2018

Three Essays on Nonlinear Effects in Development Economics: An Applied Semiparametric Approach

Jinjing Tian

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Tian, Jinjing, "Three Essays on Nonlinear Effects in Development Economics: An Applied Semiparametric Approach" (2018). *Graduate Theses, Dissertations, and Problem Reports*. 7137.
<https://researchrepository.wvu.edu/etd/7137>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Three Essays on Nonlinear Effects in Development Economics: An Applied Semiparametric Approach

Jinjing Tian

Dissertation submitted
to the College of Business and Economics
at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Feng Yao, Ph.D., Chair
Arabinda Basistha, Ph.D.
Joshua Hall, Ph.D.
Xiaoli Etienne, Ph.D.

Department of Economics

Morgantown, West Virginia
2018

Keywords: Development Economics; Semiparametric Models; Finance-growth Nexus;
Labor Share; Trade
Copyright 2018, Jinjing Tian

ABSTRACT

Three Essays on Nonlinear Effects in Development Economics: An Applied Semiparametric Approach

Jinjing Tian

The first chapter studies the finance-growth nexus with firm-level data to examine the impact of debt on firms' productivity growth in Chinese industrial firms, utilizing a newly-developed nonparametric threshold regression. It shows that while debt is overall positively related with productivity growth in Chinese industrial firms, each sample group exhibits a significant threshold effect of debt, beyond which the firms' debt exhibits diminishing returns on productivity growth. Also, the magnitude of both the estimated threshold levels and partial effects of debt is heterogeneous across ownership types and regions, with state-owned enterprises (among ownership types) and firms in the western region (among regions) having the lowest threshold values. It also finds that debt negatively affects productivity growth for state-owned enterprises (SOEs) and firms in western region after the detected threshold, indicating an inverted U-shaped relationship between external debt and firms' productivity growth.

The second chapter studies the trade impact on labor share across countries, using a fixed effect semiparametric approach. The cross-country declined labor share has been partially attributed to rising trade openness. However, the roles of export and import in the literature were not studied separately and assumed to be homogeneous across countries. It proposes two hypotheses for how export and import can affect labor share differently and nonlinearly. It empirically tests the hypotheses by re-examining the trade-labor share nexus across 96 countries during 1970-2009, and it employs a partially linear model with fixed effect that allows a general functional form of trade variables to be estimated. Results are fairly consistent with the hypotheses, showing that while export (import) share significantly declines (raises) labor share, both effects diminish as the level of export or import share increases. The indicated nonlinear effects are significant and robust by controlling for related economic, social, and political factors. Also, it finds a significant heterogeneous impact of export and import across OECD and non-OECD countries, the implication of which is also discussed.

The third chapter identifies a potential link between neoclassical trade theory and labor share literature, showing that export potentially affects labor share neutrally through itself due to its associated trade theory, and non-neutrally through three non-trade channels commonly discussed in the labor share literature. Nonetheless, most existing studies focus only on the neutral effect of trade, leaving its non-neutral effects unexplored particularly from a micro-level foundation. Given that Chinese manufacturing industries has actively engaged in exporting sector with millions of labor participation, it empirically tests the relationship by employing firm-level data in Chinese manufacturing industries during 1998-2007. To alleviate the risk of model mis-specification, it estimates the effect of export semiparametrically from a varying coefficient model with fixed effect. It implements a newly developed spline-backfitted kernel estimator to reveal the potential nonlinearity of export's impact. We show that the neutral effect of export significantly increases labor's share as expected, the non-neutral effects through the channels of firms' capital intensity, monopoly power, and capital-augmenting technological progress significantly pull down labor share, making the net effect of export less beneficial to workers and vary with firms' characteristics, regions, and time periods. The associated policy implication is given as well.

Acknowledgements

First of all, I express my genuine and great appreciation to my advisor, Dr. Feng Yao, for his generous help and guidance along the way. It is his endless knowledge of econometrics, especially nonparametric and semiparametric methods, his attitude towards hard work, and his rigorous training that help me open the door of this field and get to this point.

I am also grateful to my committee members, Dr. Arabinda Basistha, Dr. Joshua Hall, and Dr. Xiaoli Etienne, for their invaluable feedbacks and advice on my work and study. I would also like to thank the faculty, staff, and graduate students in the Department of Economics at West Virginia University.

Finally, I greatly thank my family and friends, especially my mother Dongyu Tan, my aunt Chunying Tan, and Taining Wang, for their endless love and continuous support along the way.

Contents

1	Debt and Total Factor Productivity Growth in Chinese Industrial Firms: Evidence from Nonparametric Threshold Regression	1
1.1	Introduction	1
1.2	Literature Review	4
1.2.1	Financing and Growth	4
1.2.2	Debt and Firm Productivity	6
1.2.3	The China Case	7
1.3	Data	8
1.4	Threshold Regression	10
1.4.1	Parametric threshold regression	10
1.4.2	Nonparametric threshold regression	13
1.5	Empirical Results and Discussions	15
1.5.1	Parametric Estimates	15
1.5.2	Nonparametric Estimates	16
1.5.3	By Industry	19
1.6	Robustness Check	20
1.6.1	Linear Function Specification	20
1.6.2	Parametric Threshold Regression Specification	25
1.6.3	Testing for the Existence of a Threshold	25
1.6.4	Testing for Specific Value of a Threshold	26
1.6.5	Endogeneity	27
1.6.6	Debt growth	32
1.7	Conclusion	32
1.8	Appendixes	35
2	Recasting the Trade Impact on Labor Share: A Fixed Effect Semiparametric Estimation Study	54
2.1	Introduction	54
2.2	Literature Review	57
2.3	Potential Impact of Exports and Imports	60
2.4	Data and Methodology	61
2.4.1	Data and parametric model specification	61
2.4.2	Semiparametric model specification	65
2.5	Empirical Results	70
2.5.1	Parametric model results	70
2.5.2	Semiparametric model results	72
2.6	Conclusion	81
2.7	Appendixes	82
3	Exporting Behavior and Labor Share in Chinese Manufacturing Industries: A Semiparametric Approach	93
3.1	Introduction	93
3.2	Literature Background	96
3.2.1	Labor share and export in China	96

3.2.2	Links between trade and labor share	97
3.3	Data	99
3.4	Empirical Methodologies	102
3.5	Results Discussions	105
3.5.1	Whole sample	106
3.5.2	Exporters only	108
3.5.3	Domestic vs foreign firms	111
3.5.4	Eastern and inland Region	114
3.5.5	Before vs after 2002	114
3.6	Conclusion	115
3.7	Appendixes	117

List of Figures

1.1	The rate of convergence h_1 given $p_1 = 1$ and $p_2 = 0$	46
2.1	Cross-country Labor Share Joint Density: 1970-2009	55
2.2	Constant labor share issue in PWT 8.1	62
2.3	Pooled OLS fitted line in export and import share against labor share	63
2.4	Estimated FEPLM for Trade Openness, Export Share, and Import Share	75
2.5	Gradients comparison across OECD and Non-OECD countries	78
3.1	Labor Share in Chinese Manufacturing Industries: 1998-2007	94
3.2	Exporters vs Non-exporters in capital intensity, total factor productivity, and markup. . .	96
3.3	Neutral and Non-neutral effect: whole sample	109
3.4	Neutral and Non-neutral effect: exporters	110
3.5	VCM-FE Results comparison: domestic and foreign firms	110
3.6	Gradients comparison: east and inland	112
3.7	Gradients comparison: before and after 2002	113
3.8	Whole Sample: Labor share growth rate	120

List of Tables

1.1	Data summary	11
1.2	Distribution of Debt among Firms	12
1.3	Parametric Results of Whole Sample Regression	16
1.4	Nonparametric Results for Whole Sample and Classification by Ownership	21
1.5	Nonparametric Results for Whole Sample and Classification by Region	22
1.6	Nonparametric Results for Whole Sample and Classification by Industry	23
1.7	Nonparametric Test Results For Functional Form Specification	28
1.8	Threshold Regression Result by Hansen (2000)	30
1.9	Results of Nonparametric Partial Effects of Debt Ratio on TFP Growth by Su and Ullah (2008)	31
1.10	Nonparametric Threshold Regression Results: Debt Growth	33
1.11	Simulation Results for Univariate : Sim=500, len=50	37
1.12	Simulation Results for Univariate : Sim=500, len=50	38
1.13	Simulation Results for Bivariate : Sim=500, len=50	39
1.14	Simulation Results for Bivariate : Sim=500, len=50	40
1.15	Simulation Results for Trivariate : Sim=500, len=50	41
1.16	Simulation Results for Trivariate : Sim=500, len=50	42
2.1	Summary of Selected Existing Studies: Dependent variable is labor share	59
2.2	Descriptive data summary	64
2.3	Correlation coefficient of independent variables with labor regulation index and Human capital	66
2.4	Parametric Regression Results of Trade Variable(s) Impact on Ln(Labor Share)	71
2.5	FEPLM Regression Results of Trade Variables Impact on Ln(Labor Share)	73
2.6	FEPLM Regression Results of Trade Variable(s) Impact on Ln(Labor Share)	74
2.7	Partial Linear Model Results of Trade Variables' Impact on Ln(Labor Share)	77
2.8	Nonparametric Test Results for Variable Significance, Model Specification, and Density Equality	79
3.1	Descriptive Data Summary	101
3.2	Neutrally Partial and Non-neutrally Effects from OLS-FE Estimation	107
3.3	Neutrally and Non-neutrally Partial Effect: Whole sample and exporters	108
3.4	Varying coefficient estimation results: domestic and foreign firms	111
3.5	Varying coefficient estimation results: east and inland	112
3.6	Varying coefficient estimation results: before and after 2002	113

Chapter 1

Debt and Total Factor Productivity Growth in Chinese Industrial Firms: Evidence from Nonparametric Threshold Regression

1.1 Introduction

The relationship between financial development and growth has long been debated in economics. For example, financial development can affect long-term growth by influencing saving and investment decisions and transaction costs (Levine (2005); Chen and Guariglia (2013)). Much theoretical and empirical work has reached a consensus that increased financial development leads to higher output and growth (King and Levine (1993)). Levine (2005) further illustrates that many factors contribute to economic growth through the channel of financial development, such as increasing household consumption or higher business investment.

