import values have a positive and highly significant effect (at 1% level) on the number of export varieties. The AMEs of imported varieties and log imported values are 0.01 and 0.083 respectively. This means that on average a firm importing one more variety is expected to lead to 0.01 additional export variety and for one unit increase of log import value, the exported variety is expected to increase by 0.08. Although highly significant, the magnitude of the impact of imported inputs on export varieties seems to be small. The reason may be that I have a large sample of heterogenous firms and over the period most of the firms have similar pattern for their imports and exports since changing the mix of import or export products and partners incur large costs. Besides, it also takes time for the newly imported inputs being added to the production and firms cannot expand the products and partners for their exports immediately. Hence I do not observe very large immediate impact of imported inputs on the number of export varieties. As for the control variables, except for wage which has a negative coefficient at 10% level, coefficients of age, size and foreign and HMT ownership are positive and significant at 1% and TFP is positive and significant at 5% level in Column (1) and 10% at Column (2). These results indicate that older, more productive firms and small and medium-sized firms (SME) and those with foreign and HMT ownership export more varieties. In line with the existing literature, due to the sunk and fixed costs in exporting, firms' experience and productivity and connection with foreign markets enables them to export more varieties abroad. Compared with the large firms,

[1] See Cameron and Trivedi (2010, p.639) and Hilbe (2011, p.126) for more details on marginal effects.

SMEs are more willing to expand their export scope aiming for more profit and expansion of their exports. The negative sign of wage can be explained as that Chinese exports are usually of low-skilled and technology content which do not require high skilled labour force. Firms paying low wages to employees, i.e., employing less skilled workers, are thus found to export more varieties.

In order to get a better understanding in what type of imported inputs contribute more to the exports, I separate the imported inputs by their quality differentiation and the origins. Interesting findings are observed. Results in Columns (4) to (6) show that both differentiated inputs and homogenous inputs have positive and significant effects on export varieties, but stronger effects are found for differentiated imported inputs than homogenous inputs (0.042 vs. 0.005 for the number of imported varieties in Column (4) and 0.039 vs. 0.008 for the vales of imported inputs in Column (5)). We also test whether the differences in effects of homogenous and differentiated inputs on export varieties are statistically significant. The null hypotheses that the effects of both groups of imported inputs are the same are rejected at usual level. However, when both the number of varieties and values of homogenous and differentiated imported inputs are taken into account together in Column (6), the value of differentiated inputs lose the significance and that of homogenous inputs seems to have the largest effect on the number of export varieties.

In Columns (7) to (9), I find that inputs imported from both high-income economies and other countries have a significant effect on export varieties (at 1% level). Bas and Strauss-Kahn (2014) also report similar finding that French firms expand their exported varieties when they import inputs from either developed or developing countries. The number of imported varieties from both groups have the same effect on the export scope as in Column (7). However, Bas and Strauss-Kahn (2014) find that the number of imported varieties from developed countries has a larger impact on the exported varieties. The value of inputs imported from high-income economies has a larger effect on export varieties compared to that from other countries (0.098 vs. 0.064 as in Column (8)). However, null hypotheses that the effects of imported inputs from high-income economies and other origins are the same cannot be rejected at usual level. When the number of varieties and values of inputs imported from high-income economies and other origins are taken into account together in Column (9), the value of inputs imported from high-income economies do not have a significant impact on the number of exported varieties. The value of inputs imported from non-high-income economies has a larger effect on the number of export varieties than the number of input varieties from both groups of countries. We cannot compare these findings with Bas and Strauss-Kahn (2014) as they do not look at the relationship between value of imported inputs and the number of exported varieties.

Table 5.3: Export varieties and imported inputs of Chinese manufacturing firms: fixed effects negative binomial model (AME)

Dependent variable: count of export varieties									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.010***		0.009***						
	(0.001)		(0.001)						
IMPvalue		0.083***	0.052***						
		(0.017)	(0.017)						
IMPno_diff				0.005***		0.005***			
				(0.001)		(0.001)			
IMPno_homo				0.042***		0.033***			
				(0.008)		(0.008)			
IMPvalue_diff					0.008***	0.027			
					(0.001)	(0.033)			
IMPvalue_homo					0.039***	0.101***			
					(0.007)	(0.028)			
IMPno_HI							0.010***		0.008***
							(0.001)		(0.001)
IMPno_nonHI							0.010***		0.010***
							(0.003)		(0.003)
IMPvalue_HI								0.064**	0.026
								(0.028)	(0.029)
IMPvalue_nonHI								0.098***	0.064**
								(0.025)	(0.026)
p-value(H0: diff=homo)				0.0000	0.0836				
p-value(H0: HI=nonHI)							0.9599	0.4288	

Table 5.3 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	1.525***	1.556***	1.525***	2.380***	1.599***	2.275***	1.422***	2.111***	2.039***
	(0.186)	(0.185)	(0.186)	(0.311)	(0.192)	(0.337)	(0.211)	(0.300)	(0.303)
wage	-0.121*	-0.137**	-0.133*	-0.046	-0.118	-0.083	-0.224***	-0.186*	-0.174*
	(0.070)	(0.070)	(0.070)	(0.099)	(0.072)	(0.105)	(0.078)	(0.099)	(0.099)
TFP	0.095**	0.099**	0.081*	0.062	0.070	0.050	0.055	0.009	-0.011
	(0.043)	(0.043)	(0.044)	(0.062)	(0.045)	(0.065)	(0.049)	(0.062)	(0.062)
SMALL	0.952***	0.895***	0.988***	0.558*	0.949***	0.753**	1.141***	0.750**	0.842***
	(0.253)	(0.254)	(0.254)	(0.334)	(0.258)	(0.349)	(0.272)	(0.308)	(0.309)
MEDIUM	1.412***	1.357***	1.428***	0.981***	1.414***	1.046***	1.555***	1.123***	1.194***
	(0.227)	(0.228)	(0.227)	(0.290)	(0.231)	(0.300)	(0.241)	(0.262)	(0.262)
FOREIGN	1.309**	1.438***	1.309**	-0.316	1.361**	-0.196	0.800	-0.900	-1.002
	(0.531)	(0.528)	(0.532)	(0.885)	(0.545)	(1.002)	(0.579)	(0.846)	(0.855)
HMT	1.443***	1.557***	1.451***	-0.249	1.529***	-0.117	0.816	-0.709	-0.783
	(0.536)	(0.533)	(0.537)	(0.894)	(0.551)	(1.011)	(0.586)	(0.852)	(0.861)
PRIVATE	-0.460	-0.386	-0.443	-1.300	-0.350	-0.920	-0.669	-1.578**	-1.575*
	(0.504)	(0.501)	(0.505)	(0.838)	(0.517)	(0.954)	(0.548)	(0.804)	(0.814)
COLLECTIVE	0.519	0.582	0.541	-1.369	0.536	-1.262	0.255	-0.615	-0.499
	(0.633)	(0.629)	(0.635)	(1.135)	(0.654)	(1.326)	(0.704)	(1.165)	(1.178)
Observations	28,942	28,942	28,942	28,942	28,942	28,942	28,942	28,942	28,942

Notes: Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

We now analyse the impact of imported inputs on intensive margins of exports by estimating the following equation:

$$EXPvalue_{it} = \alpha + \beta IMP_{it} + \gamma X_{it} + \eta_i + \delta_t + \varepsilon_{it} \tag{5.2}$$

where $EXPvalue$ is logarithm of export value of firm i in year t, X is a vector of control variables, including logs of firm age, average wages of employees, TFP, size and ownership dummies, η is the firm fixed effect, δ is the year fixed effect and ε is the error term. We expect IMP_{it} ($IMPno$ and $IMPvalue$ respectively) to have a positive effect on the dependent variable.

Within models (fixed effects) are used for the estimation of Equation (5.2) and Table 5.4 presents the results. Again, both the number and value of imported inputs are found to have positive significant effects on export value. As the dependent variable, export value, is in logarithm term and dependent variable $IMPno$ is in level term. The coefficient on $IMPno$ should be interpreted in this way: for a unit increase in $IMPno$ the percentage change in export value is 100β (Wooldridge, 2002, p.45). Therefore, holding all other independent variables constant, an additional imported input is associated with a 0.6 percentage increase in export value as shown in Column (1). We then look at the coefficient of import value in Column (2). As both import value and export value are in logarithm terms, the coefficient can be interpreted as the elasticity. So a 10 percentage increase in value of imported inputs leads to a 0.61 percentage increase in export value, holding other dependent variables constant.

Both the variety number and value of differentiated and homogenous imported inputs have a positive impact on export value as in Columns (4) and (5). The number of homogenous inputs has a stronger impact on export value than that of the differentiated inputs (0.014 vs 0.003 as seen in Column (4)) and the different effects of both groups are statistically significant. The value of differentiated inputs has a slightly larger effect on export value than that of homogenous inputs as shown in Column (4), but the null hypothesis that the effects of both groups are the same cannot be rejected in usual level. When both the number of varieties and values of homogenous and differentiated imported inputs are taken into account together in Column (6), the value of differentiated inputs has the largest effect on the export values.

In Columns (7) to (9), I find that inputs imported from both high-income economies and other countries have significant effects on export values (at 1% level). The variety number of inputs imported from non-high-income economies has a slightly larger effect on export values than that from high-income economies (0.008 vs. 0.005 as in Column (7)) while value of inputs imported from high-income economies has a stronger effect on export values than that from non-high-income economies (0.071 vs. 0.043 as in Column (8)). The difference in effects of imported inputs from both groups of countries are statistically significant. Finally when taken in to consideration together, both the variety number and value of inputs imported from both groups of countries have positive and significant effects on export values as indicated in Column (9). The value of inputs imported from high-income economies has a stronger impact on firms' expansion of export scopes than the other three factors.

Table 5.4: Export value and imported inputs of Chinese manufacturing firms: within model

Dependent variable: log of export values									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.006***		0.005***						
	(0.000)		(0.000)						
IMPvalue		0.061***	0.048***						
		(0.006)	(0.006)						
IMPno_diff				0.003***		0.002***			
				(0.001)		(0.001)			
IMPno_homo				0.014***		0.008**			
				(0.004)		(0.004)			
IMPvalue_diff					0.026***	0.042***			
					(0.004)	(0.012)			
IMPvalue_homo					0.020***	0.037***			
					(0.003)	(0.010)			
IMPno_HI							0.005***		0.003***
							(0.001)		(0.001)
IMPno_nonHI							0.008***		0.007***
							(0.001)		(0.002)
IMPvalue_HI								0.071***	0.057***
								(0.012)	(0.012)
IMPvalue_nonHI								0.043***	0.027***
								(0.010)	(0.010)
p-value(H0: diff=homo)				0.0058	0.1739				
p-value(H0: HI=nonHI)							0.0668	0.0999	

Table 5.4 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	0.657***	0.663***	0.652***	1.174***	0.697***	1.302***	0.634***	0.905***	0.871***
	(0.101)	(0.101)	(0.101)	(0.156)	(0.104)	(0.173)	(0.117)	(0.180)	(0.179)
wage	0.028	0.022	0.024	0.060*	0.037	0.052	0.015	0.052	0.053
	(0.025)	(0.025)	(0.025)	(0.035)	(0.026)	(0.037)	(0.028)	(0.039)	(0.039)
TFP	0.120***	0.119***	0.113***	0.101***	0.116***	0.095***	0.127***	0.105***	0.097***
	(0.015)	(0.016)	(0.015)	(0.021)	(0.016)	(0.023)	(0.018)	(0.025)	(0.025)
SMALL	-0.433***	-0.479***	-0.416***	-0.581***	-0.478***	-0.539***	-0.390***	-0.496***	-0.438***
	(0.100)	(0.100)	(0.100)	(0.129)	(0.101)	(0.136)	(0.107)	(0.133)	(0.133)
MEDIUM	-0.189**	-0.234***	-0.182**	-0.343***	-0.232**	-0.357***	-0.175*	-0.310***	-0.264**
	(0.090)	(0.090)	(0.090)	(0.113)	(0.090)	(0.118)	(0.095)	(0.113)	(0.113)
FOREIGN	0.538**	0.524**	0.531**	0.346	0.455*	0.462	0.708***	0.703	0.710*
	(0.231)	(0.231)	(0.230)	(0.379)	(0.234)	(0.432)	(0.255)	(0.431)	(0.430)
HMT	0.500**	0.495**	0.500**	0.376	0.437*	0.470	0.630**	0.684	0.686
	(0.232)	(0.232)	(0.232)	(0.382)	(0.235)	(0.435)	(0.257)	(0.433)	(0.431)
PRIVATE	0.241	0.241	0.246	0.222	0.253	0.308	0.297	0.465	0.482
	(0.201)	(0.201)	(0.200)	(0.330)	(0.203)	(0.378)	(0.219)	(0.371)	(0.370)
COLLECTIVE	0.443*	0.447*	0.448*	0.217	0.478*	0.182	0.573**	0.590	0.611
	(0.249)	(0.249)	(0.248)	(0.420)	(0.253)	(0.493)	(0.279)	(0.523)	(0.521)
Constant	12.147***	11.539***	11.603***	11.983***	12.732***	9.819***	12.945***	10.025***	10.322***
	(0.333)	(0.341)	(0.340)	(0.571)	(0.373)	(0.611)	(0.412)	(0.617)	(0.616)
Observations	31,632	31,632	31,632	31,632	31,632	31,632	31,632	31,632	31,632
R-squared	0.135	0.133	0.138	0.154	0.138	0.168	0.139	0.164	0.169

Notes: Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

5.3.2 Imported inputs and the quality of exports

The next stage is to explore the relationship between imported inputs and the quality of exports. The link between imported inputs and the quality of export varieties is estimated using the following equation:

$$ln\lambda_{ikct} = \alpha + \beta IMP_{it} + \gamma X_{it} + \lambda_{ikc} + \delta_t + \varepsilon_{ikct} \quad (5.3)$$

where $ln\lambda$ is the natural logarithm of the quality of product k exported to country c by firm i in year t; IMP represents number of imported varieties ($IMPno$), or (and) logarithm of import value ($IMPvalue$); X is a vector of control variables, including logs of firm age, average wages of employees, TFP, size and ownership dummies; λ is product-country pair fixed effect and ε is the error term. We expect $IMP_{it}$to have a positive impact on the quality of firms' exported products.