Besides this consensus on the positive relationship between financial development and growth, Aghion et al. (2005) show that financial development has a positive but vanishing impact on per capita GDP relative to the world frontier, indicating a possible threshold effect of finance on growth. By “hypothetically” dividing nations into low, middle, and high stages of financial development based on the growth of private credit, Rioja and Valev (2004a) argue that financial development can have various effects on a nation’s economic growth, depending on the level of financial development. Studies also shed light on this finance-growth nexus from micro-level evidence. According to the M-M theorem (Modigliani and Miller (1958)), a firm’s market value and investment behavior are independent on its capital structure and financing decisions in a perfect capital market. However, in the face of imperfect capital market, finance can affect a firm’s performance by influencing its decisions on investment, labor, and innovative activities

(Nucci et al. (2005); Harless and Peterson (1998)). As Levine and Warusawitharana (2014) mention, “firm-level evidence of a link between finance and productivity growth suggests an additional and vital channel through which the financial system can impact overall economic activity” (p. 1). However, the effect of this nexus remains mixed and inconclusive in empirical studies.

This paper contributes to the literature by examining the impact of firms’ use of external financing, specifically debt, on productivity growth. We evaluate the potential non-linear and/or threshold effects (if any) of this relationship using a large panel of Chinese industrial firms. To further relax functional specification and avoid model mis-specification, we use a newly-developed non-parametric threshold technique. A series of recent economic studies have concluded that productivity growth promotes a country’s sustained economic growth rather than factor accumulation (Hall and Jones (1999); Easterly and Levine (2001)). Therefore, total factor productivity (TFP) has become the focus of many studies, and it is equally important to examine to what extent finance affects TFP growth. Although many studies examine finance-growth nexus, there is less literature emphasizing the finance-TFP growth nexus. This paper will fill this gap by looking at firms’ external debt and TFP growth. While financial development in the literature is often proxied by the availability of credit in the economy from a macroeconomic perspective, recent studies employ firm-level leverage as a superior proxy for credit availability, based on micro-level evidence (Levine and Warusawitharana (2014); Gomis and Khatiwada (2017)). Likewise, our results will also provide some implications for credit availability and TFP growth across the whole economy, using micro evidence.

Levine and Warusawitharana (2014) employ a large sample of European firms during 2000-2010 to show that firms invest in productivity-increasing activities through the use of debt, which significantly increases the firms’ productivity growth. They resolve the issue of reverse causality by decomposing total factor productivity (TFP) growth into parts that are observable and unobservable by firms, and show that it is the unobservable part of TFP growth on which debt has a significant impact. Using firm-level data in Central and Eastern European (CEE) countries, Coricelli et al. (2012) explicitly examine what level of firms’ leverage would suppress productivity growth. Their panel threshold analysis shows that leverage ratios before and after the identified threshold yield opposite effects on productivity growth, consistent with the trade-off theory of firms’ capital structure. Closely related to our paper is the study by Jin et al. (2015), who utilize firms’ debt and other financial indicators to develop financial-constraint indicators for a large panel of Chinese manufacturing firms, and endogenize the relationship between financial constraints and firms’ productivity through R&D investment channels. They find a non-monotonic relationship between financial constraints and productivity in Chinese manufacturing firms, but they do not identify the quantitative indicator that maximizes TFP growth.

China is an interesting case to study, given that since 1978 large-scale reforms have been taking place for both industrial sectors (mainly state-owned enterprises) and financial sectors during the country's economic expansion. On one hand, the rapid growth of Chinese firms in recent decades has prompted many economists to attribute the growth miracle to industrial productivity growth (Chen and Guariglia (2013)).¹ On the other hand, different progress of different reforms is likely to result in firms facing unequal degrees of credit constraints due to the underdeveloped financial sector. Bai et al. (2016) show that state-owned enterprises (SOEs) typically face “soft budget” constraints that allow SOEs to easily borrow a large amount of bank loans despite SOEs' low profitability. They examine to what extent financial frictions explain resource misallocation and productivity losses for Chinese manufacturing industries. They find that non-SOE firms typically have lower debt ratios but face higher interest rates compared to SOEs. Maliszewski et al. (2016) points out that there exists excessive credit growth for corporate debt in China, posing problems on production overcapacity, reducing investment efficiency and assets quality in financial institutions, and increasing risks for the overall economy. Given the existing empirical evidence, in this paper we investigate what is the impact of firms' external financing (i.e., debt) on their TFP growth and to what extent firms' debt hurts productivity growth in China. In other words, does an “optimal” debt level exist in the sense that Chinese firms' productivity growth can be maximized? If so, are Chinese firms borrowing too much, as some investors and economists claim?

To answer these questions, we provide empirical evidence of the effect of firms' external debt on TFP growth in Chinese industrial firms. First, we examine the relationship between firms' debt and productivity growth using Chinese Annual Surveys of Industrial Production (ASIP) produced by the National Bureau of Statistics of China (NBSC) covering the time period of 1998-2007. To detect potential threshold effects (if any) of debt on productivity growth while avoiding the risk of model mis-specification, we follow Henderson et al. (2015), Henderson and Parmeter (2013), and Su and Ullah (2008) to conduct a nonparametric threshold regression with the least functional form restriction imposed, to take into account the potential nonlinearity, interaction, and endogeneity of the regressors. We further study whether the debt-productivity nexus differs across firm ownership types, regions and industries, shedding light on differences of credit availability and financial development across China.

The rest of this paper is organized as follows. Section 2 provides a literature review on the relationship between finance and growth, debt and productivity, and economic background of the Chinese economy. Section 3 gives the data description and Section 4 details the empirical methodology. Section 5 presents our estimation results and related discussions, followed by a robustness check of our empirical results in Section 6. Section 7 concludes the paper.

¹See Ding et al. (2016a) for determinants of Chinese firms' productivity.

1.2 Literature Review

1.2.1 Financing and Growth

A macro perspective

In order to understand the relationship between finance and TFP growth, it is important to have a general understanding of the finance-growth nexus. Financial development can promote economic growth by efficiently allocating resources. King and Levine (1993) construct an endogenous growth model in which financial development mobilizes savings for the most innovative activities and diversifies the risks associated with those activities. Financial intermediaries can also directly affect economic growth. Jayaratne and Strahan (1996) study branch bank reform in the U.S since the 1970s, and show that the improvement of the banking system, rather than the amount of loans, is responsible for fast economic growth. The results therefore indicate that economic growth is not spurred by a mere increase in overall savings and investment, but rather by the allocation of high-quality bank loans to profitable projects. This argument is also in line with Levine and Zervos (1998), who state that stock market liquidity and banking development are positively and robustly related to faster contemporary and future growth.

Many studies have shown that cross-country variations in output and growth can be explained by differences in TFP growth. Hall and Jones (1999) find a large amount of variation of Solow residual levels across countries. Beck et al. (2000) test the relationship between financial intermediary development and various sources of growth. While the long run links among financial intermediary development, physical capital, and saving rates are tenuous, the authors find that financial development has a large and positive impact on TFP, which eventually transmitted to overall GDP growth.

Nonetheless, some research also shows that higher financial development does not necessarily generate positive effects on economic growth. Using a panel of 70 countries, Aghion et al. (2005) find that financial development plays a significant but vanishing role in the growth rate of an economy whose financial development exceeds the level of the threshold value. Rajan (2006) also argues that while financial sectors create more credit access for firms and households, economies may be more exposed to financial sector-induced turmoil. Cecchetti and Kharroubi (2012) support this threshold effect by arguing that the financial sector will compete with other sectors in the economy for human resources and other resources.

In summary, the literature implies that the positive effect of financial development diminishes towards zero or even becomes negative as the development increases. The indicated nonlinear relationship between financial development and growth has attracted attention by several scholars. Employing both cross-country and industry-level data, Arcand et al. (2015) identify a significant threshold for which the level of financial development would negatively affect output growth. The primary reason is that “too much

finance” exists when the credit of the private sector reaches 100 percent of GDP. Rioja and Valev (2004b) also study whether the effects of financial development on growth are the same across countries. Using a large panel of developed and developing countries, they find that TFP growth in high or middle-income countries is driven by financial development, while in less developed countries it is driven by capital accumulation.

The above finding is consistent with Henderson et al. (2013), who implemented nonparametric regression to allow for heterogeneous partial effects and endogenous regressors to be taken into account. Compared with a parametric regression, a nonparametric regression permits an unknown functional form with the fewest *a priori* restrictions imposed, and thereby it is robust to model mis-specification. Henderson et al. (2013) find that in the initial stage of development, financial development may not be key in determining growth but becomes significant in the later stage of development. Their findings also relate to the recent debate on whether low-income nations should scale up their financial reforms.

A micro perspective

As the debate over the finance-growth relationship is still going on within across-countries studies, more and more researchers have switched their attention to firm-level studies, hoping to find additional and important channels through which finance can impact economic growth. The absence of perfect capital markets makes economists to believe that firms’ financing decisions are relevant in determining firms’ performance. The availability and various sources of firm-level data make studies in this area possible. Rajan and Zingales (1998) show that financial development reduces the cost of external financing for firms and thus facilitates growth. They find that sectors that are more in need of external finance grow proportionately faster in countries with better financial development. Demirgüç-Kunt and Maksimovic (1998) show that active stock markets and well-developed financial markets are important for firm growth. Firms in countries with more active stock and financial markets can more easily obtain external financing and thus grow faster. Rahaman (2011) further examines the sources of financing and firm growth. He finds that the cost differences between internal and external financing may give rise to firms’ financial constraints and eventually limit the firms’ investment decisions. He also shows that as external financial constraints faced by firms are alleviated, firms are likely to switch from internal to external finance, which is the main source of firms’ growth. Midrigan and Xu (2014) use plant-level data to evaluate the role of financial frictions on total factor productivity. They find that financial constraints can reduce TFP through distortion to firm entry and technology adoption, dispersing the returns on capital across existing producers.

Financing can also have a negative effect on a firm’s growth. From the firm capital structure view,

Lang et al. (1996) find a negative relationship between leverage and growth. They argue that higher leverage may decrease the funds available to make profitable investments and hurt firms' ability to raise funds. Similar negative arguments are also found in Gennaioli et al. (2012), who state that financial intermediaries create securities that will expose the economy to "neglected risks".

1.2.2 Debt and Firm Productivity

Another strand of literature examines the relationship between firms' debt or leverage and productivity, to shed light on the finance-growth nexus. Levine and Warusawitharana (2014) find that a one standard deviation increase in debt leads to between a 0.033 to 0.059 standard deviation increase in firms' TFP over the next period using data from European corporations; the results also hold using other productivity measures, such as labor productivity. Given the possibility that firms may obtain financing in expectation of future productivity increases, they follow Levinsohn and Petrin (2003) to cope with reverse causality by decomposing TFP growth into expected and unexpected components. The results confirm that debt mainly promotes the unexpected component, i.e., the part of TFP unknown to firms.

Debt or leverage is often considered an external financial constraint for firms when they make investment decisions (Hale and Long (2011)). Li (2015) develops a general equilibrium model in which more financially constrained firms have a higher output-to-capital ratio, and empirically supports it by a panel of Japanese firm level data. Jin et al. (2015) study the relationship between financial constraints and firm productivity for a large panel of Chinese manufacturing firms. They find that 90% of firms in their sample face financial constraints, and that state-owned enterprises (SOEs) are much less financially constrained than non state-owned enterprises (non-SOEs). They also find an inverted U-shaped relationship between financial constraints and firm productivity. Their argument regarding this nonlinearity is that "financial constraints raise productivity at the early stage by reducing the waste of cash and sub-optimal investments, but they may reduce productivity subsequently as the paucity of funds reduces value-enhancing investments." Nucci et al. (2005) examine whether exogenous factors influencing firms' capital structure affect firms' productivity through the allocation of capital between material and immaterial assets. They find a negative relationship between leverage and share of intangible assets, a positive relationship between intangible assets and productivity, and thus an overall negative relationship between leverage and firm productivity.

The relationship between debt and productivity can also be explained in terms of firms' capital structure. Coricelli et al. (2012) posit a non-monotonic relationship between firm leverage and productivity growth based on the trade-off theory of firms' capital structure, and identify a critical threshold beyond which leverage lowers firms' TPF growth. Trade-off theory typically argues that there is a benefit of

tax shields and a cost of bankruptcy with firm leverage. The theory argues that there exists an optimal capital structure that maximizes firm value, which is also positively related to productivity. They document some threshold effects among firms in Central and Eastern Europe, most of which are transition economies. This is in line with Matemilola et al. (2014), whose study identifies a threshold effect of debt ratio at 18.56% using firm-level data from South Africa. Equity ownership can also play a role in a firm's capital structure and productivity, such as in the study by Pushner (1995), who observes a negative relationship between institutional ownership and leverage and between leverage and productivity in Japanese firms.

1.2.3 The China Case

Exploring the relationship between debt and firm productivity growth in China is particularly significant. Since the economy has been undertaking reforms in the so-called “socialist market economy” and has become one of the world's largest emerging economies since the “Open Door Policy” of 1978, it is of great importance and interest to know the role of financing in China's growth miracle. This paper will provide insight into this relationship for the Chinese economy from a micro perspective.

Two of the major reforming sectors during China's economic development are the industrial sector and the financial sector. For example, in order to reduce the claims of budget and bank loans, numerous inefficient and smaller SOEs have been shut down or merged into other larger enterprises. SOEs' contribution to domestic gross output has also fallen dramatically Gang and Hope (2013). The development of private enterprises started in the 1980s, and sped up in 1992 after an important speech by Xiaoping Deng regarding the importance of private enterprises, delivered during his tour of southern China. By the 15th Congress of the Communist Party in 1997 private enterprises had become a major part of the economy, with an annual growth rate of 20%, far above the nation's 9% economic growth rate (Chow et al. (2004)).

Reforms in the financial sector also took place to meet the needs of the fast-growing economy, and various kinds of financial institutions with different ownership models emerged. However, the financial markets and main national banks have been criticized for their incapability of catching up with the fast pace of industry reforms (Curtis (2013)). Song et al. (2011) document that entrepreneur firms utilizing more productive technologies face higher financial constraints, causing the majority of their investments to be financed only through internal savings. By contrast, SOEs with low productivity can still survive due to their easy access to credit markets, or a “soft budget constraint” (Jin et al. (2015);).

Firms with soft budget constraints may be motivated to undertake inefficient activities rather than productivity-enhancing investment. Kornai et al. (2003) finds that the soft budget constraints are likely

to exist in transition economies. Given the underdeveloped stock and bond markets in China, firms' debt is mainly loans from banks Xiao (2005), and can be considered a proxy for access to credit or the level of credit available in the economy Gomis and Khatiwada (2017). Aside from firms' ownership types, the relationship between debt and productivity growth across regions may also reflect differences in regional economic and financial development. In this study we also split China's economic geography into eastern, western, central, and northeastern regions, each of which has a different market environment. The last dimension of heterogeneity we look at is industry classification. We divide the sample into manufacturing, public utility, and mining sectors according to the classification of National Bureau of Statistics of China.

Overall, the existing literature and evidence points to an inconclusive relationship between financing and growth, and between firms' debt and productivity growth. Whether these relationships are conclusive in China remains unknown. We therefore contribute to the literature by employing a large panel of Chinese industrial firms to unveil the relationship between a firm's use of financing and its productivity growth.

1.3 Data

The data in this paper is sourced from the Chinese Annual Surveys of Industrial Production, officially referred as the "all state-owned and all above-scale non-state owned industrial enterprise database", maintained by the National Bureau of Statistics of China (NBSC). It is survey data that covers all state-owned and above-scale, non-state-owned enterprises with annual sales at or above five million RMB (611,995 US dollars), including mining, manufacturing, and public utilities.² All the firms are based in 31 provinces or province-equivalent municipal cities in China. This database is becoming one of the most important data sources to study the Chinese economy (Song et al. (2011); Hsieh and Klenow (2009)). A list of 2-digit Chinese Industrial Classification (CIC) codes and names is provided in the appendix.

Following Guariglia et al. (2011) and Chen and Guariglia (2013), ownership in this paper is defined as share in paid-in capital. There are six types of paid-in capital in this data set: state, collective, legal person, private, Hongkong, Macau and Taiwan, and foreign firms. State-owned enterprises and collective enterprises are, respectively, firms with more than 50% of the firm's paid-in capital owned by the state or by urban or rural communities. Private enterprises include ownership by private persons and legal persons.³ Foreign ownership firms have more than 50 percent of paid-in capital from Hongkong, Macau and Taiwan, or foreign countries.

The regional data is grouped into four geographic regions: eastern, northeastern, central, and western

²The exchange rate of 8.17 RMB per USD is taken as the average nominal exchange rate between 1998 and 2007 from the Federal Reserve Bank.

³Legal persons are defined as various domestic institutions such as banks and research institutions. Since the goal of this ownership is to make profit, it is treated as private ownership.

regions. The eastern area is traditionally considered as the most developed area of the economy and includes provinces that are located near or along coastlines. The northeastern region includes provinces where most of the nation's old state-owned enterprises have started up. A list of provinces in each geographic region is provided in the appendix.

TFP growth is calculated based on the Levinsohn-Petrin method (Levinsohn and Petrin (2003)).⁴ Variables used in TFP growth calculation include *output* as a firm's real gross industrial product, *capital* as real total fixed capital, *labor* as real sum of wages and benefits, and *material* as the real intermediate material firms used in production.

Our key variable *debt ratio* is defined as a firm's total debt divided by total assets, with other control covariates chosen based on the literature on firms' productivity growth: *CF ratio* is defined as a firm's cash flow divided by total assets, *Age* is the difference between a firm's observation year and establishment year, *Firm size* is the logarithm of a firm's real total asset, and *Export* is the export intensity, defined as total export divided by total output. Real variables are calculated using price indexes from National Bureau of Statistics of China. The data is trimmed by dropping the highest and lowest 1% of each variable to remove outliers and also dropping firms with missing values or obvious measurement errors. The final sample is a balanced panel data with 8,319 industrial firms doing business from 1998-2007, resulting in 83190 observations.⁵

Table (3.1) provides descriptive statistics of the data, classified by ownership types, regions, and industries. Data across ownership groups show that firms on average have similar productivity growth rates of around 26.2%, with foreign firms having the highest rate (27.9%) and SOEs having the lowest rates, both at mean (25.3%) and median (5.4%). The CF ratio and rare engagement in export activities also indicate the low profitability of SOEs relative to firms of other ownership types.

Data across regions reveals that the eastern region, the most developed area in China, has the lowest mean productivity gain but the highest mean debt ratio. Northeastern region seems to enjoy the highest (mean) TFP growth. It may initially seem counterintuitive given the fact that many "heavy" industries and SOEs are located here since this region was initially targeted as the nation's "industrial base". However, SOEs in this region have begun reforms since early 1990s and many policies have established to attract high-tech industries such as electronic manufacturing. Our data shows actually in this region foreign firms accounted for 40% of all observations.⁶ Zhou and Hall (2017) identify the positive relationship between measures of marketization and entrepreneurship across Chinese provinces. Northeastern region's marketization indexes are above the national average.

⁴Stat command *Levpet* is used to calculate TFP growth. Please refer to Petrin et al. (2004) for more details.

⁵The observations drops to 74861 because data of year 1998 was dropped after calculating growth rates of TFP

⁶Please refer to the appendix for percentage distributions of different ownership types among regions.