Prices are widely used as a proxy for quality in the international trade literature (see for example, Baldwin and Harrigan, 2011; Cadot et al., 2014; Hallak, 2006; Kugler and Verhoogen, 2012; Manova and Zhang, 2012; Schott, 2004; Swenson and Chen, 2014). The assumption for such proxy is that better quality products require better inputs and consumers are prepared to pay higher prices for better quality. Unit values are then used as surrogates for prices. The advantage of this approach is that unit values are convenient to calculate, simply dividing the value by the quantity of each product. However, unit values could reflect not just quality, but also variations in manufacturing costs (e.g., materials and labor costs), transportation costs and tariff costs. Hence, products of the same quality may have different unit values simply because of wage differentials across different markets. Also, unit values may not be appropriate proxies for quality for products with style or brand effects. For example, although having the same unit prices, women's blouses produced in Italy are more popular in the US market than those produced in Brazil because of the US consumers' love of brand and such unit values cannot reflect the different quality of these two varieties. Also as evidenced bySilver (2007), substantial bias exists in unit value import or export indices for the representation of export and import price changes.

In this chapter, I measure a variety's quality following Khandelwal et al. (2013)with information on prices and market shares is accounted for.[1] The intuition is that conditional on price, a variety with a higher quantity is assumed to possess higher quality. For varieties with the same prices, the better quality one will have a larger market share. Khandelwal et al. (2013) establish a utility function where consumers' preferences incorporate quality (λ) as:

$$U = (\int_{\zeta \in \Omega} (\lambda_\zeta q_\zeta)^{\frac{\sigma-1}{\sigma}}) d\zeta)^{\frac{\sigma}{\sigma-1}} \tag{5.4}$$

where λ is quality and q is quantity of variety ζ, σ is the elasticity of substitution across varieties ($\sigma > 0$) and Ω is the set of all varieties imported. As I define a variety as a product k to country c exported by firm i in year t, the demand equation is given by:

$$q_{ikct} = \lambda_{ikct}^{\sigma-1} p_{ikct}^{-\sigma} P_{ct}^{\sigma-1} Y_{ct} \tag{5.5}$$

where p_{ikct} is the export price of product k in country c of firm i at time t, P_{ct} and Y_{ct} are the

[1] Khandelwal (2010) and Amiti and Khandelwal (2013) use same methodology for product quality estimation and they define the quality of an imported variety to its market share after controlling for exporter size and price.

overall price index and expenditure of country c at time t. Taking natural logarithms of the above equation I can define the quality as the residual from the following:

$$lnq_{ikct} + \sigma lnp_{ikct} = \alpha_k + \alpha_{ct} + \varepsilon_{ikct} \quad (5.6)$$

where α_k is the product fixed effect that captures differences across products, α_{ct} is country-year fixed effect which controls for price index P and expenditure Y of country c at time t and $(\sigma - 1)ln\hat{\lambda}_{ikct} = \hat{\varepsilon}_{ikct}$. Hence the log of quality can be estimated by $ln\hat{\lambda}_{ikct} = \hat{\varepsilon}_{ikct}/(\sigma - 1)$ with a given value of σ. Amiti and Khandelwal (2013) note that defining quality to be inclusive of a residual from conditioning price on market share in the utility function is analogous to the productivity literature that interprets total factor productivity as the residual from conditioning output on observable inputs in the production function. Following Khandelwal et al. (2013), I set elasticity of substitution between varieties, σ, equal to 4 in estimating the quality of each variety.[1]

Results of regressions on Equation (5.3) are presented in Table 5.5. Similar patterns are observed as previously found for the export varieties and value: importing more inputs leads to an increase of the export quality. After controlling for firm other characteristics e.g., age, wage, TFP, size and ownership, and year fixed effects, the results in Columns (1) to (3) show that the import varieties and import values have a positive and highly significant effect (at 1% level) on the quality of export varieties.

When the imported inputs are separated into differentiated and homogenous inputs in Columns

[1]Broda et al. (2006) estimate the elasticities of substitution for 73 countries and report the median σ as 3.4 for China. Anderson and van Wincoop (2004) review trade barriers and survey different estimates of elasticity of substitution and values of σ range from 5 to 20, but they use $\sigma = 4$ in the estimation of trade costs and price dispersion.

(4) to (6), both differentiated inputs and homogenous inputs have positive and significant effects on the quality of export varieties, but stronger effects are found for the variety number of homogenous imported inputs than that of differentiated inputs in Column (4) and the value of differentiated inputs than that of homogenous inputs in Column (5). The difference in effects of both differentiated and homogenous inputs on exort quality are statistically significant at usual level. When both the number of varieties and values of homogenous and differentiated imported inputs are taken into account together in Column (6), the value of differentiated inputs has strongest effects on the quality of exported varieties, followed by the variety number of imports and the value of homogeous inputs does not have a significant effect.

In Columns (7) to (9), I find that inputs imported from both high-income economies and other countries have significant effects on the quality of export varieties. The variety number of inputs imported from non-high-income economies has a larger effect on export quality than that from high-income economies as in Column (7) while value of inputs imported from high-income economies has a slightly stronger effect on export values than that from non-high-income economies as in Column (8). The null hypotheses that the effects of imported inputs from high-income economies and other origins are the same are rejected at usual level. When the number of varieties and values of inputs imported from high-income economies and other origins are taken into account together in Column (9), the significance of the value of inputs imported from non-high-income economies on export quality drop to 10% level and the number of varieties imported from non-high-income economies and value of inputs imported from high-income economies have same effect on export quality. do not have a significant impact on the number of exported varieties. The value of inputs imported from non-high-income economies has a larger effect on the number of export varieties than the number of input varieties from both groups of countries.

Finally, I also looked at the effect of firms' average quality of imports on the quality of export varieties in Column (10). The firm-level imports quality is constructed as the weighted average quality of import varieties which is estimated using the same method as specified above. The weights are the value of each imported variety over total imports value of each year. A positive and significant effect of average import quality on exports quality is observed, indicating firms that import better quality inputs export higher quality products.

Table 5.5: Export quality and imported inputs of Chinese manufacturing firms: within model

Dependent variable: quality of exported varieties										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IMPno	0.001***		0.000***							
	(0.000)		(0.000)							
IMPvalue		0.009***	0.007***							
		(0.001)	(0.001)							
IMPno_diff				0.000***		0.000***				
				(0.000)		(0.000)				
IMPno_homo				0.004***		0.003***				
				(0.001)		(0.001)				
IMPvalue_diff					0.005***	0.004***				
					(0.001)	(0.001)				
IMPvalue_homo					0.002***	0.001				
					(0.001)	(0.001)				
IMPno_HI							0.000***		0.000***	
							(0.000)		(0.000)	
IMPno_nonHI							0.002***		0.002***	
							(0.000)		(0.000)	
IMPvalue_HI								0.003***	0.002***	
								(0.001)	(0.001)	
IMPvalue_nonHI								0.002***	0.001*	
								(0.001)	(0.001)	
IMPquality										0.007***
										(0.002)
p-value(H0: diff=homo)				0.0010	0.0035					
p-value(H0: HI=nonHI)							0.0005	0.0581		

Table 5.5 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
age	0.124***	0.114***	0.118***	0.121***	0.112***	0.116***	0.121***	0.115***	0.117***	0.118***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
wage	-0.007	-0.006	-0.006	-0.007	-0.007	-0.006	-0.006	-0.007	-0.006	-0.008
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
TFP	0.033***	0.032***	0.032***	0.033***	0.033***	0.032***	0.032***	0.034***	0.032***	0.034***
	(0.004)	(0.004)	(0.004)	(0.004)	-0.004	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
SMALL	-0.077***	-0.080***	-0.076***	-0.077***	-0.079***	-0.075***	-0.078***	-0.079***	-0.076***	-0.081***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
MEDIUM	-0.027	-0.031*	-0.028	-0.027	-0.030	-0.027	-0.027	-0.029	-0.026	-0.031
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
FOREIGN	-0.191***	-0.199***	-0.196***	-0.191***	-0.189***	-0.186***	-0.191***	-0.198***	-0.196***	-0.190***
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)
HMT	-0.168***	-0.175***	-0.172***	-0.168***	-0.166***	-0.162***	-0.168***	-0.174***	-0.171***	-0.167***
	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
PRIVATE	-0.110**	-0.111**	-0.109**	-0.110**	-0.108**	-0.106**	-0.111**	-0.112**	-0.111**	-0.107**
	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
COLLECTIVE	-0.072	-0.074	-0.072	-0.076	-0.070	-0.071	-0.073	-0.078	-0.076	-0.065
	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)
Constant	-0.279***	-0.343***	-0.348***	-0.278***	-0.306***	-0.318***	-0.276***	-0.286***	-0.296***	-0.253***
	(0.091)	(0.092)	(0.092)	(0.091)	(0.091)	(0.091)	(0.091)	(0.091)	(0.091)	(0.090)
Observations	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033
R-squared	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.001	0.002	0.001

Notes: Robust standard errors clustered at the firm-HS6-country level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

5.4 Endogeneity Concerns

We now address the potential endogeneity issue. It could be argued that firms' import performance such as what product to import, where from and how much to import, is simultaneously determined with the dependent variable, export performance (number of varieties or values of exports). Rational firms will always estimate their input needs taking into account their exports. This could lead to reverse causality. Moreover, some unobserved effects, such as managerial strategy or foreign network, can affect firms' exports and also impact on the decision to import. Such unobserved heterogeneity is also a source of endogeneity. The instrumental variables (IV) and general method of moments (GMM) estimators are traditionally applied to deal with endogeneity concerns of this type.

The first part of our analysis is a count data model and the techniques for addressing the endogeneity in count data models are less developed compared to those for models with continuous dependent variables. Mullahy (1997) discusses IV, two-stage quasi- maximum likelihood (2SQML) and GMM estimations of count data models when some explanatory variables are endogenous. Windmeijer and Santos Silva (1997) apply the GMM estimator in the study of demand for health care where the dependent variable is the number of visits to doctors with the health index as a potential endogenous regressor. Baum et al. (2003) discuss the advantages of GMM and IV and state that "if heteroskedasticity is present, the GMM estimator is more efficient than the simple IV estimator, whereas if heteroskedasticity is not present, the GMM estimator is no worse asymptotically than the IV estimator." We test the heteroskedasticity of the errors, and the null hypothesis cannot be rejected in count data (dependent variable as number of exported varieties) as well as the continuous data cases (i.e., the export values and quality of exports as dependent variables). We thus choose GMM estimator for the analysis.

A brief introduction to GMM estimator is presented below. Let us define the following equation:

$$y_{it} = \beta X_{it} + \gamma Z_{it} + \epsilon_{it} \tag{5.7}$$

where y_{it} is firm export performance ($EXPno$, $EXPvalue$ and $EXPquality$ respectively), X_{it} includes strictly exogenous regressors, Z_{it} are endogenous regressors and predetermined regressors (which may include lag of y) and $\epsilon_{it} = \nu_i + \mu_{it}$ with ν_i as the fixed effects and μ_{it} the idiosyncratic shocks and $E[\nu_i] = E[\mu_{it}] = E[\nu_i \mu_{it}] = 0$. As previous sections, firm imported inputs (number of varieties and values) are the main variables of interest. Firm age, wage, TFP, size, ownership and year dummies are included in the explanatory variable set to control for firm characteristics and time effects.

Firm-specific time-invariant effects can be removed by differencing the above equation as:

$$\Delta y_{it} = \beta \Delta X_{it} + \gamma \Delta Z_{it} + \Delta \mu_{it} \tag{5.8}$$

However, applying OLS to 5.8 will produce biased and inconsistent parameter estimates. Even if μ_{it} are serially independent, any variable in ΔZ will still be endogenous as it may correlate with $\mu_{i(t-1)}$. A range of instrumental variable estimators have been proposed to provide consistent estimation. For example, Anderson and Hsiao (1982) proposed a two-stage-least-square (2SLS) estimator which uses $Z_{i(t-2)}$ or $\Delta Z_{i(t-2)}$ as an instrument for $\Delta Z_{i(t-1)}$ since either $Z_{i(t-2)}$ or $\Delta Z_{i(t-2)}$ is uncorrelated with $\Delta \mu_{it}$, as long as $\mu_{it} C$ is not serially correlated.

Anderson and Hsiao (1982) estimator yields consistent but not necessarily efficient estimates because it does not allow for all the available moment conditions or account for the structure of $\Delta\mu_{it}$.

Deeper lags of the endogenous variable can be used as additional instruments to improve efficiency because they introduce more information. Arellano and Bond (1991) construct a different number of instruments in each time period based on the moment equations that exist between lagged levels of Z_{it} and the first difference of the error term μ_{it}. This estimator is known as Arellano-Bond GMM estimator or first-difference GMM estimator. Performance for this first-difference GMM estimator improves as T increases. But this is also a drawback for this estimator because the deeper lags are used, the longer period of the sample needs to be.

Arellano and Bover (1995) and Blundell and Bond (1998) propose an alternative system GMM estimator. This method uses lagged levels as instruments for the regression in differences and lagged differences as instruments for the regression in levels. The system GMM estimator addresses the weak instrument problem that arises from using lagged levels of persistent explanatory variables as instruments for the regression in differences (Blundell and Bond, 1998). The system GMM estimator is applied to estimate the impact of imported inputs on export performance of Chinese manufacturing firms.[1] [2]

One of the challenges of the GMM estimator is the choice of good instruments which should

[1]See Roodman (2009) for the detailed introduction to how to choose between and apply difference and system GMM estimators. I follow Roodman (2009) and assess whether a first-difference or system GMM estimator to be used by applying OLS, fixed effects estimator and first-difference GMM estimator to the model respectively and compare the coefficients of lagged dependent variable obtained for each specification. The coefficient on lagged dependent variable is 1.053 obtained from OLS (which is upward biased), 0.525 from FE (which is downward biased) and 0.310 from first-difference GMM. As the coefficient obtained from first-difference GMM does not fall between those obtained from OLS and FE estimators. A system GMM estimator is thus chosen.