When the sample is classified by industry, mining industry is having the highest TFP growth and debt ratios compared with other two industries. Zhu et al. (2018) show that technological progress has been the most important contributor to total factor productivity in Chinese mining and quarrying industries. To get a better idea of how debt is distributed among Chinese firms, a more detailed distribution of debt ratio by ownership types, regions, and industries is presented in Table (1.2). Generally speaking, there are no obvious differences in debt ratios among firms across regions. However, foreign firms (among ownership types) and firms in public utility industry (among industries) having the lowest debt ratios at each quantile, respectively.

1.4 Threshold Regression

1.4.1 Parametric threshold regression

To study the relationship between finance and growth in China and examine the potential threshold effects of external finance (i.e., debt ratio) on productivity growth, we employ both parametric and nonparametric threshold regression. The parametric threshold regression is by Hansen (2000), which detects endogenously the existence of threshold by splitting the sample into two “regimes”, each of which is used to estimate the coefficients of threshold variable that is potentially different across regimes. Specifically, we consider the model:

$$TFPG_{it} = \alpha_1 L_{it} I(L_{it} \leq \gamma) + \alpha_2 L_{it} I(L_{it} > \gamma) + \mathbf{X}_{it} \boldsymbol{\beta} + e_{it} \quad (1.1)$$

where $i = 1, \dots, N$ and $t = 1, \dots, T$ index a total of N firms over T time periods in our pool dataset, and e_{it} is the additive zero-mean error term. Our dependent variable $TFPG$ is firm’s total factor productivity calculated by Levinsohn and Petrin (2003) and Petrin et al. (2004). L is our key explanatory variable of firm’s debt ratio that has potential threshold effect on $TFPG$, followed by an indicator function of its argument, $I(\cdot)$. Debt ratio has been argued to be a good proxy for firms’ external finance conditions (Levine and Warusawitharana (2014); Gomis and Khatiwada (2017)). This is true particularly in China, whose capital market is still underdeveloped and most of firms’ short-term and long-term debts come from bank loans (Xiao (2005)).

Matrix X contains control variables. We follow the firm governance literature to include firms’ size ($\ln \text{total asset}$); firms’ age in logarithm form ($\ln \text{Age}$); export intensity (export), and cash flow ratio (CF ratio).

For any value of γ , say $\gamma = \gamma_0$, we obtain the OLS estimates $\beta(\gamma_0)$ and $\alpha(\gamma_0)$, and calculate the sum

Table 1.1: Data summary

	Whole sample																			
	State	Private	Foreign	Collective	Eastern	Western	Northeastern	Central	Manufacturing	Public Utility	Mining									
TFP growth	25%	-0.200	-0.184	-0.183	-0.211	-0.196	-0.203	-0.205	-0.237	-0.167	-0.204	-0.178	-0.140							
	50%	0.066	0.054	0.078	0.062	0.073	0.062	0.062	0.059	0.096	0.067	0.044	0.114							
	75%	0.432	0.386	0.432	0.452	0.438	0.423	0.452	0.490	0.468	0.436	0.351	0.509							
	mean	0.262	0.253	0.259	0.279	0.257	0.251	0.280	0.315	0.299	0.263	0.226	0.323							
Debt ratio	25%	0.348	0.379	0.428	0.278	0.407	0.366	0.348	0.305	0.348	0.351	0.308	0.353							
	50%	0.530	0.558	0.601	0.462	0.584	0.546	0.540	0.506	0.547	0.534	0.477	0.541							
	75%	0.693	0.725	0.749	0.631	0.740	0.710	0.703	0.701	0.716	0.696	0.653	0.699							
	mean	0.516	0.548	0.579	0.458	0.563	0.532	0.523	0.502	0.528	0.519	0.482	0.520							
CF ratio	25%	0.031	0.015	0.032	0.042	0.032	0.032	0.022	0.021	0.029	0.032	0.021	0.030							
	50%	0.063	0.038	0.063	0.084	0.068	0.065	0.046	0.053	0.067	0.065	0.045	0.030							
	75%	0.119	0.070	0.118	0.144	0.137	0.119	0.087	0.104	0.160	0.122	0.075	0.067							
	mean	0.098	0.050	0.106	0.104	0.122	0.095	0.067	0.077	0.140	0.100	0.057	0.164							
Age	25%	8.0	15.0	7.0	7.0	9.0	7.0	8.0	8.0	8.0	7	15	11							
	50%	11.0	30.0	11.0	9.0	13.0	11.0	13.0	12.0	14.0	11	26	20							
	75%	18.0	43.0	17.0	12.0	21.0	16.0	29.0	19.0	30.0	15	37	37							
	mean	15.6	29.8	14.7	9.9	16.7	14.4	19.4	17.0	19.4	14.4	27.7	24.5							
Total asset	25%	46.2	76.1	35.3	70.8	32.6	45.2	64.0	52.9	32.0	44.5	92.4	40.2							
	50%	107.8	200.9	77.4	175.2	63.9	101.2	152.6	127.1	85.2	102.1	232.8	114.9							
	75%	302.8	587.7	205.0	500.0	134.8	275.0	410.0	394.6	263.2	282.2	645.8	389.5							
	mean	694.2	1734.5	504.2	780.6	232.0	535.7	780.7	1171.3	1023.9	611.9	1509.1	1445.0							
Export	25%	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000							
	50%	0.000	0.000	0.000	0.574	0.000	0.013	0.000	0.000	0.000	0.008	0.000	0.000							
	75%	0.709	0.000	0.319	0.950	0.000	0.794	0.396	0.820	0.000	0.787	0.000	0.000							
	mean	0.299	0.046	0.219	0.515	0.129	0.335	0.246	0.334	0.081	0.328	0.000	0.030							
Observations	74,871	12,251	30,626	16,924	10,672	53,232	8,640	4,771	8,151	67,878	5,166	1,827								

Note: Data source is from NBS. The table reports 25th, 50th, 75th, and sample means. TFP growth is calculated industry by industry using Levinsohn and Petrin (2003); Debt ratio is a firm's total debt divided by total asset; CF ratio is a firm's cash flow divided by total asset; Age is the difference between a firm's data year and establishing year; Total asset is a firm's total assets deflated by price deflators from NBS; Export is total export divided by total output.

Table 1.2: Distribution of Debt among Firms

	10%	25%	50%	75%	90%
Whole sample	0.187	0.348	0.530	0.693	0.817
<i>By Ownership</i>					
State	0.227	0.379	0.558	0.725	0.849
Private	0.262	0.428	0.601	0.749	0.859
Foreign	0.140	0.278	0.462	0.631	0.765
Collective	0.233	0.407	0.584	0.740	0.856
<i>By Geography</i>					
Eastern	0.203	0.366	0.546	0.710	0.834
Western	0.183	0.348	0.540	0.703	0.833
Northeastern	0.146	0.305	0.506	0.701	0.849
Central	0.172	0.348	0.547	0.716	0.841
<i>By Industry</i>					
Manufacturing	0.188	0.351	0.534	0.696	0.818
Public Utility	0.178	0.308	0.478	0.653	0.790
Mining	0.188	0.353	0.541	0.699	0.818

Data source: NBSC

of squared errors $S_n(\beta(\gamma), \alpha(\gamma))$, where $\alpha = [\alpha_1, \alpha_2]$ defined in (1). Estimation of $[\alpha, \beta]$ in (1) requires γ to belong to a bounded set $\Gamma = [\underline{\gamma}, \bar{\gamma}]$, and we set the upper ($\bar{\gamma}$) and lower bound ($\underline{\gamma}$) of γ , respectively, to be 0.05 and 0.95 percentile of our threshold variable L . Therefore, our objective is to search over a grid point of $\gamma \in \Gamma$ to obtain $\hat{\gamma}$ that satisfies:

$$\hat{\gamma} = \underset{\{\gamma \in \Gamma\}}{\operatorname{argmin}} S_n(\beta(\gamma), \alpha(\gamma)), \quad (1.2)$$

and we evaluate the significance of the estimated threshold by testing the null hypothesis $H_0 : \alpha_1 = \alpha_2$ using the following feasible test statistic:

$$LR_0 = \frac{S_0(\hat{\gamma}) - S_1(\hat{\gamma})}{\hat{\sigma}^2} \quad (1.3)$$

where S_1 (S_0) denotes the residual sum of squares in (1) with (without) threshold, and $\hat{\sigma}^2$ is the estimated residual variance from model *with* threshold. Clearly, LR_0 converges to zero if and only if the null is true. We follow Hansen (1996, 1999) to improve the power by obtaining its p-value from the bootstrapped empirical distribution of the test statistics.

1.4.2 Nonparametric threshold regression

While parametric methods possess the advantage of parsimony, the potential disadvantage of model misspecification makes its nonparametric counterparts alluring. We thereby follow Henderson et al. (2015) to estimate γ nonparametrically by allowing for all covariates to enter the regression via unknown but smoothing functions $m_1(\cdot)$ or $m_2(\cdot)$:

$$TFPG_{it} = m_1(\mathbf{Z}_{it}) + e_{it}, \text{ if } L_{it} \leq \gamma \quad (1.4)$$

$$TFPG_{it} = m_2(\mathbf{Z}_{it}) + e_{it}, \text{ if } L_{it} \geq \gamma \quad (1.5)$$

where $\mathbf{Z} = [L, \mathbf{X}]$ contains all regressors in our study, and we again split the sample into regime 1 or 2 (i.e., equation (4) or (5)) based on the threshold value of L . The key difference between (4)-(5) and (1) is that the former permits any potential nonlinearity and interaction of variables in $\mathbf{Z}$ to be estimated, therefore robust to potential model mis-specification. We can rewrite (4)-(5) in a compact manner:

$$TFPG_{it} = \alpha(\mathbf{Z}_{it}) + \beta \cdot I(L_{it} > \gamma) + e_{it} \quad (1.6)$$

where

$$\alpha(\mathbf{Z}_{it}) = \begin{cases} m_1(\mathbf{Z}_{it}), & \text{if } L_{it} \leq \gamma \\ m_2(\mathbf{Z}_{it}) - \beta, & \text{if } L_{it} \geq \gamma \end{cases}$$

where $I(\cdot)$ is the indicator function of its condition, and β is interpreted as the jump size of the regression function at $L_{it} = \gamma$. Similar as Hansen (2000), we estimate γ in (6) by concentrating out both $\alpha(\cdot)$ and β in the following three-stage procedure.

For a given value of γ we first obtain $\hat{\alpha}_h(\mathbf{Z}; \gamma, \hat{\beta}_h(\gamma)) \equiv \hat{\alpha}_0$, a local-linear estimator evaluated at $\mathbf{Z}$ from

$$(\hat{\alpha}_0, \hat{\alpha}_1) = \underset{\{\alpha_0, \alpha_1\}}{\operatorname{argmin}} \sum_{i=1}^N \sum_{t=1}^T [\tilde{y}_{it} - \alpha_0 - (\mathbf{Z}_{it} - \mathbf{Z})\alpha_1]^2 K_H(\mathbf{Z}_{it}, \mathbf{Z}) \quad (1.7)$$

where $\tilde{y}_{it} = [y_{it} - \hat{\beta}_h(\gamma) \cdot I(\mathbf{Z}_{it} > \gamma)]$, $\alpha_1 = [\alpha_{11}, \dots, \alpha_{1d}]'$ is a $(d \times 1)$ vector of the first-order gradient estimates of $\mathbf{Z}$ given $\tilde{y}_{it}$, and $\hat{\beta}_h(\gamma) \equiv \hat{b}$ from

$$\hat{b} = \underset{\{b\}}{\operatorname{argmin}} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{\alpha}_h(\mathbf{Z}_{it}; \gamma, b) - bI_{it}(\gamma)]^2 \hat{f}(\mathbf{Z}_{it})^2$$

where $\hat{f}(\mathbf{Z}_{it}) = \frac{1}{NT} \sum_{l=1}^N \sum_{m=1}^T K_H(\mathbf{Z}_{lm}, \mathbf{Z}_{it})$ is the estimated joint density function at $\mathbf{Z}_{it}$. The solution

is given by

$$\hat{\beta}_h(\gamma) = \left(\frac{1}{NT} \sum_{l=1}^N \sum_{m=1}^T \tilde{I}_{lm}(\gamma)^2 \right)^{-1} \frac{1}{NT} \sum_{l=1}^N \sum_{m=1}^T \tilde{I}_{lm}(\gamma) y_{lm}^* \quad (1.8)$$

where for $l = 1, \dots, N$ and $m = 1, \dots, T$, $y_{lm}^* = (1/NT) \sum_{i=1}^N \sum_{t=1}^T K_H(\mathbf{Z}_{it}, \mathbf{Z}_{lm})(y_{it} - y_{lm})$ with $(\mathbf{Z}_{lm}, y_{lm})$ the evaluation points of choice,

$$\tilde{I}_{lm}(\gamma) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T K_H(\mathbf{Z}_{it}, \mathbf{Z}_{lm}) [I_{it}(\gamma) - I_{lm}(\gamma)]$$

is an indicator function defined in (6) (i.e., $I_{it}(\gamma) = I(L_{it} > \gamma)$), and

$$K_H(\mathbf{Z}_{it}, \mathbf{Z}) = \frac{1}{|H|} \prod_{s=1}^d k\left(\frac{Z_{it}^s - Z^s}{h_s}\right), |H| = h_1 \dots h_d$$

is a product kernel of univariate kernel function $k\left(\frac{Z_{it}^s - Z^s}{h_s}\right)$ that is symmetric and integrated to one.

The intuition of local linear estimator in (7) is that for each evaluation point we “locally” fit a linear line through $\mathbf{Z}$ based on the its neighbor observation near to $\mathbf{Z}$. The kernel function $K_H(\mathbf{Z}_{it}, \mathbf{Z})$ produces higher (lower) weights for points closer to (further away from) $\mathbf{Z}$. We use bandwidth h_s to control the range for which points are consider as the neighbor points around $Z^s \in \mathbf{Z}$. Therefore, a larger bandwidth implies more equal weight assigned to the neighbor points of $\mathbf{Z}$, and vice versa. We then connect each of these linear lines to obtain the estimate of unknown function as well as gradient. In fact, it can be shown that local linear estimator collapses to OLS estimator if the unknown function is truly linear (i.e., the local linear regression fits a linear line using all sample points as neighbor points of $\mathbf{Z}$, so $\{h_s\}_{s=1}^d$ must be diverge to infinity).

Finally, with both estimates in (7) and (8) we obtain an approximate estimate of

$$\hat{M}_n(\gamma, \hat{\alpha}_h(\gamma, \hat{\beta}_h(\gamma))) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left[y_{it} - \hat{\alpha}_h(\mathbf{Z}_{it}; \gamma, \hat{\beta}) - \hat{\beta}_h(\gamma) I_{it}(\gamma) \right]^2 W(L_{it}) \quad (1.9)$$

where $W(\cdot)$ is a trimming function for the threshold variable to eliminate extreme value of L . We repeat steps (7)-(9) until we find an optimal threshold estimate $\hat{\gamma}$ by searching over a grid of points $\gamma \in \Gamma$ that satisfies:

$$\hat{\gamma} = \underset{\{\gamma \in \Gamma\}}{\operatorname{argmin}} |\hat{M}_n(\gamma, \hat{\alpha}_h(\gamma, \hat{\beta}_h(\gamma)))| \quad (1.10)$$

Once we obtain $\hat{\gamma}$, we split our sample data accordingly based on (4)-(5), and proceed to estimate the

first-order gradient of $\mathbf{Z}$ in $m_j(\mathbf{Z}_{it})$ with $j = 1, 2$ via local linear estimator of $\hat{\theta}_1$ from:

$$(\hat{\theta}_0, \hat{\theta}_1) = \underset{\{\theta_0, \theta_1\}}{\operatorname{argmin}} \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \theta_0 - (\mathbf{Z}_{it} - \mathbf{Z})\theta_1]^2 K_H(\mathbf{Z}_{it}, \mathbf{Z}) \quad (1.11)$$

and our gradients of $\mathbf{Z}$ is observation-specific as opposed to constant partial effects in parametric regression.

We note that the original paper lacks the simulation result for multivariate case. Therefore we perform a Monte-Carlo simulation to check the consistency of the proposed estimator in both univariate and multivariate setting. The results reported in appendix shows that the bias of estimated threshold level $\hat{\gamma}$ is very close to zero for trivariate regression even if the sample size is limited to 400. Given the large sample size of our empirical data, we are thus confident to proceed to implement nonparametric threshold regression in the study below.

1.5 Empirical Results and Discussions

1.5.1 Parametric Estimates

Table (1.3) presents our parametric threshold regression for the whole sample, with debt ratio as our main variable. For comparison, linear and quadratic model specifications are also included. The estimation in (1) gives a threshold of 0.3742 (i.e., debt ratio = 0.3742) along with its p-value of 0.078, which is close to the 28th quantile of debt ratios in the sample.⁷ When *debt ratio* enters the regression linearly, its effect is positive at the 1% significance level. In other words, a one-percentage point increase in *debt ratio* increases TFP growth by 0.2%. Comparing results in Columns (1) and (2), we see that *debt ratio* has a significant nonlinear effect on TFP growth once a threshold effect is considered. This indicates that when the *debt ratio* exceeds the estimate threshold $\hat{\gamma} = 0.3742$, the effect of debt decreases from 0.3074 to 0.2077. Although direct evidence for the “excessive” debt with negative partial effects passing the threshold is not found, the results still suggest that debt ratio has a heterogeneous and diminishing marginal impact on TFP growth.

Notably, the signs and magnitude of the other control variables remain similar across the linear, quadratic, and threshold models. *Age* is negatively associated with firms’ TFP growth in China, because as firms get older their ability to promote productivity, such as innovation, is reduced. This result is in line with most studies in the literature (Chen and Guariglia (2013); Jin et al. (2015)). *CF ratio*, measured by the ratio of cash flow to total assets, is significantly positively related to TFP growth. It is argued in

⁷When firm fixed effect is controlled using Hansen (1999), a threshold value of 0.4812 is detected with 10% significance. The partial effect of each variable is quantitatively and qualitatively similar as the one we find here.

Table 1.3: Parametric Results of Whole Sample Regression

Variable	[1]	[2]	[3] ($\hat{\gamma} = 0.3742$)
Debt ratio	0.1965*** (0.0274)	0.2746*** (0.0884)	
Debt ratio-squared		-0.0785 (0.0843)	
Debt ratio*I(Debt $\leq \gamma$)			0.3074*** (0.0358)
Debt ratio*I(Debt $> \gamma$)			0.2077*** (0.0205)
Ln(age)	-0.0768*** (0.0111)	-0.0767*** (0.0111)	-0.0640*** (0.0088)
Exp	-0.1937*** (0.0183)	-0.1934*** (0.0183)	-0.1690*** (0.0174)
CF ratio	0.3952*** (0.0276)	0.3948*** (0.0276)	0.3520*** (0.0210)
Ln(total asset)	0.0045 (0.0098)	0.0044 (0.0098)	0.0029 (0.0088)
Constant	0.3482*** (0.0550)	0.3333*** (0.0573)	0.3790*** (0.0390)
F-stat	74.1492***	74.1505***	79.1339***
R^2	0.0850	0.0850	0.0934
Observations	74,871	74,871	74,871

Note: Standard errors in parentheses. For linear and quadratic model, time and individual fixed effects are taken into account. ***p<0.01, **p<0.05, *p<0.1

corporate-governance literature that cash flow can be considered one of a firm's internal financing sources. Much of the literature argues that Chinese firms, especially non-state-owned enterprises, face external financial constraints, so they have to seek help from their internal savings (Song et al. (2011);). In our case, it is evident that a firm's cash flow serves as a source of internal financing and helps to promote productivity growth. *Exports* are negatively associated with firms' TFP growth, which contradicts the positive correlation between exports and firms' total factor productivity argued in most of the trade literature. However, Fu (2005) argues that the low skill and low technology content of exported products, the intensive use of unskilled labor, and the low-price competition in Chinese export industries are likely to discourage incentives for innovation. Finally, *firm size* is statistically insignificant, though it has a positive sign.