[2]The GMM estimation in the chapter is carried out in STATA with the command -xtabond2-.

be both relevant and valid: correlated with the endogenous regressors and at the same time orthogonal to the errors (Baum et al., 2003). To test for the validity of instruments, Hansen-*J* tests can be employed (Hansen, 1982). The null hypothesis is that the instruments, as a group, are exogenous. Roodman (2009) discusses the Hansen tests of overidentifying restrictions and concludes that Hansen-Sargan tests may have very little power in a model containing a very large set of excluded instruments under which case the difference-in-Sargan tests can be employed.

The system GMM results are presented in Table 5.6 to Table 5.8. Lagged levels dated as early as *t*-2 are potentially valid instruments if the error term in the original specification is iid (Arellano and Bond, 1991). However Hansen -*J* tests reject the validity of the *t*-2 instruments. Deeper lags of the endogenous variables, i.e., third or (and) fourth lags, are used as instruments. I treat all right-hand side variables except firm age and year dummies as potentially endogenous variables. Levels of these variables dated *t*-3 and earlier are used as instruments in the first-differenced equations and first-differences of these same variables lagged once are used as additional instruments in the level equations. In addition, firm-level imports tariffs are used as instrument. Firms' imported input tariffs are good instruments as they are on the one hand closely related with the firms' imports and on the other hand uncorrelated with firms' exports. AR(1), AR(2) and AR(3) tests for first-, second- and third-order autocorrelation in the first-differenced residuals are performed. Evidence of third-order autocorrelation in the first-differenced residuals would imply second-order serial correlation in levels, which would imply the *t*-3 instruments invalid (Brown and Petersen, 2009). In such case, only the *t*-4 instruments are employed. Due to the short period of the data, no further lagged instruments are available.

Table 5.6 reports the results from the regressions of imported inputs on export varieties.

First order autocorrelation in the first-differenced residuals are found in four out of seven specifications and no evidence of AR(2) and AR(3) tests in all specifications. Hansen J tests pass the overidentification restriction and difference-in-Hansen (C) tests pass the instrument validity in all specifications. Results in Column (1) show that the number of imported varieties have a positive and significant effect on the number of varieties exported while the value of imports negatively affects the number of exported varieties. Firm age, wage, TFP, size and ownership type have significant impact on firms' exported varieties. In Columns (2) to (4), the imported inputs are categorized into homogenous and differentiated goods. Column (2) indicates that the number of varieties of both homogenous and differentiated do not have a significant impact on the exported varieties, but Columns (3) and (4) show that the import value of differentiated inputs have a positive and significant effect on the exported varieties. In Columns (5) to (7), imported inputs are grouped by the country of origin. Neither the number of imported inputs from high-income economies or from other economies has a significant effect on the exported varieties as observed in Column (5). However, in Column (6) I find that value of inputs imported high-income economies have a positive and significant impact on the number of export varieties while the value of inputs imported from non-high-income economies has no effect on exported varieties. Finally in Column (7), numbers of inputs from both high-income economies and other origins and the value of imports from high-income economies have significant impacts on the exported varieties. The impact of value of imports from high-income economies are much higher than the others, indicating the higher value of inputs imported from high-income economies, the more varieties the firm exports.

Table 5.7 presents the GMM estimates from regressions on imported inputs and export value. AR(1) are observed in all specifications and AR(2) are found in four out of seven specifications, but no evidence of AR(3) in all specifications. Low p-values of Hansen J tests are obtained in most specifications. As discusses in Roodman (2009) the Hansen tests may have very little power in a model containing a very large set of excluded instruments under which case the difference-in-Sargan tests can be employed. The difference-in-Hansen (C) test passes the instrument validity in three specifications (Columns (3), (4) and (6)). Given the short period of sample, I cannot test the validity of instruments of longer lags of explanatory variables for the other specifications. A positive and 10% significant relationship between the number of imported varieties and export values is found in Columns (3) and (4). To the contrast, significant (at 5% level) and positive impacts are observed in inputs from non-high-income economies. When all included in Column (7), only the value of inputs imported has a positive and significant (at 1% level) effect on export value. When imported inputs are divided into homogenous and differentiated, only the variety number or value of differentiated inputs have a positive and significant impact on the export values while the variety number or value of homogenous imported varieties have no significant effect. When both the variety number and value of both homogenous and differentiated of imports are included in the regression, only the value of differentiated inputs have a positive and significant effect on firms' export values.

GMM estimation is repeated to examine the link between the imported inputs and the quality of export varieties and results are presented in Table 5.8. AR(1) are observed in all specifications, AR(2) are found in six out of eight specifications and AR(3) exist in four out of eight specifications. For specifications where the third-order autocorrelation of first-differenced residuals is present, only the t-4 instruments are applied. Low p-values of Hansen J tests are found in all eight specifications, rejecting the overidentification of instruments. However, the difference-in-Hansen test passes the instrument validity in all specifications. We cannot test the validity of instruments of longer lags due to the short period of data.

In Column (1) both the variety number and value of imported inputs have negative and significant impact on the quality of exports. When the imported inputs are divided into homogenous and differentiated groups, the number of differentiated varieties imported does not have a significant impact on export quality, and that of differentiated varieties imported has a negative effect, but the magnitude is rather small (-0.003 only) as shown in Column (2). When the value of imports of both groups are examined in Column (3), the value of differentiated imported inputs has a significant (at 1% level) and positive impact on the export quality, while that of homogenous inputs has a positive and 5% significant impact. In Column (4) where both variety number and value of homogenous and differentiated inputs are included in the regression, only the value of differentiated inputs has a positive and significant effect on export quality, indicating higher value of differentiated inputs imported, the better quality of goods the Chinese manufacturing firms export. Imported inputs are grouped by their origins in Columns (5) to (7). Neither the coefficients of variety numbers imported from high-income economies or other origins are significant as in Column (5) and only the value of inputs imported from non-high-income economies are positive (0.216) at 1% level in Column (6). When the variety number and value of inputs imported from both high-income economies and other origins are included in the regression in Column (7), the numbers of inputs imported from both groups have negative impact on the export quality, and the value of inputs imported from non-high-income economies has positive and significant effect on export quality while that from high-income economies does not have significant effect. Finally, the relationship between the imports quality and exports quality is examined in Column (8), a positive and significant (1% level) impact is found for imports quality on exports quality. So Chinese manufacturing firms are found to export better quality goods with high quality inputs imported. Such finding is in line with Fan et al. (2015).

Table 5.6: Imported inputs and exports of Chinese manufacturing firms: GMM estimation (I)

Dependent variable: count of export varieties							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IMPno	0.126**						
	(0.063)						
IMPvalue	-3.659***						
	(0.625)						
IMPno_diff		0.302		0.307			
		(0.387)		(0.569)			
IMPno_homo		0.042		-0.193			
		(0.075)		(0.174)			
IMPvalue_diff			5.850**	6.049**			
			(2.323)	(2.538)			
IMPvalue_homo			-3.435	-4.055			
			(3.707)	(3.873)			
IMPno_HI					0.018		-0.304*
					(0.101)		(0.172)
IMPno_nonHI					0.217		0.631*
					(0.179)		(0.355)
IMPvalue_HI						7.492**	6.011**
						(3.722)	(2.890)
IMPvalue_nonHI						-3.251	-2.360
						(3.447)	(2.688)

Table 5.6 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
age	-1.515*	-1.434*	-3.322**	-3.383**	-1.193	-6.506*	-4.324*
	(0.814)	(0.784)	(1.355)	(1.396)	(0.838)	(3.337)	(2.243)
wage	3.380***	-0.905	-4.697	-1.165	-0.479	-10.342**	-6.485*
	(1.126)	(0.887)	(4.972)	(3.457)	(1.086)	(5.037)	(3.695)
TFP	4.610***	2.025***	0.944	2.072	2.078***	-1.546	-0.402
	(0.672)	(0.614)	(2.038)	(1.675)	(0.625)	(2.415)	(1.793)
SMALL	-22.228***	-16.663***	-17.900*	-25.672***	-15.389***	-9.898	-10.141
	(4.971)	(4.386)	(9.783)	(9.635)	(4.327)	(9.990)	(7.403)
MEDIUM	-11.464**	-8.830**	-11.038	-17.183**	-7.813*	-3.807	-4.541
	(4.853)	(4.075)	(6.979)	(7.694)	(4.057)	(6.293)	(5.754)
FOREIGN	2.120	1.203	-15.852*	-13.476*	1.051	10.596	8.477
	(3.651)	(3.293)	(8.374)	(8.138)	(3.258)	(10.541)	(7.434)
HMT	5.537	5.108	-9.933	-9.580	4.377	31.032	21.833
	(3.530)	(3.149)	(8.171)	(8.260)	(3.232)	(19.347)	(13.302)
PRIVATE	9.809***	9.787***	3.917	2.853	9.629***	20.058***	17.205***
	(3.533)	(3.153)	(5.845)	(6.185)	(3.114)	(7.675)	(5.978)
COLLECTIVE	6.896*	9.122***	-0.344	-1.370	8.912**	23.261**	19.245**
	(3.921)	(3.526)	(8.830)	(9.514)	(3.509)	(9.791)	(7.515)
Observations	30,675	30,675	30,675	30,675	30,675	30,675	30,675
p-value: AR(1)	0.333	0.094	0.082	0.080	0.090	0.120	0.126
p-value: AR(2)	0.267	0.252	0.956	0.989	0.265	0.664	0.734
p-value: AR(3)	0.989	0.739	0.848	0.899	0.711	0.432	0.553
p-value: *J*-test	0.060	0.133	0.514	0.637	0.105	0.189	0.467
p-value: *C*-test	0.311	0.169	0.337	0.273	0.151	0.307	0.654

Notes: AR(*n*) is the Arellano-Bond test for *n*th order of the autocorrelation. *J*-test is the Hansen test for overidentification. *C*-test is the difference-in-Hansen test which tests the validity of subsets of instruments. Robust standard errors are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

Table 5.7: Imported inputs and exports of Chinese manufacturing firms: GMM estimation (II)

Dependent variable: log of export values							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IMPno	0.017*						
	(0.010)						
IMPvalue	0.165						
	(0.620)						
IMPno_diff		0.047*		-0.017			
		(0.028)		(0.048)			
IMPno_homo		0.011		0.007			
		(0.008)		(0.012)			
IMPvalue_diff			0.590**	0.536**			
			(0.244)	(0.244)			
IMPvalue_homo			-0.495	-0.430			
			(0.400)	(0.362)			
IMPno_HI					0.018**		-0.006
					(0.008)		(0.005)
IMPno_nonHI					0.005		0.011
					(0.017)		(0.008)
IMPvalue_HI						-0.101	0.262***
						(0.344)	(0.081)
IMPvalue_nonHI						0.456*	0.005
						(0.254)	(0.054)

Table 5.7 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
age	0.423	0.008	-0.117	-0.079	0.198*	0.502	-0.181
	(0.648)	(0.138)	(0.384)	(0.348)	(0.112)	(0.317)	(0.408)
wage	-8.378	-4.823***	-5.553	-5.236*	-5.922***	-4.924	-1.020
	(8.558)	(1.114)	(3.504)	(2.850)	(1.583)	(5.790)	(0.669)
TFP	2.910*	1.410	2.732*	2.103	2.653**	1.143	-0.533
	(1.595)	(0.893)	(1.544)	(1.543)	(1.194)	(1.340)	(0.379)
SMALL	6.126	1.906	5.109	4.035	4.571**	2.679	-1.848**
	(4.307)	(2.396)	(4.369)	(3.836)	(2.009)	(2.700)	(0.867)
MEDIUM	3.468	2.714	4.792	4.025	2.612**	1.267	-0.544
	(2.152)	(1.743)	(2.988)	(2.682)	(1.041)	(1.372)	(0.696)
FOREIGN	-0.144	0.148	-2.031	-1.753	-0.191	-0.834	0.862
	(1.081)	(0.522)	(1.412)	(1.234)	(0.288)	(1.260)	(2.990)
HMT	-0.850**	-0.368	-2.491*	-2.212*	-0.523*	-2.346*	-0.618
	(0.406)	(0.473)	(1.336)	(1.187)	(0.288)	(1.425)	(3.281)
PRIVATE	-1.649	-0.692	-2.253*	-1.878	-1.066*	-1.056	0.706
	(1.415)	(0.594)	(1.292)	(1.163)	(0.558)	(0.745)	(2.683)
COLLECTIVE	-2.711	-1.221	-3.290*	-2.854*	-1.736**	-1.905	0.417
	(2.251)	(0.751)	(1.740)	(1.546)	(0.712)	(1.197)	(6.590)
Observations	30,675	30,675	30,675	30,675	30,675	30,675	30,675
p-value: AR(1)	0.265	0.000	0.004	0.004	0.001	0.058	0.000
p-value: AR(2)	0.322	0.040	0.312	0.274	0.065	0.032	0.000
p-value: AR(3)	0.744	0.751	0.268	0.283	0.785	0.873	0.734
p-value: *J*-test	0.019	0.000	0.054	0.011	0.001	0.000	0.000
p-value: *C*-test	0.036	0.046	0.760	0.933	0.001	0.714	0.017

Notes: AR(n) is the Arellano-Bond test for nth order of the autocorrelation. *J*-test is the Hansen test for overidentification. *C*-test is the difference-in-Hansen test which tests the validity of subsets of instruments. Robust standard errors are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

Table 5.8: Imported inputs and exports of Chinese manufacturing firms: GMM estimation (III)