1.5.2 Nonparametric Estimates

Whole sample

Columns 1-2 in Table (1.4) present results from the nonparametric threshold regression for the whole sample. Following Henderson et al. (2015), we estimate functions $m_1(\cdot)$ and $m_2(\cdot)$ nonparametrically by splitting the data according to our estimated threshold value. For the whole sample debt ratio shows a threshold of 0.5122, lower than both the median and mean values observed in the data. However, since

the estimated coefficient of each variable in Z is evaluated at each observation (i.e., see (11)), it is tricky to present a graphical representation in higher-dimensional data, particularly when unknown interactions among variables exist. Therefore we estimate $\hat{m}(\cdot)$ by holding all other regressors at their median values, and report the 50th quantile estimate of the gradient distributions along with their asymptotic standard errors.⁸

Clearly, partial effects of debt are both positive and significant. Holding all other regressors at their median levels, a one percentage point increase in firms' debt will increase TFP growth by 0.361 of a percentage point before the threshold, and 0.191% after the threshold. As for the mean value of the estimated coefficient, the effect in the left regime is 0.360 while that in the right regime is 0.156, an even larger drop in marginal returns of debt indicating a strong diminishing return of debt on firms' TFP growth. We argue that firms with debt exceeding the threshold value should pay attention to this relationship, because the association of debt to TFP growth starts to decline with increasing debt. This finding is consistent with Jin et al. (2015), who find an inverted U-shaped relationship between firms' financial constraints and TFP growth in China. It also raises concerns that the linear functional form in parametric estimation may be misspecified. The remaining covariates have the sign and significance of the coefficients approximately similar to their parametric results.⁹

In case that the region of northeast is an outlier since it has counterintuitively high TFP growth, the whole sample without northeastern region is used. A slightly lower threshold value of 0.4532 is detected with other covariates having quantitatively and qualitatively effects on TFP growth as in the original sample.

By Ownership

Table (1.4) also reports results across ownership types. For state-owned enterprises (SOEs) a threshold value of 0.2931 is detected, nearly half of the median debt ratio of 0.558 in the data (close to 16th percentile in the sample). Debt is positively related to SOEs' TFP growth, but the effects are neither statistically nor economically significant. When debt passes the threshold value, there is a drop in marginal return of debt on productivity growth. This raises the criticism that state-owned enterprises are borrowing too much, since SOEs' debt does not contribute much to TFP growth. Another appealing finding is that the firm size is negatively correlated with TFP growth. The result confirms the fact that in China, state-owned enterprises usually own a large amount of assets compared with other enterprises, but the bigger size does not make a positive contribution to SOEs' productivity growth in our results. Private firms,

⁸See Pagan and Ullah (1999) for more details on the asymptotic standard error formula.

⁹We do notice that, contradictory to OLS estimation, exports are now significantly positively related to TFP growth, with different effects in magnitude before and after threshold. The change in magnitude and signs of parameters may be mainly driven by a greater flexibility of the functional form.

which accounted for more than 40.9% of all observations in our sample, have an estimated threshold value of 0.5085 (close to the 35th percentile in the sample). Before that value, debt is significantly and positively related to TFP growth. Every one percent point increase in debt ratio is associated with an 0.27% increase in TFP growth. The effect is still positive after passing the threshold value, but it drops to 0.128%. Contrary to SOEs, private firms' size is positively related with productivity growth.

Foreign firms show a threshold value of 0.2903 (close to the 25th percentile). Debt's impact on TFP growth before the threshold is 0.036, and drops to -0.283 after the threshold. The opposite signs of coefficients indicate an inverted U-shaped relationship between debt and productivity growth. Foreign firms have the lowest average debt ratio (0.4579) compared with other types of enterprises, since they may be biased in China's financial markets or they have other sources of financing from abroad (Chen and Guariglia (2013)). However, both the mean and median values of debt ratio for this group far exceed the detected threshold value. Finally, collective firms have significantly positive effects of debt, both before and after the detected threshold of 0.5154 (close to 49th percentile). Since the impact drops slightly from 0.220 before the threshold to 0.196, it indicates that debt plays an important role in promoting collective firms' productivity growth. We can also see that cash flow positively impacts firms' TFP growth for all types of ownership. The result is consistent with Chen and Guariglia (2013), who argue that firms in China are more likely to use their cash flows as internal financing to support their innovation and production activities.

By Region

Table (1.5) presents the results of threshold regressions across different regions. The eastern region is considered the most economically developed one. We detect a threshold value of 0.5137, which is close to the median debt value of 0.5304. Debt ratio is highly significant and positive before the threshold in this area: a one percentage point increase in debt ratio raises TFP growth by 0.593 of a percentage point. But this effect drops significantly to 0.133 after passing the threshold. These results again support our argument that debt exhibits a diminishing return on TFP growth. Exports are significantly and positively associated with productivity growth in this region. Given the fact that most of the exporting industries locate their factories in the eastern region, these results provide evidence that exports boost firms' TFP growth in these areas, in line with other research Ma et al. (2014). The coefficients of cash flow ratios are positive and highly significant, again consistent with Chen and Guariglia (2013). A one percentage point increase in cash flow increases TFP growth by 0.582% before the threshold value and 0.547% after the threshold value.

Another region where debt plays a highly significant and positive role before the threshold value is

the western region. For this region, a threshold value of 0.3945 is detected, far exceeding the mean and median values in the data (close to the 32.5th percentile in the data). Before this value, a partial effect of 0.923 is associated with TFP growth, which drops to 0.21 after the threshold value. The partial effect before the threshold is the highest among the four regions, although the detected threshold is the lowest. This result indicates that when starting initially with small debt, firms in the western region may use it to invest in productivity-enhancing activities, but when their debt exceeds the threshold value, these firms may face financial distress and use more of their money to pay back debt, resulting in a huge drop in debt's contribution to productivity growth. Exports are significantly negatively correlated with TFP growth for firms in the western region. In recent years, local governments in these areas have implemented policies that encourage firms, especially manufacturing firms, to move their production lines to the west in order to boost the local economy and employment. However, our results don't show a positive relationship between the two variables.

For firms in the northeastern region, a threshold value of 0.510 (close to the 55th percentile in the sample) is detected. Debt also exhibits diminishing partial effects from 0.183 before the threshold value to 0.133 after it, although neither of the effects is significant. Age and firm size have negative effects on TFP growth in this region. One reason is that most state-owned or above-scale enterprises in the northeastern region were established at earlier stages of development than firms in other regions, because the northeastern area was initially targeted as the nation's industrial base. So, firms in this region are usually older and larger than firms in other areas. But these "advantages" apparently do not contribute to and even have harmful effects on firms' productivity growth.

Finally, firms in the central region have a threshold value of 0.4837 (close to 43th percentile in the sample). However, unlike firms in other regions that exhibit a diminishing partial effect of debt ratio, firms in the central region show an opposite effect: a one percent increase in debt is associated with -0.385 of a percentage change in TFP growth before the threshold and an 0.08 of a percent change after the threshold value, though the latter effect is not significant. Again, firms' cash flow also have highly significant and positive effects on productivity growth.

1.5.3 By Industry

The sample is split according to three broad industry categories in the data: manufacturing, public utility sector, and mining sector. For manufacturing industry, a threshold value of 0.5723 (close to the 55th percentile in the data) is detected. Debt is significantly positively associated with TFP growth in both regimes but again shows diminishing returns passing the threshold value. Age and export are significantly negatively associated with TFP growth, consistent with Fu (2005)'s argument that exported goods in

manufacturing are low-skilled and low-technology content products and thus hinder firms' TFP growth. CF ratio and firm size are all positively associated with firm' growth. For public utility sector, a threshold value of 0.3955 (close to the 37th percentile in the data). Debt is significantly positively correlated with TFP growth in regime 1, with partial effect being 0.4001, while it decreases to 0.0754 and becomes insignificant in regime 2. Firm size is negatively associated with TFP growth in public sector. Public utility sector has the highest level of total asset compared with the other two industries, and apparently it has become a drag for their productivity growth. But due to the sector's special characteristics (i.e., most of them are SOEs or collective firms), its goal is not to make profits or innovations. The other regressors are not statistically significant. As for mining industry, a threshold value of 0.4998 is detected (close to the 44th percentile in the data). The partial effect turns to negative although it becomes insignificant.

1.6 Robustness Check

In this section we conduct three nonparametric tests to evaluate the significance of our nonparametric results for both threshold and gradient estimates. The first is consistent model specification test by Zheng (1996) where the null functional form is parametrically specified against a nonparametric alternative, and the last three tests are modified version of F-test by Ullah (1985) where the null is 1) parametric threshold regression, 2) a nonparametric regression *without* threshold, and 3) a nonparametric regression *with* a specific threshold.¹⁰

1.6.1 Linear Function Specification

The results in Table (1.3) and Table (1.4) indicate nonlinearity of variables not seen in their parametric estimation, therefore implying a potentially mis-specified function with linear restriction. To justify the use of nonparametric regression, we follow Zheng (1996) to test whether functions $m_1(\cdot)$ and $m_2(\cdot)$ in (4) and (5) are truly linear function of $\mathbf{Z}$ (i.e., $m_j(Z) = Z\beta_j$, $j = 1, 2$). Equivalently, we test whether $E(U|Z) = 0$. Using the law of iteration we have the following infeasible test statistic under the null and the alternative:

$$\begin{cases} H_0 : I = E(UE(U|Z)f(Z)) = 0 \\ H_1 : I = E(UE(U|Z)f(Z)) > 0 \end{cases}$$

where $f(\cdot)$ is the density of its argument, which can be consistently estimated by its kernel estimator $\hat{f}(Z) = (1/NT) \sum_{i=1}^N \sum_{t=1}^T K_h(\mathbf{Z}_{it}, \mathbf{Z})$. We estimate I by replacing U with its parametric residuals (i.e., OLS residuals) from either regime 1, regime 2, or whole regime, and replacing the conditional mean with

¹⁰Henderson and Parmeter (2013) conduct similar tests in their paper to test threshold values.

Table 1.4: Nonparametric Results for Whole Sample and Classification by Ownership

Ownership Classification												
Variable	Whole		State				Private		Foreign		Collective	
	$\hat{\gamma} = 0.5122$		$\hat{\gamma} = 0.2931$		$\hat{\gamma} = 0.5085$		$\hat{\gamma} = 0.2903$		$\hat{\gamma} = 0.5154$			
Threshold estimate	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
Debt ratio	0.3610*** (0.0670)	0.1910*** (0.0577)	0.0443 (0.0648)	0.0200 (0.0654)	0.2699** (0.0693)	0.1284** (0.0592)	0.0355 (0.1419)	-0.2892* (0.1514)	0.2198** (0.0933)	0.1957** (0.0810)		
Ln(age)	-0.0715** (0.0309)	0.045* (0.0231)	-0.1417 (0.1817)	0.0467 (0.0547)	-0.0050 (0.0498)	-0.0003 (0.0191)	-0.0651 (0.2761)	-0.1296 (0.0792)	0.0665 (0.0672)	-0.0323 (0.0252)		
Export	0.3046*** (0.0521)	0.1708*** (0.0390)	-3.3582*** (0.8034)	1.4824*** (0.0865)	0.7233*** (0.0900)	-0.0873** (0.0354)	-0.2670 (0.2880)	0.0360 (0.0637)	-0.4513*** (0.1121)	0.2591*** (0.0683)		
CF ratio	0.3231** (0.1301)	1.0461*** (0.1011)	2.7380 (1.8369)	0.7351** (0.3426)	0.0207 (0.2030)	0.3400*** (0.0684)	0.3355 (1.0253)	1.1134*** (0.3307)	0.4231*** (0.1677)	2.5366*** (0.1020)		
Ln(total asset)	0.0171 (0.0156)	0.0236** (0.0114)	-0.0215 (0.0848)	-0.0269*** (0.0090)	0.0399 (0.0255)	0.0316*** (0.0107)	0.0137 (0.0834)	-0.0346 (0.0257)	0.0135 (0.0308)	-0.0060 (0.0183)		
Observations	35,406	39,465	1,943	10,308	10,936	19,690	4,521	12,403	4,128	6,544		
R-squared	0.1404	0.1143	0.2075	0.2094	0.2016	0.1344	0.2527	0.2675	0.2640	0.1815		

Note: The estimated threshold effect of debt ratio ($\hat{\gamma}$) from (8) is reported for whole sample and sub-sample based on ownership. Wild-bootstrapped standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 1.5: Nonparametric Results for Whole Sample and Classification by Region

Variable	Region Classification									
	Whole		Eastern		Western		Central		Northeastern	
	$\hat{\gamma} = 0.5122$		$\hat{\gamma} = 0.5137$		$\hat{\gamma} = 0.3945$		$\hat{\gamma} = 0.4837$		$\hat{\gamma} = 0.5099$	
Threshold estimate	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
Debt ratio	0.3610*** (0.0670)	0.1910*** (0.0577)	0.5927*** (0.0657)	0.1327** (0.0614)	0.9233*** (0.1271)	0.2097** (0.0911)	-0.3851*** (0.1075)	0.0800 (0.0903)	0.1834 (0.1864)	0.1329 (0.1272)
Ln(age)	-0.0715** (0.0309)	0.0453* (0.0231)	-0.0478 (0.0359)	0.0222 (0.0307)	0.1176 (0.1337)	-0.0306 (0.0305)	-0.1069 (0.0347)	0.0466*** (0.0348)	-0.0616 (0.0501)	-0.0018 (0.0486)
Export	0.3046*** (0.0521)	0.1708*** (0.0390)	0.3014*** (0.0564)	0.3843*** (0.0394)	-1.4682*** (0.2416)	-0.8572*** (0.0580)	0.6319*** (0.1415)	-0.2006** (0.0945)	0.6292*** (0.0924)	-0.4266*** (0.0835)
CF ratio	0.3231** (0.1301)	1.0461*** (0.1011)	0.5822*** (0.1429)	0.5465*** (0.1009)	-0.9089 (0.9810)	0.4608 (0.2497)	0.7878*** (0.1653)	0.3719*** (0.1105)	-0.0083 (0.2896)	2.4602*** (0.2568)
Ln(total asset)	0.0171 (0.0156)	0.0236** (0.0114)	0.0540*** (0.0196)	0.0421*** (0.0131)	-0.0074 (0.0691)	-0.0117 (0.0196)	0.0082 (0.0242)	-0.0087 (0.0144)	-0.0069 (0.0269)	-0.0514** (0.0242)
Observations	35,406	39,465	24,874	25,358	2,807	5,833	3,520	4,631	2,561	2,210
R-squared	0.1404	0.1143	0.1693	0.1355	0.2294	0.2189	0.2223	0.1802	0.2977	0.2296

Note: The estimated threshold effect of debt ratio ($\hat{\gamma}$) from (8) is reported for whole sample and sub-sample based on ownership. Wild-bootstrapped standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 1.6: Nonparametric Results for Whole Sample and Classification by Industry

Variable	Industry Classification									
	Whole		Manufacturing		Public Utility		Mining			
	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
Threshold estimate	$\hat{\gamma} = 0.5122$		$\hat{\gamma} = 0.5723$		$\hat{\gamma} = 0.3955$		$\hat{\gamma} = 0.4998$			
Debt ratio	0.3610*** (0.0670)	0.1919*** (0.0577)	0.2164*** (0.0318)	0.0902* (0.0503)	0.4001*** (0.1324)	0.0754 (0.1429)	0.3442*** (0.1443)	-0.0379 (0.0654)		
Ln(age)	-0.0715** (0.0309)	0.045* (0.0231)	-0.0282*** (0.0073)	-0.0173* (0.0098)	0.0071 (0.0450)	-0.0369 (0.0491)	-0.0422 (0.0484)	-0.0510 (0.0432)		
Export	0.3046*** (0.0521)	0.1708*** (0.0390)	-0.0450*** (0.0123)	-0.0459*** (0.0127)	-0.2880 (0.5722)	1.7699 (2.8918)	-0.1878 (0.1941)	-0.4517 (0.2885)		
CF ratio	0.3231** (0.1301)	1.0461*** (0.1011)	0.2729*** (0.0242)	0.4289*** (0.0451)	0.8510** (0.4152)	0.1366 (0.3285)	0.0735 (0.0902)	-0.0129 (0.0623)		
Ln(total asset)	0.0171 (0.0156)	0.0236** (0.0114)	0.0070*** (0.0012)	0.0193*** (0.0098)	-0.0330* (0.0172)	-0.0344* (0.0169)	-0.0867*** (0.0167)	-0.0142 (0.0189)		
Observations	35,406	39,465	37,939	29,939	1,927	3,239	810	1,017		
R-squared	0.1404	0.1143	0.3541	0.3034	0.0934	0.0994	0.0897	0.0972		

Note: The estimated threshold effect of debt ratio ($\hat{\gamma}$) from (8) is reported for whole sample and sub-sample based on ownership. Wild-bootstrapped standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1

the local constant estimator of U on $\mathbf{Z}$:

$$\hat{E}(U|\mathbf{Z}) = \frac{\sum_{i=1}^N \sum_{t=1}^T K_h(\mathbf{Z}_{it}, \mathbf{Z}) U_{it}}{\sum_{i=1}^N \sum_{t=1}^T K_h(\mathbf{Z}_{it}, \mathbf{Z})}$$

We now have our feasible test statistic that leaves out the diagonal terms to eliminate its asymptotic bias:

$$\hat{I} = \frac{1}{N^2|H|} \sum_{i=1}^N \sum_{t=1}^T \sum_{\substack{j=1 \\ \{j,k\} \neq \{i,t\}}}^N \sum_{k=1}^T \hat{u}_{it} \hat{u}_{jk} K_h(\mathbf{Z}_{it}, \mathbf{Z}_{jk}) \quad (1.12)$$

with its variance:

$$\hat{\sigma}^2 = \frac{2}{N^2|H|} \sum_{i=1}^N \sum_{t=1}^T \sum_{\substack{j=1 \\ \{j,k\} \neq \{i,t\}}}^N \sum_{k=1}^T \hat{u}_{it}^2 \hat{u}_{jk}^2 K_h^2(\mathbf{Z}_{it}, \mathbf{Z}_{jk}) \quad (1.13)$$

so we construct the standardized version of test statistic:

$$\hat{J} = (NT\sqrt{|H|}) \frac{\hat{I}}{\hat{\sigma}} \quad (1.14)$$

which converges to the standard normal distribution under the null. Please refer to the appendix for its bootstrapping procedure.

In sum, we are testing whether the unknown function of Z is truly linear in Regime 1, 2, or whole regime without sample splitting. The p-value of testing linear functional form in regime 1, 2, or whole regime is reported in *Left*, *Right*, and *Total*, respectively, in Panel a-c, Table (1.7). We clearly reject the null for total regime without sample splitting for all but state-owned enterprises at 1% significant level.

By splitting the sample into regimes based on estimated threshold, we also see that both left and right regimes of the threshold reject the null of linear model specification in whole sample data. However, various patterns are presented across ownership groups. While we reject linear model in both left and right regime for private enterprises, we cannot reject the null in either side for foreign and state-owned enterprises respectively. For collective enterprises, the left regime fails to reject a linear model specification while the right regime of threshold rejects the linear functional form. For regional groups, only northeastern region fail to reject the linear model specification in either sides. However, since both regimes of eastern, western and central regions reject the linear functional form as well as their whole sample data without splitting, it is justified for the use of nonparametric model rather than for a linear functional form specification. All industry groups reject linear functional form model specification.