Dependent variable: quality of export varieties								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IMPno	-0.003*							
	(0.001)							
IMPvalue	-0.136***							
	(0.049)							
IMPno_diff		-0.003**		-0.003				
		(0.002)		(0.002)				
IMPno_homo		0.004		0.009				
		(0.006)		(0.007)				
IMPvalue_diff			0.619***	0.447***				
			(0.124)	(0.079)				
IMPvalue_homo			-0.125**	-0.043				
			(0.051)	(0.034)				
IMPno_HI					-0.003		-0.008***	
					(0.002)		(0.002)	
IMPno_nonHI					0.002		-0.007***	
					(0.002)		(0.002)	
IMPvalue_HI						0.007	0.031	
						(0.041)	(0.034)	
IMPvalue_nonHI						0.295***	0.216***	
						(0.055)	(0.043)	
IMPquality								0.336***
								(0.083)

Table 5.8 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
age	-0.232***	-0.069**	-0.114	0.011	-0.083***	-0.103	-0.013	-0.101***
	(0.059)	(0.028)	(0.073)	(0.045)	(0.031)	(0.072)	(0.062)	(0.030)
wage	1.926***	0.827***	1.584**	0.358	1.106***	0.658	0.370	0.977***
	(0.432)	(0.242)	(0.676)	(0.343)	(0.286)	(0.603)	(0.525)	(0.306)
TFP	0.113	0.069	-2.508***	-1.256***	-0.205	-0.796***	-0.137	-0.291*
	(0.187)	(0.144)	(0.543)	(0.361)	(0.171)	(0.239)	(0.197)	(0.163)
SMALL	-0.747	0.058	-2.380**	-0.191	-0.229	0.571	0.844*	0.179
	(0.503)	(0.408)	(1.129)	(0.691)	(0.434)	(0.614)	(0.484)	(0.495)
MEDIUM	0.353	0.612*	-2.806***	-0.917	0.470	0.836*	0.926***	0.522
	(0.379)	(0.321)	(1.015)	(0.618)	(0.322)	(0.443)	(0.343)	(0.371)
FOREIGN	-0.545***	-0.305***	-0.243	-0.296**	-0.417***	-1.008***	-0.426***	-0.322***
	(0.132)	(0.104)	(0.208)	(0.147)	(0.099)	(0.183)	(0.133)	(0.103)
HMT	-0.663***	-0.492***	-0.120	-0.262*	-0.596***	-1.266***	-0.769***	-0.409***
	(0.108)	(0.094)	(0.199)	(0.137)	(0.083)	(0.177)	(0.114)	(0.094)
PRIVATE	-0.344***	-0.263***	0.709***	0.320*	-0.278***	-0.264*	-0.129	-0.108
	(0.109)	(0.102)	(0.264)	(0.166)	(0.092)	(0.142)	(0.116)	(0.097)
COLLECTIVE	-0.344***	-0.336***	0.791**	0.321*	-0.361***	-0.231*	-0.148	-0.034
	(0.118)	(0.112)	(0.308)	(0.191)	(0.101)	(0.138)	(0.117)	(0.118)
Observations	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033
p-value: AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p-value: AR(2)	0.015	0.004	0.381	0.038	0.003	0.017	0.147	0.000
p-value: AR(3)	0.990	0.076	0.034	0.137	0.105	0.053	0.035	0.605
p-value: *J*-test	0.000	0.000	0.015	0.000	0.000	0.000	0.000	0.000
p-value: *C*-test	0.299	0.157	0.861	0.100	0.948	0.220	0.945	0.823

Notes: AR(n) is the Arellano-Bond test for nth order of the autocorrelation. *J*-test is the Hansen test for overidentification. *C*-test is the difference-in-Hansen test which tests the validity of subsets of instruments. Robust standard errors are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

5.5 Robustness Checks

In order to verify the impact of imported inputs on firms' export performance I conduct several robustness checks. Firstly, I do log-linear estimation of the count data by taking logs of the numbers of firms' export and import varieties. Secondly, as it may take time for firms to realize the impact of the imported inputs on export performance and in order to mitigate the simultaneity problem, I run regressions with measures of imported inputs lagged one year. We expect lagged imported inputs sill have a significant impact on current export performance. Thirdly, I estimate the quality of exported varieties using alternative different value of elasticity of substitution. Moreover, I use unit values as proxy of quality. Finally, rather than looking at only the firms that both import and export, I run all regressions using the full sample which contains all firms that either import or export. Results across different specifications show a positive and significant link between imported inputs and firm export performance.

5.5.1 Log linear model of count data

As discussed in Section 5.3, simple linear estimators are not appropriate for count data, a popular practice in the literature is to take logarithms of count variables and do the estimation in a linear setting, see for example Bas and Strauss-Kahn (2014). Taking logs of count variables avoids the restriction that the dependent variable can only be non-negative integer values in linear models. We check whether our results are robust using a log transformation of the count variables. Table 5.9 reports the results of regressions of imported inputs and log of the number of export varieties. Estimates across different specifications show a positive and significant relationship between the use of imported inputs and the number of export varieties. Larger coefficients are found in log-linear models than those obtained in negative binomial models with level count variables in Table 5.3. However, in a log-log model with *log(y)* as dependent variable and *log(x)* as independent variable, the coefficient β of *log(x)* is the estimated elasticity of *y* with respect to *x* (Wooldridge, 2002, p.45). Results in Columns (1) and (2) imply that one percent increase in firms' number of imported input varieties and value of imports is associated with 0.114 and 0.02 percent increase in the number of export varieties respectively. When the variety number and value of imports are included in Column (3), the variety number of imports has a positive and significant effect (0.114 at 1% level) on the number of exported varieties while the value of imports loses the significance. Imported inputs are categorized into homogenous and differentiated inputs in Columns (4) to (6). Coefficients of different groups of inputs are positive and highly significant and the magnitude are very similar as shown in Columns (4) and (5). However, in Column (6), the numbers of imported varieties of two groups of inputs have positive and significant effects on export varieties while the value of both homogenous and differentiated inputs have negative but small impacts. In Columns (7) to (9), inputs are categorized by their origins. Coefficients of inputs imported from high-income economies and other origins are positive and highly significant. The inputs imported from high-income economies have larger impact on exported varieties than that imported from other origins, both for the number of imported varieties and value of imports as in Column (7) and Column (8) respectively. However, in Column (9), the numbers of imported varieties of two groups of inputs have positive and significant effects on export varieties while the value of inputs of these two groups have negative but small effects.

Table 5.9: Export varieties and imported inputs of Chinese manufacturing firms: log-linear model

Dependent variable: log of the count of export varieties									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.114***		0.114***						
	(0.008)		(0.009)						
IMPvalue		0.020***	-0.000						
		(0.003)	(0.003)						
IMPno_diff				0.074***		0.099***			
				(0.009)		(0.015)			
IMPno_homo				0.079***		0.087***			
				(0.008)		(0.009)			
IMPvalue_diff					0.008***	-0.004**			
					(0.001)	(0.002)			
IMPvalue_homo					0.009***	-0.003**			
					(0.001)	(0.002)			
IMPno_HI							0.088***		0.105***
							(0.008)		(0.010)
IMPno_nonHI							0.069***		0.076***
							(0.008)		(0.011)
IMPvalue_HI								0.008***	-0.005***
								(0.001)	(0.002)
IMPvalue_nonHI								0.007***	-0.002
								(0.001)	(0.001)
Constant	1.225***	1.207***	1.228***	1.251***	1.301***	1.279***	1.267***	1.324***	1.290***
	(0.145)	(0.148)	(0.147)	(0.145)	(0.146)	(0.145)	(0.144)	(0.146)	(0.144)
Observations	30,675	30,675	30,675	30,675	30,675	30,675	30,675	30,675	30,675
Number of firms	10,290	10,290	10,290	10,290	10,290	10,290	10,290	10,290	10,290
R-squared	0.130	0.120	0.130	0.132	0.122	0.132	0.133	0.122	0.133

Notes: All independent variables are in natural logs. Firm age, wage, TFP, size, ownership and year dummies are included in all specifications. Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

5.5.2 Lagged measures of imported inputs as independent variables

Since it may take time for firms to realize the impact of the imported inputs on export performance and in order to mitigate the simultaneity problem, I run regressions with all independent variables lagged one year. We expect lagged measures of imported inputs sill have a significant impact on current export performance.

Results are presented in Table 5.10. A positive link between the lagged number and value imported inputs on export value is observed. Lagged number and value of differentiated imported inputs have a positive and significant effect on export value while lagged homogenous imported inputs have no significant impact on export value. Finally, lagged inputs imported from both high-income economies and other countries have a positive and significant effect on export value.

Table 5.10: Export value and imported inputs of Chinese manufacturing firms: lagged independent variables

Dependent variable: log of export value									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.004***		0.003***						
	(0.001)		(0.001)						
IMPvalue		0.027***	0.018*						
		(0.009)	(0.009)						
IMPno_diff				0.003***		0.003***			
				(0.001)		(0.001)			
IMPno_homo				0.005		0.005			
				(0.004)		(0.004)			
IMPvalue_diff					0.014***	0.010**			
					(0.005)	(0.005)			
IMPvalue_homo					0.003	0.000			
					(0.003)	(0.003)			
IMPno_HI							0.003***		0.003***
							(0.001)		(0.001)
IMPno_nonHI							0.005***		0.004***
							(0.002)		(0.002)
IMPvalue_HI								0.009**	0.007*
								(0.004)	(0.004)
IMPvalue_nonHI								0.006*	0.003
								(0.003)	(0.003)
Constant	14.462***	14.198***	14.251***	14.458***	14.364***	14.358***	14.462***	14.402***	14.384***
	(0.413)	(0.419)	(0.419)	(0.332)	(0.336)	(0.336)	(0.332)	(0.334)	(0.334)
Observations	18,948	18,948	18,948	18,948	18,948	18,948	18,948	18,948	18,948
R-squared	0.096	0.095	0.097	0.096	0.095	0.097	0.096	0.095	0.097
Number of firms	7,037	7,037	7,037	7,037	7,037	7,037	7,037	7,037	7,037

Notes: All independent variables lagged one year. Firm age, wage, TFP, size, ownership and year dummies are included in all specifications. Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

5.5.3 Alternative elasticity of substitution for quality estimation

We have shown in Section 5.3.2 that quality can be estimated as the residual of a utility function with a given value of elasticity of substitution between goods. We now test whether the positive and significant link between export quality and imported inputs holds for different values of elasticity of substitution. Broda and Weinstein (2006) report that the mean elasticity of substitution between SITC 5-digit goods is 7.5 for the period between 1972 and 1988 and 6.6 between 1990 and 2001. Rose and van Wincoop (2001) use $\sigma = 5$ for the elasticity of substitution between goods for 143 countries in the study of the impact of currency unions on trade and welfare for 1980 to 1990. We estimate the variety quality with $\sigma = 6$ and run regressions of Equation (5.3) and present the results in Table 5.11. Same as previous tables, Columns (1) to (3) examine the impact of imported varieties or (and) value on export quality. A positive and significant impact of imported inputs on the quality of export varieties is still observed. Inputs are divided into homogenous and differentiated inputs in Columns (4) to (6) and whether imported from high-income economies or other countries in Columns (7) to (9). The relationship between the quality of exported varieties and the quality of imported inputs is examined in Columns (10). Compared with previous results obtained with $\sigma = 4$, the coefficients here are smaller, indicating that with a larger elasticity of substitution between varieties, the magnitude of the impact of imported inputs on the quality of export varieties is smaller.

Table 5.11: Export quality and imported inputs of Chinese manufacturing firms: within model alternative σ

Dependent variable: quality of export varieties										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IMPno	0.0003***		0.0003***							
	(0.000)		(0.000)							
IMPvalue		0.005***	0.004***							
		(0.001)	(0.001)							
IMPno_diff				0.0002***		0.0002***				
				(0.000)		(0.000)				
IMPno_homo				0.002***		0.002***				
				(0.001)		(0.001)				
IMPvalue_diff					0.003***	0.002***				
					(0.000)	(0.000)				
IMPvalue_homo					0.001***	0.001				
					(0.000)	(0.000)				
IMPno_HI							0.0002***		0.0002***	
							(0.000)		(0.000)	
IMPno_nonHI							0.001***		0.001***	
							(0.000)		(0.000)	
IMPvalue_HI								0.002***	0.001***	
								(0.000)	(0.000)	
IMPvalue_nonHI								0.001***	0.001*	
								(0.000)	(0.000)	
IMPquality										0.004***
										(0.001)
Constant	-0.307***	-0.344***	-0.349***	-0.306***	-0.319***	-0.327***	-0.305***	-0.311***	-0.318***	-0.287***
	(0.051)	(0.052)	(0.052)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
Observations	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033
Number of firms	389,842	389,842	389,842	389,842	389,842	389,842	389,842	389,842	389,842	389,842
R-squared	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Notes: Firm age, wage, TFP, size, ownership and year dummies are included in all specifications. Robust standard errors clustered at firm-HS6-country level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

5.5.4 Unit values as proxy of quality

The widespread use of unit values as a proxy for quality is noticed in the international trade literature (see for example, Baldwin and Harrigan, 2011; Hallak, 2006; Manova and Zhang, 2012; Schott, 2004; Swenson and Chen, 2014). We now test whether our findings are robust with unit values as a measure of quality of varieties. Unit values are obtained by dividing the export (import) values by their respective quantities. Results in Table 5.12 show that unit values of exports increase with more varieties of inputs imported or (and) higher unit values of inputs. According to Wooldridge (2002, p.45), in a log-level model with log export unit values as the dependent variable and the level count of imported inputs as dependent variable in Column (1), the coefficient on $IMPno$ can be interpreted as for a unit increase in imported inputs the percentage change in the unit value of the exports is $100 * \beta = 0.01$. Again I estimate the impact of imported inputs on the unit values of export varieties for different groups of inputs by the differentiation and origin of the inputs. Results in Columns (2)-(9) are similar to those with quality estimated by Khandelwal et al. (2013) method in Table 5.5. In Column (10) I test the relationship between the average unit value of imported inputs of firms and unit values of export varieties and find that a ten percent increase in firms' average unit values of imported inputs leads to a 0.01 percent increase in the unit values of their export varieties.

Table 5.12: Export quality and imported inputs of Chinese manufacturing firms: unit values as proxy for quality

Dependent variable: log of unit values of export varieties										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IMPno	0.0001***		0.0001***							
	(0.000)		(0.000)							
IMPvalue		0.002**	0.001*							
		(0.001)	(0.001)							
IMPno_diff				0.001**		0.001**				
				(0.001)		(0.001)				
IMPvalue_homo				0.00003		0.00003				
				(0.000)		(0.000)				
IMPvalue_diff					0.001***	0.001*				
					(0.000)	(0.000)				
IMPvalue_homo					0.001	0.0005				
					(0.000)	(0.000)				
IMPno_HI							0.00007**		0.00006*	
							(0.000)		(0.000)	
IMPno_nonHI							0.0004*		0.0004	
							(0.000)		(0.000)	
IMPvalue_HI								0.001**	0.001*	
								(0.000)	(0.000)	
IMPvalue_nonHI								0.0004	0.0002	
								(0.000)	(0.000)	
IMPuv										0.001**
										(0.001)
Constant	3.752***	3.739***	3.737***	3.753***	3.749***	3.747***	3.753***	3.749***	3.747***	3.752***
	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.049)
Observations	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033	633,033
Number of firms	389,842	389,842	389,842	389,842	389,842	389,842	389,842	389,842	389,842	389,842
R-squared	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Notes: Firm age, wage, TFP, size, ownership and year dummies are included in all specifications. Robust standard errors clustered at the firm-HS6-country level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

5.5.5 Large sample tests

Finally, I examine the link between imported inputs and export performance using a large sample which includes all firms that either import or export. This way I take into account all exporters (importers) independently from their import (export) activities. For firms that do not export (import), the count varieties and values of their exports (imports) are set to zero, and the log values are the natural logarithms of export (import) values plus one.[1] Due to the overdispersion of the count varieties, I use the fixed effects negative binomial estimator for the regressions. Results of AME are presented in Table 5.13 with coefficients presented in Table 5A.3 in the Appendix. Again, the number of varieties and log values of imported inputs are found to have positive and significant effects on firms' export extensive margins. Both homogenous and differentiated inputs have positive impact on firms' number of exported varieties. More varieties of differentiated inputs have larger effects on the number of exported varieties, but larger values of homogenous inputs leads to more exported varieties than that of the differentiated inputs. Also, inputs from both high income countries and other countries lead to the expansion of export varieties and values. However, inputs imported from high income countries have a larger effect on export varieties.

We also look at the impact of imported inputs on the intensive margins of exports, the export values. Given that the data are censored, i.e., the values of dependent variable are the actual values of exports for firms that export while the values of dependent variable are not observed for firms that do not export, I use a Tobit model for the regressions.[2] Results of AME

[1]In this way I include all non-exporters and non-importers for regressions with their log export/import values and their log values are all zero as obtained by log(0+1).

[2]Truncated regressions are not appropriate here as in the truncated case neither the dependent variable nor the explanatory variables are observed for individuals in the truncated region (Baum, 2006, p.262). In our case, I do observe the values of the explanatory variables for non-exporters.

are reported in Table 5.14 with coefficients presented in Table 5A.4 in the Appendix. Similar results of imported inputs on export values are found to those on export varieties in Table 5.13.

Table 5.13: Export varieties and imported inputs: full sample with negative binomial model (AME)

Dependent variable: count of export varieties									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.006***		0.004***						
	(0.000)		(0.000)						
IMPvalue		0.126***	0.124***						
		(0.002)	(0.003)						
IMPno_diff				0.005***		0.005***			
				(0.001)		(0.001)			
IMPno_homo				0.037***		0.030***			
				(0.006)		(0.006)			
IMPvalue_diff					0.079***	0.036			
					(0.025)	(0.026)			
IMPvalue_homo					0.118***	0.079***			
					(0.021)	(0.022)			
IMPno_HI							0.010***		0.008***
							(0.001)		(0.001)
IMPno_nonHI							0.010***		0.010***
							(0.002)		(0.002)
IMPvalue_HI								0.061***	0.021
								(0.022)	(0.023)
IMPvalue_nonHI								0.075***	0.039*
								(0.020)	(0.020)

Table 5.13 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	0.626***	0.614***	0.621***	2.307***	2.282***	2.293***	1.208***	1.940***	1.878***
	(0.034)	(0.037)	(0.038)	(0.239)	(0.260)	(0.264)	(0.151)	(0.230)	(0.231)
wage	0.014	-0.048**	-0.053***	-0.045	-0.089	-0.085	-0.230***	-0.113	-0.108
	(0.019)	(0.020)	(0.020)	(0.078)	(0.084)	(0.084)	(0.059)	(0.078)	(0.078)
TFP	0.083***	0.055***	0.044***	0.063	0.073	0.058	-0.007	0.017	-0.008
	(0.012)	(0.012)	(0.013)	(0.048)	(0.052)	(0.052)	(0.036)	(0.048)	(0.049)
SMALL	0.656***	0.740***	0.810***	0.511**	0.573**	0.647**	1.121***	0.354	0.487**
	(0.083)	(0.086)	(0.086)	(0.261)	(0.276)	(0.276)	(0.208)	(0.240)	(0.241)
MEDIUM	0.901***	0.885***	0.944***	0.988***	0.905***	0.962***	1.445***	0.791***	0.897***
	(0.078)	(0.081)	(0.081)	(0.227)	(0.239)	(0.238)	(0.186)	(0.205)	(0.205)
FOREIGN	2.511***	2.203***	2.196***	0.547	0.344	0.263	1.612***	0.467	0.354
	(0.119)	(0.128)	(0.130)	(0.685)	(0.786)	(0.791)	(0.425)	(0.657)	(0.664)
HMT	2.172***	1.966***	1.970***	0.391	0.265	0.192	1.355***	0.505	0.421
	(0.119)	(0.128)	(0.130)	(0.693)	(0.794)	(0.799)	(0.432)	(0.663)	(0.670)
PRIVATE	0.065	0.066	0.046	-0.253	-0.274	-0.282	0.257	-0.620	-0.630
	(0.110)	(0.119)	(0.121)	(0.653)	(0.752)	(0.757)	(0.405)	(0.629)	(0.637)
COLLECTIVE	0.053	0.096	0.076	-0.157	-0.689	-0.615	1.226**	-0.000	0.115
	(0.131)	(0.141)	(0.143)	(0.888)	(1.055)	(1.062)	(0.528)	(0.939)	(0.949)
Observations	82,109	82,109	82,109	18,595	16,132	16,132	27,258	14,712	14,712

Notes: Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

Table 5.14: Export values and imported inputs of Chinese manufacturing firms: full sample with Tobit model (AME)

Dependent variable: log of export values									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.026***		0.012***						
	(0.001)		(0.001)						
IMPvalue		0.201***	0.190***						
		(0.002)	(0.002)						
IMPno_diff				0.010***		0.009***			
				(0.001)		(0.001)			
IMPno_homo				0.024***		0.032***			
				(0.006)		(0.006)			
IMPvalue_diff					0.037**	-0.025			
					(0.018)	(0.019)			
IMPvalue_homo					-0.028*	-0.064***			
					(0.015)	(0.016)			
IMPno_HI							0.014***		0.013***
							(0.001)		(0.001)
IMPno_nonHI							0.014***		0.016***
							(0.002)		(0.003)
IMPvalue_HI								-0.031*	-0.091***
								(0.018)	(0.019)
IMPvalue_nonHI								-0.019	-0.062***
								(0.016)	(0.017)

Table 5.14 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	0.126***	0.094***	0.095***	0.586***	0.742***	0.730***	0.338***	0.633***	0.610***
	(0.021)	(0.020)	(0.020)	(0.125)	(0.137)	(0.136)	(0.100)	(0.135)	(0.134)
wage	0.058***	0.023**	0.017	-0.130**	-0.116*	-0.143**	-0.234***	-0.055	-0.089
	(0.011)	(0.011)	(0.011)	(0.060)	(0.065)	(0.065)	(0.051)	(0.066)	(0.066)
TFP	0.126***	0.102***	0.096***	0.057	0.135***	0.099**	0.023	0.112***	0.070*
	(0.007)	(0.007)	(0.007)	(0.036)	(0.039)	(0.039)	(0.031)	(0.040)	(0.040)
SMALL	-1.144***	-1.031***	-0.863***	-0.650***	-0.899***	-0.599***	-0.351*	-0.928***	-0.566***
	(0.076)	(0.074)	(0.074)	(0.215)	(0.223)	(0.224)	(0.180)	(0.211)	(0.212)
MEDIUM	-0.352***	-0.354***	-0.210***	0.027	-0.283	-0.041	0.116	-0.230	0.048
	(0.073)	(0.072)	(0.072)	(0.196)	(0.202)	(0.203)	(0.165)	(0.188)	(0.188)
FOREIGN	3.531***	2.967***	2.902***	-0.591	-0.115	-0.358	0.218	0.051	-0.272
	(0.057)	(0.055)	(0.055)	(0.409)	(0.450)	(0.449)	(0.288)	(0.419)	(0.417)
HMT	2.883***	2.436***	2.404***	-1.359***	-0.835*	-1.031**	-0.469	-0.574	-0.834**
	(0.056)	(0.055)	(0.055)	(0.412)	(0.454)	(0.452)	(0.291)	(0.422)	(0.420)
PRIVATE	0.474***	0.443***	0.442***	0.409	0.609	0.588	0.952***	0.778*	0.750*
	(0.048)	(0.047)	(0.047)	(0.410)	(0.454)	(0.452)	(0.283)	(0.420)	(0.418)
COLLECTIVE	0.272***	0.252***	0.250***	0.506	0.154	0.141	1.146***	0.301	0.281
	(0.052)	(0.050)	(0.050)	(0.564)	(0.673)	(0.671)	(0.387)	(0.623)	(0.619)
Observations	254,416	254,416	254,416	23,923	20,341	20,341	34,492	19,096	19,096

Notes: Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

5.6 Conclusions

Compared to the extensive studies on different aspects of firm-level exports, import activities, the other important side of international trade, has received far less attention in the literature. Little has been done to explore the link between imports and exports. In this chapter I have examined various aspects of the relationship between imported inputs and export performance. Using firm production survey data and transaction-level international trade data of Chinese manufacturing firms over the period of 2002 to 2006, I observe a positive and significant link between imported inputs and export performance.

Our results across different specifications show that importing more varieties and larger value of inputs promotes firms' exports, including the extensive and intensive margins of exports. Imported inputs of differentiated products and inputs imported from high-income economies are found to have stronger impacts on export performance. Quality upgrading of export varieties is also observed with the increase of imported inputs varieties, values and average quality. We address the possible endogeneity of imports by employing GMM estimations. We also perform several robustness checks using alternative methods, different measures of quality or larger samples.

Our findings have potentially important policy implications. Since the growth of the Chinese economy over the past two decades is largely driven by its exports and the use of imported inputs improves firms' export performance, policy makers should consider promote imports alongside continued export promotion. Chinese manufacturing firms would benefit from further trade liberalization. Examples include lowering barriers to importing by providing more information on sources of intermediate inputs, lowering tariffs and arranging for such firms to attend overseas trade fairs and similar events and possible support mechanisms that policy makers could employ. Recent senior trade delegations from China to the West and the subsequent highly publicised deals make it clear that China is interested in importing high quality intermediate inputs and attracting further foreign direct investment. Support from central and local governments in lowering the substantial costs of importing would be beneficial to potential firms to participate in the imports market. Firms should also be encouraged to import more differentiated inputs and source inputs from high-income economies as these inputs are found to have stronger effects in promoting exports.

Appendix 5A

Table 5A.1: Definition of variables

Variable	Definition
EXPno	count of export varieties
EXPvalue	natural log of export value at firm level
IMPno	count of import varieties
IMPno_diff	count of import differentiated varieties
IMPno_home	count of import homogenous varieties
IMPno_HI	count of varieties imported from high-income economies
IMPno_nonHI	count of varieties imported from non high-income economies
IMPlue	natural log of import value at firm level
IMPvalue	natural log of import value at firm level
IMPvalue_diff	natural log of import value of differentiated goods
IMPvalue_home	natural log of import value of homogenous goods
IMPvalue_HI	natural log of inputs imported from high-income economies
IMPvalue_nonHI	natural log of inputs imported from non high-income economies
EXPquality	quality of export varieties
IMPquality	average quality of imports at firm-level
EXPuv	unit values of export varieties
IMPuv	average unit values of imports at firm-level
age	log of a firm's age (the report year minus the founded year of a firm)
wage	log of average wage of employees of a firm (total wage bill in 1,000RMB divided by the number of employees)
TFP	total factor productivity of a firm estimated by the method of De Leocker(2010)
SMALL	a size dummy which equals 1 if a firm is a small firm and 0 otherwise.
MEDIUM	a size dummy which equals 1 if a firm is a medium firm and 0 otherwise.
LARGE	a size dummy which equals 1 if a firm is a large firm and 0 otherwise.

Table 5A.2: Export varieties and imported inputs: negative binomial model

Dependent variable: count of export varieties									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.001***		0.001***						
	(0.000)		(0.000)						
IMPvalue		0.011***	0.007***						
		(0.002)	(0.002)						
IMPno_diff				0.001***		0.001***			
				(0.000)		(0.000)			
IMPno_homo				0.005***		0.004***			
				(0.001)		(0.001)			
IMPvalue_diff					0.001***	0.003			
					(0.000)	(0.004)			
IMPvalue_homo					0.005***	0.012***			
					(0.001)	(0.003)			
IMPno_HI							0.001***		0.001***
							(0.000)		(0.000)
IMPno_nonHI							0.001***		0.001***
							(0.000)		(0.000)
IMPvalue_HI								0.008**	0.003
								(0.004)	(0.004)
IMPvalue_nonHI								0.013***	0.008**
								(0.003)	(0.003)

Table 5A.2 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	0.193***	0.198***	0.192***	0.277***	0.199***	0.265***	0.177***	0.275***	0.264***
	(0.023)	(0.023)	(0.023)	(0.035)	(0.023)	(0.038)	(0.026)	(0.038)	(0.038)
wage	-0.015*	-0.017**	-0.017*	-0.005	-0.015	-0.010	-0.028***	-0.024*	-0.023*
	(0.009)	(0.009)	(0.009)	(0.012)	(0.009)	(0.012)	(0.010)	(0.013)	(0.013)
TFP	0.012**	0.013**	0.010*	0.007	0.009	0.006	0.007	0.001	-0.001
	(0.005)	(0.005)	(0.005)	(0.007)	(0.006)	(0.008)	(0.006)	(0.008)	(0.008)
SMALL	0.120***	0.114***	0.124***	0.065*	0.118***	0.088**	0.142***	0.098**	0.109***
	(0.032)	(0.032)	(0.032)	(0.039)	(0.032)	(0.041)	(0.034)	(0.040)	(0.040)
MEDIUM	0.178***	0.172***	0.180***	0.114***	0.176***	0.122***	0.193***	0.147***	0.154***
	(0.029)	(0.029)	(0.028)	(0.034)	(0.029)	(0.035)	(0.030)	(0.034)	(0.034)
FOREIGN	0.165**	0.183***	0.165**	-0.037	0.169**	-0.023	0.100	-0.117	-0.130
	(0.067)	(0.067)	(0.067)	(0.103)	(0.068)	(0.117)	(0.072)	(0.110)	(0.110)
HMT	0.182***	0.198***	0.183***	-0.029	0.190***	-0.014	0.102	-0.092	-0.101
	(0.068)	(0.068)	(0.067)	(0.104)	(0.068)	(0.118)	(0.073)	(0.111)	(0.111)
PRIVATE	-0.058	-0.049	-0.056	-0.151	-0.043	-0.107	-0.083	-0.206**	-0.204*
	(0.064)	(0.064)	(0.064)	(0.097)	(0.064)	(0.111)	(0.068)	(0.105)	(0.105)
COLLECTIVE	0.065	0.074	0.068	-0.159	0.067	-0.147	0.032	-0.080	-0.065
	(0.080)	(0.080)	(0.080)	(0.132)	(0.081)	(0.154)	(0.088)	(0.152)	(0.152)
Constant	1.370***	1.238***	1.303***	1.461***	1.386***	1.292***	1.535***	1.265***	1.396***
	(0.107)	(0.109)	(0.109)	(0.158)	(0.109)	(0.181)	(0.117)	(0.176)	(0.177)
Observations	28,942	28,942	28,942	28,942	28,942	28,942	28,942	28,942	28,942
log likelihood	-59,651	-59,692	-59,647	-31,921	-57,419	-28,226	-48,540	-27,014	-26,982

Notes: Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

Table 5A.3: Export varieties and imported inputs: full sample with negative binomial model

Dependent variable: count of export varieties									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.003***		0.002***						
	(0.000)		(0.000)						
IMPvalue		0.049***	0.047***						
		(0.001)	(0.001)						
IMPno_diff				0.001***		0.001***			
				(0.000)		(0.000)			
IMPno_homo				0.006***		0.005***			
				(0.001)		(0.001)			
IMPvalue_diff					0.013***	0.006			
					(0.004)	(0.004)			
IMPvalue_homo					0.019***	0.012***			
					(0.003)	(0.003)			
IMPno_HI							0.002***		0.001***
							(0.000)		(0.000)
IMPno_nonHI							0.002***		0.002***
							(0.000)		(0.000)
IMPvalue_HI								0.011***	0.004
								(0.004)	(0.004)
IMPvalue_nonHI								0.013***	0.007*
								(0.003)	(0.004)

Table 5A.3 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	0.261***	0.236***	0.235***	0.369***	0.361***	0.359***	0.215***	0.346***	0.332***
	(0.014)	(0.014)	(0.014)	(0.036)	(0.039)	(0.039)	(0.026)	(0.039)	(0.039)
wage	0.006	-0.018**	-0.020***	-0.007	-0.014	-0.013	-0.041***	-0.020	-0.019
	(0.008)	(0.008)	(0.008)	(0.012)	(0.013)	(0.013)	(0.010)	(0.014)	(0.014)
TFP	0.034***	0.021***	0.017***	0.010	0.012	0.009	-0.001	0.003	-0.001
	(0.005)	(0.005)	(0.005)	(0.008)	(0.008)	(0.008)	(0.006)	(0.009)	(0.009)
SMALL	0.273***	0.284***	0.306***	0.082**	0.091**	0.101**	0.199***	0.063	0.086**
	(0.034)	(0.033)	(0.033)	(0.042)	(0.043)	(0.043)	(0.037)	(0.043)	(0.043)
MEDIUM	0.376***	0.340***	0.357***	0.158***	0.143***	0.151***	0.257***	0.141***	0.159***
	(0.032)	(0.031)	(0.030)	(0.036)	(0.038)	(0.037)	(0.033)	(0.036)	(0.036)
FOREIGN	1.047***	0.846***	0.831***	0.088	0.054	0.041	0.287***	0.083	0.063
	(0.047)	(0.047)	(0.047)	(0.110)	(0.124)	(0.124)	(0.075)	(0.117)	(0.117)
HMT	0.906***	0.755***	0.745***	0.063	0.042	0.030	0.241***	0.090	0.075
	(0.048)	(0.048)	(0.048)	(0.111)	(0.125)	(0.125)	(0.076)	(0.118)	(0.119)
PRIVATE	0.027	0.025	0.017	-0.040	-0.043	-0.044	0.046	-0.111	-0.112
	(0.046)	(0.046)	(0.046)	(0.105)	(0.119)	(0.119)	(0.072)	(0.112)	(0.113)
COLLECTIVE	0.022	0.037	0.029	-0.025	-0.109	-0.096	0.218**	-0.000	0.020
	(0.055)	(0.054)	(0.054)	(0.142)	(0.167)	(0.166)	(0.094)	(0.168)	(0.168)
Constant	-0.891***	-0.777***	-0.753***	0.753***	0.484**	0.623***	0.955***	0.547***	0.738***
	(0.078)	(0.077)	(0.077)	(0.166)	(0.189)	(0.190)	(0.122)	(0.186)	(0.187)
Observations	82,109	82,109	82,109	18,595	16,132	16,132	27,258	14,712	14,712
log likelihood	-164,727	-163,043	-162,946	-35,507	-31,233	-31,196	-54,034	-29,796	-29,740

Notes: Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

Table 5A.4: Export values and imported inputs of Chinese manufacturing firms: full sample with Tobit model

Dependent variable: log of export values									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IMPno	0.036***		0.017***						
	(0.001)		(0.001)						
IMPvalue		0.278***	0.264***						
		(0.002)	(0.002)						
IMPno_diff				0.010***		0.010***			
				(0.001)		(0.001)			
IMPno_homo				0.026***		0.033***			
				(0.006)		(0.006)			
IMPvalue_diff					0.039**	-0.026			
					(0.019)	(0.019)			
IMPvalue_homo					-0.029*	-0.067***			
					(0.016)	(0.017)			
IMPno_HI							0.015***		0.014***
							(0.001)		(0.001)
IMPno_nonHI							0.015***		0.017***
							(0.003)		(0.003)
IMPvalue_HI								-0.033*	-0.095***
								(0.019)	(0.019)
IMPvalue_nonHI								-0.020	-0.065***
								(0.017)	(0.018)

Table 5A.4 (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age	0.175***	0.131***	0.132***	0.616***	0.776***	0.763***	0.356***	0.662***	0.637***
	(0.029)	(0.028)	(0.028)	(0.131)	(0.144)	(0.143)	(0.105)	(0.141)	(0.139)
wage	0.080***	0.032**	0.024	-0.137**	-0.121*	-0.149**	-0.247***	-0.058	-0.093
	(0.016)	(0.015)	(0.015)	(0.063)	(0.068)	(0.068)	(0.054)	(0.069)	(0.069)
TFP	0.176***	0.142***	0.133***	0.060	0.141***	0.104**	0.024	0.117***	0.073*
	(0.010)	(0.010)	(0.010)	(0.037)	(0.041)	(0.041)	(0.032)	(0.042)	(0.042)
SMALL	-1.590***	-1.430***	-1.198***	-0.683***	-0.940***	-0.626***	-0.369*	-0.971***	-0.591***
	(0.105)	(0.103)	(0.103)	(0.226)	(0.233)	(0.234)	(0.190)	(0.221)	(0.221)
MEDIUM	-0.489***	-0.491***	-0.291***	0.028	-0.296	-0.043	0.122	-0.241	0.050
	(0.102)	(0.099)	(0.100)	(0.206)	(0.211)	(0.212)	(0.174)	(0.197)	(0.197)
FOREIGN	4.909***	4.118***	4.031***	-0.621	-0.120	-0.374	0.230	0.053	-0.284
	(0.079)	(0.077)	(0.077)	(0.430)	(0.471)	(0.469)	(0.303)	(0.438)	(0.435)
HMT	4.008***	3.381***	3.339***	-1.428***	-0.874*	-1.077**	-0.494	-0.600	-0.871**
	(0.078)	(0.076)	(0.076)	(0.433)	(0.475)	(0.473)	(0.306)	(0.442)	(0.439)
PRIVATE	0.659***	0.615***	0.614***	0.430	0.637	0.615	1.002***	0.814*	0.784*
	(0.067)	(0.065)	(0.065)	(0.431)	(0.475)	(0.473)	(0.298)	(0.439)	(0.436)
COLLECTIVE	0.378***	0.350***	0.347***	0.531	0.161	0.147	1.207***	0.315	0.293
	(0.072)	(0.070)	(0.070)	(0.593)	(0.704)	(0.701)	(0.407)	(0.651)	(0.646)
Constant	1.846***	1.927***	1.774***	11.390***	10.597***	11.803***	11.423***	11.504***	12.898***
	(0.167)	(0.163)	(0.163)	(0.687)	(0.780)	(0.782)	(0.535)	(0.756)	(0.757)
Observations	254,416	254,416	254,416	254,416	254,416	254,416	254,416	254,416	254,416
log likelihood	-696,733	-691,387	-691,119	-71,180	-60,243	-60,170	-102,561	-56,861	-56,760

Notes: Robust standard errors clustered at the firm level are reported in parentheses. ***, ** and * denote significance levels at 0.01, 0.5 and 0.1 respectively.

Chapter 6

Conclusions

This chapter concludes the thesis. Firstly, I summarize the empirical results and policy implications presented in Chapters 3 to 5 in Section 6.1. Section 6.2 then discusses the limitations of the research and proposes potential future research related to the topics studied in the thesis.

6.1 Summary of Results

This thesis consists of three main empirical studies on firm-level exporting, importing and heterogeneity. Annual survey data of Chinese industrial enterprises and customs data of China for the period 2002 to 2006 are used for the studies. In order to investigate the relationships between firms' exporting / importing behavior and firms' characteristics, I match survey data and customs data. Introduction of the two datasets and the method of data construction are presented in Chapter Two, followed by the three empirical studies.

Chapter Three studies the dynamics of the decision to export of Chinese manufacturing firms. Sunk entry costs, firm heterogeneity and the initial conditions problem are investigated and dynamic random effects Probit models are used to examine firms' propensity to export. It is one of the first few papers that takes into account firms' import activity when studying their decision to export. Results show that exporting involves large sunk entry costs and firms that have past export experiences have a larger propensity to export in the current period. Firms' participation in imports market has a positive and significant impact on their decision to export. Unobserved firm heterogeneity and the initial conditions problem should be properly addressed, otherwise the size of the sunk costs would be overstated. We apply Wooldridge (2005) approach to address both issues.

Chapter Four examines the relationship between importing and productivity of Chinese manufacturing firms. Propensity score matching and matching difference-in-differences techniques are employed to test the hypotheses of self-selection into importing and learning-from-importing. A bi-directory causality is observed: more productive firms self-select into the import market and after import-market entry the average firm experiences productivity gains. Results show that compared to the matched non-importing firms, indigenous import starters display stronger learning effects from importing than entrants that are owned by foreign investors. We also find that small and medium sized import starter experience higher productivity growth in comparison to the large new importers. Import starters that source their inputs from high income economies have larger productivity gains than those that start to import only from non-high income economies. In terms of self-selection, it is small private firms and foreign firms who appear to increase productivity prior to import entry.

In Chapter Five, I explore the connections between the firms' exporting and importing activities in more depth. Results across different specifications show that the use of imported inputs has a positive and significant effect on firms' export performance: firms importing more inputs raise the number of export varieties and the value of exports and firms upgrade their export quality with better quality of imported inputs. The impacts of imported inputs on firms' export performance are larger for those differentiated inputs and imported from high-income economies than the homogenous inputs and those imported form other countries.

The results of the above empirical studies have potentially important policy implications. Firstly, results in Chapter Three show that exporting incurs large sunk entry costs, policies aiming to assist firms in overcoming barriers to exporting, such as providing information about potential markets and providing financial support to facilitate firms' the entry costs, should be effective in promoting more export entrants. Secondly, results from all three studies show that firms' importing activity should not be neglected as imports market participation increases the propensity of exporting, improves firm productivity and raises firms' export scopes and quality. As Chinese manufacturing firms appear to benefit from importing, China could derive additional productivity gains from further trade liberalization and might want to consider promoting trade liberalisation alongside continued export promotion. Examples include simplifying the regulation and procedure of importing and lowering barriers to importing by providing more information on sources of intermediate inputs, lowering tariffs and arranging for such firms to attend overseas trade fairs and similar events and possible support mechanisms that policy makers could employ. Chinese manufacturing firms will benefit from further trade liberalization on imports market.

6.2 Limitations and Future Research

The limitations of the thesis are mainly due to the fact that the panel data used are of relatively short period. In Chapter Three I apply Wooldridge (2005) method to address the initial conditions problem which uses within-means of time-varying explanatory variables. However, Akay (2012) argues that the Wooldridge (2005)'s solution to the initial conditions problem works very well only for the panels of moderately long duration (longer than 5 to 8 periods) and Rabe-Hesketh and Skrondal (2013) also show that the Wooldridge method performs poorly for short panels (up to 5 periods). Alternatively, Akay (2012) suggests Heckman's reduced-form approximation for short panels (shorter than 5 periods). The drawbacks of Heckman estimator are that it requires separate programming and it is computationally cumbersome. We have used a high performance computing cluster for estimations, but the regressions still failed to converge even after a very lengthy time period. Hence, I need a longer period of data in order to get more accurate estimates using Wooldridge (2005)'s solution.

In Chapter Five, I use two-step systems GMM to address the endogeneity issues. The GMM estimator performances well when I test the impact of imported inputs on the number of export varieties. However, Hansen and difference-in Sargan tests failed when I moved to look at that on export value and quality. The reason may be the instruments, the third and fourth lags of explanatory variables, are not good instruments as they may still be correlated with the endogenous variable. Deeper lags of the dependent variable can be used as additional instruments to improve efficiency because they introduce more information. So if I have a longer period of data, I can do GMM estimation with deeper lags as instruments.

To improve the thesis, I may construct a longer period of data from both sources by adding data of either earlier years before 2002 or more recent years after 2006. With data before 2002, I can compare the Chinese manufacturing firms' performance in international trade before and after its entry into WTO, while with data after 2006 till 2013 I can study how Chinese firms are affected in the international markets by the recent economic crisis and how they react to it after that.

Given the rich information contained in the data, the future research can be carried out on topics related to the impact of trade liberalization, such as tariff reduction, on exporters and importers' performance, the environmental issues with international trade and the complex trade network at firm level.

Bibliography

Acharya, R. C. and W. Keller (2009). Technology transfer through imports. *Canadian Journal of Economics 42(4)*, 1411–1448. 75, 116

Ahn, J., A. Khandelwal, and S.-J. Wei (2011). The Role of Intermediaries in Facilitating Trade. *Journal of International Economics 84(4)*, 73–85. 44

Aitken, B., G. H. Hanson, and A. E. Harrison (1997). Spillovers, foreign investment, and export behavior. *Canadian Journal of Economics 43(1)*, 103–132. 34

Akay, A. (2012). Finite-sample comparison of alternative methods for estimating dynamic panel data models. *Journal of Applied Econometrics 27(7)*, 1189–1204. 185

Albarran, P., R. Carrasco, and A. Holl (2013). Domestic transport infrastructure and firmsŠ export market participation. *Small Business Economics 40(4)*, 879–898. 26, 29, 33

Alessandria, G. and H. Choi (2007). Do Sunk Costs of Exporting Matter for Net Export Dynamics? *The Quarterly Journal of Economics 122(1)*, 289–336. 33

Alvarez, R. and R. Lopez (2005). Exporting and performance: evidence from Chilean plants. *Canadian Journal of Economics 38 (4)*, 1384 –1400. 26

Alvarez, R. and R. A. Lopez (2008). Entry and Exit in International Markets: Evidence from Chilean Data. *Review of International Economics 16(4)*, 692–708. 31

Amiti, M. and A. K. Khandelwal (2013). Import Competition and Quality Upgrading. *The Review of Economics and Statistics 92(2)*, 476–490. 75, 116, 138, 139

Amiti, M. and J. Konings (2007). Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. *The American Economic Review 97(5)*, 1611–1638. 17, 27, 75, 121

Anderson, J. E. and E. van Wincoop (2004). Trade Costs. *Journal of Economic Literature 42(3)*, 691–751. 139

Anderson, T. and C. Hsiao (1982). Formulation and Estimation of Dynamic Models Using Panel Data. *Journal of Econometrics 18(1)*, 47–82. 145, 146

Andersson, M., H. Loof, and S. Johansson (2008). Productivity and International Trade-Firm Level Evidence from a Small Open Economy. *Review of World Economics 144(4)*, 774–801. 72, 74, 85

Arellano, M. and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies 58(2)*, 277–297. 146, 147

Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics 68*, 29–51. 146

Aristei, D., D. Castellani, and C. Franco (2013). Firms' exporting and importing activities: is there a two-way relationship? *Review of World Economics 149(1)*, 55–84. 27, 28, 72, 116

Arkolakis, C., S. Demidova, P. J. Klenow, and A. Rodriguez-Clare (2008). Endogenous Variety and the Gains from Trade. *American Economic Review 98(2)*, 444–450. 122

Arnold, J. M. and K. Hussinger (2005). Export Behavior and Firm Productivity in German Manufacturing: A Firm-Level Analysis. *Review of World Economics 141(2)*, 219–243. 14, 29, 71, 115, 120

Arulampalam, W. and M. B. Stewart (2009). Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators. *Oxford Bulletin of Economics and Statistics 71(5)*, 659–681. 41, 42

Augier, P., O. Cadot, and M. Dovis (2013). Imports and TFP at the firm level: the role of absorptive capacity. *Canadian Journal of Economics 46(3)*, 956–981. 27, 72, 75, 83, 109, 116

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine 28(25)*, 3083–3107. 81, 82, 83

Baldwin, J. R. and W. Gu (2004). Participation in Export Markets and Productivity Performance in Canadian Manufacturing. *Canadian Journal of Economics 36 (3)*, 634–657. 71, 115

Baldwin, R. and J. Harrigan (2011). Zeros, Quality, and Space: Trade Theory and Trade Evidence. *American Economic Journal: Microeconomics 3(2)*, 60–88. 137, 165

Bas, M. (2012). Input-trade Liberalization and Firm Export Decisions: Evidence from Argentina. *Journal of Development Economics 97(2)*, 481–493. 27

Bas, M. and V. Strauss-Kahn (2014). Does importing more inputs raise exports? Firm-level evidence from France. *Review of World Economics 150(2)*, 241–275. 17, 27, 75, 116, 117, 121, 129, 130, 158

Baum, C. F. (2006). *An Introduction to Modern Econometrics Using Stata*. StataCorp LP, College Station, Texas: Stata Press. 167

Baum, F., M. E. Schaffer, and S. Stillman (2003). Instrumental variables and GMM: Estimation and testing. *The Stata Journal 3(1)*, 1–31. 144, 147

Bellone, F., P. Musso, L. Nesta, and S. Schiavo (2010). Financial Constraints and Firm Export Behaviour. *The World Economy 33 (3)*, 347–373. 26, 29, 33, 50

Berman, N., P. Martin, and T. Mayer (2012). How do Different Exporters React to Exchange Rate Changes? *The Quarterly Journal of Economics 127 (1)*, 437–492. 115

Bernard, A. B., J. B.Jensen, and P. K. Schott (2005). Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods. *NBER Working Paper No. 11404*. 72, 116

Bernard, A. B. and J. B. Jensen (2004). Why Some Firms Export. *Review of Economics and Statistics 86(2)*, 561–569. 14, 26, 28, 30, 31, 33, 37, 115

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007). Firms in International Trade. *Journal of Economic Perspectives 21(3)*, 105–130. 72

Bernard, A. B. and J. Wagner (2001). Export Entry and Exit by German Firms. *Review of World Economics 137(1)*, 105–123. 27, 29

Blalock, G. and F. Veloso (2007). Imports, Productivity Growth, and Supply Chain Learning. *World Development 35 (7)*, 1134–1151. 74, 116

Bloch, H. and J. T. McDonald (2001). Import Competition and Labor Productivity. *Journal of Industry, Competition and Trade 1*(3), 301–319. 100

Blundell, R. and S. Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics 87(1)*, 115–143. 146

Blundell, R. and M. Costa Dias (2000). Evaluation Methods for Non-Experimental Data. *Fiscal Studies 21(4)*, 427–468. 81

Brandt, L., J. V. Biesebroeck, and Y. Zhang (2012). Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing. *Journal of Development Economics 97(2)*, 339–351. 14, 15, 44, 120

Broda, C., J. Greenfield, and D. Weinstein (2006). From Groundnuts to Globalization: A Structural Estimate of Trade and Growth. *NBER Working Paper No. 12512*. 139

Broda, C. and D. E. Weinstein (2006). Globalization and the Gains from Variety. *The Quarterly Journal of Economics 121(2)*, 541–585. 122, 163

Brown, J. R. and B. C. Petersen (2009). Why has the investment-cash flow sensitivity declined so sharply? Rising R&D and equity market developments. *Journal of Banking & Finance 33(5)*, 971–84. 147

Buck, T., I. Filatotchev, N. demina, and M. Wright (2010). Exporting Activity in Transitional Economies: An Enterprise-Level Study. *Journal of Development Studies 37(2)*, 44–66. 32, 33

Bugamelli, M. and L. Infante (2003). Sunk Costs to Exports: A Role for Industrial Districts? *Local Economies and Internationalization in Italy Conference*, 343–372. 31, 34

Bustos, P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms. *American Economic Review 101(1)*, 304–40. 115

Cadot, O., C. Carrere, and V. Strauss-Kahn (2014). OECD imports: diversification of suppliers and quality search. *Review of World Economics 150*, 1–24. 116, 137

Caggese, A. and V. Cunatc (2013). Financing constraints, firm dynamics, export decisions, and aggregate productivity. *Review of Economic Dynamics 16(1)*, 177–193. 33

Caliendo, M. and S. Kopeinig (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys 22(1)*, 31–72. 79, 81

Cameron, A. C. and P. K. Trivedi (2005). *MICROECONOMETRICS: Methods and Applications* . New York: Cambridge University Press. 126

Cameron, A. C. and P. K. Trivedi (2010). *Microeconometrics Using Stata* (Revised ed.). Stata Press. 127, 128

Campa, J. M. (2004). Exchange rates and trade: How important is hysteresis in trade? *European Economic Review 48(3)*, 527–548. 29, 33

Castellani, D., F. Serti, and C. Tomasi (2010). Firms in International Trade: Importers and Exporters Heterogeneity in Italian Manufacturing Industry. *The World Economy 33(3)*, 424–457. 47, 71, 74, 89, 93, 116

Chamberlain, G. (1984). Panel Data. In Z. Griliches and M. Intrilligator (Eds.), *Handbook of Econometrics* (1 ed.), Volume 2, Chapter 22, pp. 1247–318. Amsterdam, North-Holland: Elsevier. 40

Chaney, T. (2013). Liquidity Constrained Exporters. *NBER Working Paper No. 19170*. 33, 50

Chen, B. and H. Ma (2012). Import Variety and Welfare Gain in China. *Review of International Economics 20(4)*, 807–820. 116, 122

Colantonea, I. and R. Crino (2014). New Imported Inputs, New Domestic Products. *Journal of International Economics 92(1)*, 147–165. 27

Cole, M. A., R. J. Elliott, and S. Virakul (2010). Firm Heterogeneity, Origin of Ownership and Export Participation. *The World Economy 33(2)*, 264–291. 32

Dai, M. and M. Yu (2013). Firm R&D, Absorptive Capacity and Learning by Exporting: Firm-level Evidence from China. *The World Economy 26(9)*, 1131–1145. 29

Das, S., M. J. Roberts, and J. R. Tybout (2007). Market Entry Costs, Producer Heterogeneity, and Exporting Dynamics. *Econometrica 75(3)*, 837–873. 20, 26, 27, 29, 31, 33, 34

De Loecker, J. (2007). Do Exports Generate Higher Productivity? Evidence from Slovenia. *Journal of International Economics 73(1)*, 69–98. 17, 20, 44, 69, 73, 78, 115, 120

Dixon, R., A. Guariglia, and R. Vijayakumaran (2015). Managerial ownership, corporate governance and firms' exporting decisions: evidence from Chinese listed companies. *The European Journal of Finance, forthcoming*. 33

Drakos, K. and P. T. Konstantinou (2013). Investment decisions in manufacturing: assessing the effects of real oil prices and their uncertainty. *Journal of Applied Econometrics 28(1)*, 151–165. 52

Du, J. and S. Girma (2007). Finance and Firm Export in China. *Kyklos 60(1)*, 37–54. 71

Egger, P. and T. Url (2006). Public Export Credit Guarantees and Foreign Trade Structure: Evidence from Austria. *World Economy 29 (4)*, 399–418. 52

Elliott, R. and Y. Zhou (2013). State-owned Enterprises, Exporting and Productivity in China: A Stochastic Dominance Approach. *World Economy 36(8)*, 1000–Ű1028. 14

Fan, H., L. Y. A., and Y. S. R. (2015). Trade Liberalization, Quality, and Export Prices. *Review of Economics and Statistics, forthcoming*. 121, 151

Feng, L., Z. Li, and D. L. Swenson (2012). The Connection between Imported Intermediate Inputs and Exports: Evidence from Chinese Firms. *NBER Working Paper No. 18260*. 29, 72, 116, 117, 121

Filatotchev, I., N. Demina, M. Wright, and T. Buck (2001). Strategic choices, export orientation and corporate governance in privatized firms in Russia, Ukraine and Belarus: Theory and empirical evidence. *Journal of International Business Studies 32(4)*, 853–871. 33

Filatotchev, I., N. Isachenkova, and T. Mickiewicz (2007). Corporate Governance, Managers' Independence, Exporting, and Performance of Firms in Transition Economies. *Emerging Markets Finance & Trade 43(5)*, 62–77. 33

Fugazza, M. and A. McLaren (2014). Market Access, Export Performance and Survival: Evidence from Peruvian Firms. *Review of International Economics 22(3)*, 599–624. 27, 52

Galdon-Sanchez, J. E. and J. A. J. Schmitz (2002). Competitive Pressure and Labor Productivity: World Iron-Ore Markets in the 1980's. *The American Economic Review 92*(4), 1222–1235. 100

Garrido, M. M., A. S. Kelley, J. Paris, K. Roza, D. E. Meier, R. S. Morrison, and M. D. Aldridge (2014). Methods for Constructing and Assessing Propensity Scores. *Health Services Research 49(5)*, 1701–1720. 78, 79, 82, 83

Girma, S., D. Greenaway, and R. Kneller (2004). Does Exporting Increase Productivity? A Microeconometric Analysis of Matched Firms. *Review of International Economics 12(5)*, 855–866. 14, 71, 80, 115, 120

Girma, S., R. Kneller, and M. Pisu (2007). Do exporters have anything to learn from foreign multinationals? *European Economic Review 51(4)*, 993–1010. 78

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2009). Trade Liberalization and New Imported Inputs. *American Economic Review: Papers and Proceedings 99(2)*, 494–500. 74

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010). Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. *The Quarterly Journal of Economics 125 (4)*, 1727–1767. 27, 75, 116, 121

Gorg, H., M. Henry, and E. Strobl (2008). Grant Support and Exporting Activity. *The Review of Economics and Statistics 90(1)*, 168–174. 33

Greenaway, D., A. Guariglia, and R. Kneller (2007). Financial Factors and Exporting Decisions. *Journal of International Economics 73(2)*, 377–395. 26, 29, 32, 33, 37

Greenaway, D. and R. Kneller (2004). Exporting and Productivity in the United Kingdom. *Oxford Review of Economic Policy 20(3)*, 325–371. 31

Greenaway, D. and R. Kneller (2008). Exporting, Productivity and Agglomeration. *European Economic Review 52(5)*, 919–939. 26, 34

Hallak, J. C. (2006). Product quality and the direction of trade. *Journal of International Economics 68(1)*, 238–265. 137, 165

Haller, S. A. (2012). Intra-firm trade, exporting, importing, and firm performance. *Canadian Journal of Economics 45(4)*, 1397–1430. 47

Halpern, L., M. Koren, and A. Szeidl (2005). Imports and Productivity. *Centre for Economic Policy Research Discussion Paper No. 5139*. 27, 71, 72, 74, 116

Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50(3)*, 1029–1054. 147

Heckman, J. J. (1981a). Heterogeneity and State Dependence. In S. Rosen (Ed.), *Studies in Labor Markets*, pp. 91–139. Chicago: University of Chicago Press. 39

Heckman, J. J. (1981b). The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 179–195. Cambridge, MA.: MIT Press. 39, 41

Heckman, J. J., H. Ichimura, and P. Todd (1998). Matching as an Econometric Evaluation Estimator. *Review of Economic Studies 65(2)*, 261–294. 77

Heckman, J. J., H. Ichimura, and P. E. Todd (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies 64(4)*, 605–654. 104

Hilbe, J. M. (2011). *Negative Binomial Regression* (2nd ed.). Cambridge University Press. 128

Ho, D. E., K. Imai, G. King, and E. A. Stuart (2007). Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis 15(3)*, 199–236. 82

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: a review. *Review of Economic and Statistics 86(1)*, 4–29. 78

Kanamori, T. and Z. Zhao (2004). Private Sector Development in the People's Republic of China . *Asian Development Bank Institute (ADBI) Working Paper No.27*. 15

Kasahara, H. and B. Lapham (2013). Productivity and the Decision to Import and Export: Theory and Evidence. *Journal of International Economics 89(2)*, 297–316. 20, 28, 31, 71, 74, 93, 116

Kasahara, H. and J. Rodrigue (2008). Does the Use of Imported Intermediates Increase Productivity? Plant-level Evidence. *Journal of Development Economics 87(1)*, 106–118. 17, 27, 72, 73, 75, 116

Khandelwal, A. K. (2010). The long and short (of) quality ladders. *Review of Economic Studies 77(4)*, 1450–1476. 122, 138

Khandelwal, A. K., P. K. Schott, and S. Wei (2013). Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters. *American Economic Review 103(6)*, 2169–95. 138, 139, 165

Kim, S. and D. Park (2011). Ownership Structure and Export Performance: Firm-Level Evidence from the Republic of Korea. *Asian Development Bank, Economics Working Paper Series No. 295*. 32, 33

Kraay, A. (1999). Exports and Economic Performance: Evidence From a Panel of Chinese Enterprises. *Revue d'Economie du Developpement 1(2)*, 183–207. 71

Kugler, M. and E. Verhoogen (2009). Plants and Imported Inputs: New Facts and an Interpretation. *American Economic Review: Papers & Proceedings 99(2)*, 501–507. 71, 116

Kugler, M. and E. Verhoogen (2012). Prices, plant size and product quality. *Review of Economic Studies 79(1)*, 307–339. 137

Lawless, M. (2009). Firm export participation: entry, spillovers and tradability. *Applied Economics 41(5)*, 665–675. 26, 29

Lawless, M. and K. Whelan (2014). Where Do Firms Export, How Much and Why? *The World Economy 37(8)*, 1027–1050. 115

Levinsohn, J. . (1993). Testing the Imports as Market Discipline Hypothesis. *Journal of InternationalEconomics 35*(1), 1–22. 100

Levinsohn, J. and A. Petrin (2003). Estimating Production Functions using Inputs to Control for Unobservables. *Review of Economic Studies 70(2)*, 317–341. 17, 21, 55

Li, J. and A. He (2014). China is now No 1 in trade. *China Daily USA, Jan 13, available at http://usa.chinadaily.com.cn/epaper/2014-01/13/content_17232280.htm*, p.1. 1

Lileeva, A. and D. Trefler (2010). Improved Access to Foreign Markets Raises Plant-level Productivity For Some Plants. *Quarterly Journal of Economics 125(3)*, 1051–1099. 71, 121

Lo Turco, A. and D. Maggioni (2013). On the Role of Imports in Enhancing Manufacturing Exports. *The World Economy 36(1)*, 93–120. 116, 117

Lo Turco, A. and D. Maggioni (2014). Imports, Exports and the Firm Product Scope: Evidence From Turkey. *The World Economy DOI: 10.1111/twec.12201*. 116

Lu, J., Y. Lu, and Z. Tao (2010). Exporting behavior of foreign affiliates: Theory and evidence. *Journal of International Economics 81(2)*, 197–205. 29, 71

Lu, J., B. Xu, and X. Liu (2009). The Effects of Corporate Governance and Institutional Environments on Export Behaviour in Emerging Economies. *Management International Review 49(4)*, 455–478. 33

MacDonald, J. M. (1994). Does Import Competition Force Efficient Production. *Review of Economicsand Statistics 76*(4), 721–27. 100

MacGarvie, M. (2006). Do firms learn from international trade? *The Review of Economics and Statistics 88 (1)*, 46–60. 71, 116

Manez, J. A., M. E. Rochina-Barrachina, and J. A. Sanchis (2008). Sunk Costs Hysteresis in Spanish Manufacturing Exports. *Review of World Economics 144(2)*, 272–294. 26, 29, 33

Manova, K. (2013). Credit Constraints, Heterogeneous Firms, and International Trade. *Review of Economic Studies 80*, 711–744. 29, 33

Manova, K. and Z. Yu (2012). Firms and Credit Constraints along the Global Value Chain: Processing Trade in China. *NBER Working Paper No. 18561*. 44

Manova, K. and Z. Zhang (2009). China's Exporters and Importers: Firms, Products and Trade Partners. *NBER Working Paper No. 15249*. 29, 72, 116

Manova, K. and Z. Zhang (2012). Export Prices Across Firms and Destinations. *The Quarterly Journal of Economics 127(1)*, 379–436. 115, 137, 165

Martin, J. and I. Mejean (2014). Low-wage country competition and the quality content of high-wage country exports. *Journal of International Economics 93(1)*, 140–152. 122

Mayneris, F. and S. Poncet (2013). Chinese Firms' Entry to Export Markets: The Role of Foreign Export Spillovers. *The World Bank Economic Review 27*, 1–30. 34

McCann, F. (2009). Importing, Exporting and Productivity in Irish Manufacturing. *Working Papers No.200922, School Of Economics, University College Dublin*. 72, 75, 116

Minetti, R. and S. C. Zhu (2011). Credit Constraints and Firm Export: Microeconomic Evidence from Italy. *Journal of International Economics 83(2)*, 109–125. 33

Mullahy, J. (1997). Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior. *The Review of Economics and Statistics 79(4)*, 586–593. 144

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica 46(1)*, 69–85. 40, 42

Muuls, M. (2015). Exporters, importers and credit constraints. *Journal of International Economics 95(2)*, 333–343. 33

Muuls, M. and M. Pisu (2009). Imports and Exports at the Level of the Firm: Evidence from Belgium. *The World Economy 32(5)*, 692–734. 28, 47, 72, 74, 85, 93, 116

Olley, S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64(6)*, 1263–1298. 17, 20, 21

Orme, C. D. (1997). The initial conditions problem and two-step estimation in discrete panel datamodels. *mimeo, University of Manchester*. 42

Parra, M. D. and I. Martinez-Zarzoso (2014). Imported intermediate inputs and Egyptian exports: Exploring the links. *Working Papers 2014/09, Economics Department, Universitat Jaume I, Castellon (Spain)*. 116

Rabe-Hesketh, S. and A. Skrondal (2013). Avoiding biased versions of WooldridgeŠs simple solution to the initial conditions problem. *Economics Letters 120(1)*, 346–349. 185

Rauch, J. E. (1999). Networks Versus Markets in International Trade. *Journal of International Economics 48(1)*, 7–35. 122

Requena-Silvente, F. (2005). The Decision to Enter and Exit Foreign Markets: Evidence from U.K. SMEs. *Small Business Economics 25(3)*, 237–253. 29

Roberts, M. J. and J. R. Tybout (1997). The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs. *American Economic Review 87(4)*, 545–564. 29, 31, 32, 36, 37, 41, 71, 115

Roodman, D. (2009). How to do xtabond2: An Introduction to difference and system GMM in Stata. *The Stata Journal 9(1)*, 86–136. 146, 147, 149

Rose, A. K. and E. van Wincoop (2001). National Money as a Barrier to International Trade: The Real Case for Currency Union. *American Economic Review 91(2)*, 386–390. 163

Rosenbaum, P. and D. Rubin (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistician 39(1)*, 33–38. 81, 83

Ruane, F. and J. Sutherland (2005). Export performance and destination characteristics of Irish manufacturing industry. *Review of World Economics 141(3)*, 442–459. 71

Rubin, D. (2001). Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology 2*, 169–188. 113

Schott, P. (2004). Across-product versus within-product specialization in international trade. *Quarterly Journal of Economics 119*, 647–678. 137, 165

Seker, M. (2012). Importing, Exporting, and Innovation in Developing Countries. *Review of International Economics 20(2)*, 299–314. 116

Sianesi, B. (2004). An evaluation of the Swedish system of active labour market programmes in the 1990s. *Review of Economics and Statistics 86(1)*, 133–155. 83

Silva, A., O. Afonso, and A. P. Africano (2013). Economic performance and international trade engagement: the case of Portuguese manufacturing firms. *International Economics and Economic Policy 10(4)*, 521–547. 47

Silver, M. (2007). Do Unit Value Export, Import, and Terms of Trade Indices Represent or Misrepresent Price Indices? *IMF Working Paper No. 07/121*. 138

Silverman, B. (1998). *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall/CRC. 79, 104

Sinani, E. and B. Hobdari (2010). Export Market Participation with Sunk Costs and Firm Heterogeneity. *Applied Economics 42(25)*, 3195–3207. 26, 34

Smith, J. A. and P. E. Todd (2005). Rejoinder. *Journal of Econometrics 125(1)*, 365–375. 82

Stewart, M. B. (2007). Inter-related dynamics of unemployment and low-wage employment. *Journalof Applied Econometrics 22(3)*, 511–531. 41, 52

Sun, X. and J. Hong (2011). Exports, Ownership and Firm Productivity: Evidence from China. *The World Economy 34 (7)*, 1199–1215. 29, 71

Swenson, D. and H. Chen (2014). Multinational Exposure and the Quality of New Chinese Exports. *Oxford Bulletin of Economics and Statistics 76(1)*, 41–66. 137, 165

Topalova, P. and A. Khandelwal (2011). Trade Liberalization and Firm Productivity: The Case of India. *Review of Economics and Statistics 93(3)*, 995–1009. 74, 121

Upward, R., Z. Wang, and J. Zheng (2010). Weighing China's export basket: The domestic content and technology intensity of Chinese exports. *Journal of Comparative Economics 41(2)*, 527–543. 44

Van Biesebroeck, J. (2005). Exporting Raises Productivity in Sub-Saharan Manufacturing Plants. *Journal of International Economics 67(2)*, 373–391. 17, 20, 26, 73

Vogel, A. and J. Wagner (2010). Higher Productivity in Importing German Manufacturing Firms: Self-selection, Learning from Importing, or Both? *Review of World Economics 145(4)*, 641–665. 27, 29, 47, 72, 75, 80, 85, 89, 93, 109, 116

Wagner, J. (2007). Exports and Productivity: A Survey of the Evidence from Firm Level Data. *The World Economy 30(1)*, 60–82. 71

Wagner, J. (2012). International trade and firm performance: a survey of empirical studies since 2006. *Review of World Economics 148(2)*, 235–267. 32

Wang, Z. and Z. Yu (2012). Trading Partners, Traded Products and Firm Performances of China's Exporter-Importers: Does Processing Trade Make a Difference? *The World Economy 35(12)*, 1795–1824. 27, 29, 44, 116, 119

Windmeijer, F. A. G. and J. M. C. Santos Silva (1997). Edogeneity in Count Data Models: an Application to Demand for Health Care. *Journal of Applied Econometrics 12(3)*, 281–Ű294. 144

Wooldridge, J. M. (2002). *Introductory Econometrics: A Modern Approach* (2nd ed.). South-Western Publishing Co. 133, 159, 165

Wooldridge, J. M. (2005). Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics 20(1)*, 39–54. 41, 42, 53, 183, 185

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). The MIT Press, Massachusetts. 39

Yang, Y. and S. Mallick (2010). Export Premium, Self-selection and Learning-by-Exporting: Evidence from Chinese Matched Firms. *The World Economy 33(10)*, 1218–1240. 29, 71

Yi, J. and C. Wang (2012). The decision to export: Firm heterogeneity, sunk costs, and spatial concentration. *International Business Review 21 (3)*, 766–781. 29, 34, 37, 71

Yu, M. (2014). Processing Trade, Tariff Reductions, and Firm Productivity: Evidence from Chinese Firms. *The Economic Journal*, doi: 10.1111/ecoj.12127. 17, 44, 119, 121