1.6.2 Parametric Threshold Regression Specification

To examine if the conditional mean function of debt ratio in our study can be sufficiently explained by the parametric threshold regression in (1), we test null functional form of Hansen's parametric threshold model against our nonparametric threshold regression as alternative. We use the test statistic on the basis of F-type test by Ullah (1985) which compares the residual sum of squares difference between the null (RSS_0) and alternative model (RSS_1):

$$\hat{T} = (RSS_0 - RSS_1)/RSS_1 \quad (1.15)$$

which converges to zero if and only if the null is true. See the appendix for detailed bootstrapped procedure. We report the p-values in " H_0 : Hansen Model" under *Threshold test* in Table (1.7). For the whole sample, we reject the null hypothesis that parametric threshold model is correctly specified. For ownership groups, collective and private enterprises reject the null, while it is not the case for state-owned and foreign firms. For regional groups, only the northeastern group fails to reject the null hypothesis of Hansen's parametric threshold model. The results strongly indicate that a nonparametric threshold regression without function restriction is suitable in our study. Again all industry groups reject the use of Hansen's model.

1.6.3 Testing for the Existence of a Threshold

To evaluate the significance of our estimated threshold, we again adopt the F-type test by Ullah (1985) again to test the null hypothesis where a threshold does not exist in the nonparametric regression. Let $m(\mathbf{Z}_{it}, \gamma) = \alpha(\mathbf{Z}_{it}) + \beta \cdot I(\mathbf{L}_{it} > \gamma)$ defined in (6), we are interested in testing:

$$H_0 : m(\mathbf{Z}_{it}, \gamma) = m(\mathbf{Z}_{it}) \quad (1.16)$$

where the function $m(\mathbf{Z}_{it})$ stands for a unknown function without threshold structure specification (i.e., $E(Y|Z) = m(\mathbf{Z})$). The corresponding test statistic is similar as the previous case:

$$\hat{T} = (RSS_0 - RSS_1)/RSS_1 \quad (1.17)$$

with RSS_0 and RSS_1 be the residual sum of squares under the current null (16) and its alternative. See Appendix 2.3 for detailed bootstrap procedure. We report the p-values in $H_0 : \gamma_{\text{not exist}}$ in Table (1.7) under "Threshold Test". Clearly, for whole sample data the null hypothesis is rejected, strongly indicating the existent of threshold detected for debt ratio. For ownership groups, collective and private

firms reject the null hypothesis, but we did not find evidence of significant threshold value in state-owned and foreign firms. This is expected, since both group of firms fail to reject Hansen's parametric threshold model. For regional groups, only northeastern region fails to reject the null hypothesis of no threshold value, which is again consistent with the test of Hansen's parametric threshold model. For industry groups, only manufacturing industry suggests the existence of threshold value while the other two groups reject the threshold existence.

1.6.4 Testing for Specific Value of a Threshold

Finally, to test if the detected threshold value is a specific value, we conduct the F-type test with the null hypothesis where the debt ratio has its threshold value equal to its median.¹¹ In other words, we test

$$H_0 : \gamma = Debt_{50} \quad (1.18)$$

with our test statistic

$$\hat{T} = (RSS_0 - RSS_1)/RSS_1 \quad (1.19)$$

where RSS_0 and RSS_1 are the residual sum of squares obtained from the null and alternative, respectively. See Appendix 2.4 for detailed bootstrap procedure. We report p-values of the test in $H_0 : \gamma = Debt_{50}$ in Table (1.7), and we clearly reject the null hypothesis for both whole sample and ownership groups, suggesting that debt ratio processes a significant threshold effect not equal to any arbitrary level or the confidence interval is not very wide that contains other values (in this case, the median value in the data) but the detected one from our model. For regional groups, the central and western regions reject the null hypothesis, while eastern and northeastern are the opposite case. The results for the latter two groups are in fact expected, since the estimated threshold values in eastern region (0.5137) and northeastern region (0.5099) are very close to their median values 0.5463 and 0.5056, respectively. For industry groups, only manufacturing reject the null of hypothesis of having threshold at median value.

In sum, the test results indicate that all groups within the sample aside from state-owned enterprises (SOEs), foreign enterprises, firms of northeastern region, and firms in public utility sector and mining sector should use non-parametric threshold analysis to estimate the effect of debt on TFP growth. In other words, first three groups of firms can be estimated more efficiently by parametric threshold regression, and we provide the results in Table 16. The threshold values are: 0.2784 for state-owned enterprises, 0.3284 for foreign firms and 0.4841 for firms in northeastern region. The partial effects of debt are

¹¹The test is conducted because we can not directly compute the confidence intervals of our threshold values. In case that the confidence interval is too wide that contains many values, here we pick the median of debt ratio in the sample to do the test.

different from nonparametric regressions. For SOEs, for instance, the coefficient of debt is positive before the detected threshold but becomes negative afterwards, indicating an inverted U-shaped relationship. It also indicates that China's SOEs have "soft budget constraint" problem (Bai et al. (2016)). Ding et al. (2016b) also poses the possibility that higher leverage in China is likely to result in over-investment, which may ultimately hurt TFP growth. The signs of other variables are consistent with other studies (Chen and Guariglia (2013); Guariglia et al. (2011)). The thresholds detected for foreign firms and firms in the northeastern region are not statistically significant. Firms in public utility sector and mining sector do not exhibit significant threshold effects of debt on TFP growth.

1.6.5 Endogeneity

To cope with endogeneity of debt ratio, we apply a two-stage local polynomial estimator by Su and Ullah (2008) to estimate the partial effects of debt in nonparametric threshold regression. We follow them to consider a triangular equation:

$$y_{it} = m(x_{it}, \mathbf{Z}_{it}^1) + u_{it} \quad (1.20)$$

$$x_{it} = g(\mathbf{Z}_{it}^1, \mathbf{Z}_{it}^2) + \epsilon_{it} \quad (1.21)$$

where x is the endogenous variable, or debt ratio in our case; $\mathbf{Z}^1$ is a vector of exogenous variable: age, cash flow, export and firm's size, and $\mathbf{Z}^2$ contains a relevant and valid instrumental variable. We choose the one-year lag of debt ratio to several studies in the literature (See, for example, Coricelli et al. (2012)). Assume $E(\epsilon|Z^1, Z^2) = 0$ and $E(u|Z^1, Z^2, \epsilon) = E(u|\epsilon)$, we observe that

$$\begin{aligned} E(y|x, \mathbf{Z}^1, \mathbf{Z}^2, \epsilon) &= m(x, \mathbf{Z}^1) + E(u|x, \mathbf{Z}^1, \mathbf{Z}^2, \epsilon) \\ &= m(x, \mathbf{Z}^1) + E(u|x - g(\mathbf{Z}^1, \mathbf{Z}^2), \mathbf{Z}^1, \mathbf{Z}^2, \epsilon) \\ &= m(x, \mathbf{Z}^1) + E(u|\mathbf{Z}^1, \mathbf{Z}^2, \epsilon) \\ &= m(x, \mathbf{Z}^1) + E(u|\epsilon) \end{aligned} \quad (1.22)$$

which is in fact an additive model that takes the form:

$$E(y|x, \mathbf{Z}^1, \mathbf{Z}^2, \epsilon) = M(x, \mathbf{Z}^1, \epsilon) \quad (1.23)$$

Table 1.7: Nonparametric Test Results For Functional Form Specification

Panel a	Collective (0.5154)	Foreign (0.2903)	Ownership group State (0.2931)	Private (0.5085)
Functional form test				
<i>Left</i>	0.6950	0.9150	0.4650	0.0200
<i>Right</i>	0.0875	0.1475	0.9200	0.0750
<i>Total</i>	0.0075	0.0475	0.3225	0.0000
Threshold Test				
H_0 : Hansen model	0.0960	0.9975	0.7789	0.0005
H_0 : $\gamma_{\text{not exist}}$	0.0910	0.5420	0.3770	0.0410
H_0 : $\gamma = Debt_{50}$	0.0954	0.0402	0.0000	0.0704
Panel b	Central (0.4837)	Eastern (0.5137)	Regional group Northeastern (0.5099)	Western (0.3944)
Functional form test				
<i>Left</i>	0.0100	0.0000	0.9900	0.0230
<i>Right</i>	0.0880	0.0025	0.9225	0.0000
<i>Total</i>	0.0050	0.0075	0.0975	0.0200
Threshold Test				
H_0 : Hansen model	0.0450	0.0000	0.1497	0.0201
H_0 : $\gamma_{\text{not exist}}$	0.0910	0.0450	0.8391	0.0000
H_0 : $\gamma = Debt_{50}$	0.0603	0.1658	0.7588	0.0050
Panel c	Manufacturing (0.5723)	Public Utility (0.3955)	Industry group Northeastern (0.4998)	
Functional form test				
<i>Left</i>	0.0000	0.0025	0.0075	
<i>Right</i>	0.0075	0.0700	0.0150	
<i>Total</i>	0.0000	0.0350	0.0200	
Threshold Test				
H_0 : Hansen model	0.0050	0.0350	0.0525	
H_0 : $\gamma_{\text{not exist}}$	0.0175	0.4150	0.2225	
H_0 : $\gamma = Debt_{50}$	0.0550	0.8550	0.4325	
Panel d	Whole Sample (0.5122)			
Functional form test				
<i>Left</i>	0.0000			
<i>Right</i>	0.0000			
<i>Total</i>	0.0000			
Threshold Test				
H_0 : Hansen model	0.0000			
H_0 : $\gamma_{\text{not exist}}$	0.0025			
H_0 : $\gamma = Debt_{50}$	0.0905			

Note: The OLS functional form is tested based on the conditional mean test by Li and Wang (1998) whose null functional form is a correctly specified parametric model. The threshold tests are based on F-type related test by Ullah (1985) whose null functional form is chosen based on our specification: H_0 : *Hansen Model* test the specification of the parametric threshold model by Hansen (2000), H_0 : $\gamma = 0$ tests the validity of fully nonparametric model without threshold, and H_0 : $\gamma = Debt_{50}$ test the significance of a threshold at the median level. All procedures of test implementation are provided in Appendix.

where $M(x, \mathbf{Z}^1, \epsilon) = m(x, \mathbf{Z}^1) + w(\epsilon)$ with $w(\epsilon) \equiv E(u|\epsilon)$. To fully identify the function of our interest $m(x, \mathbf{Z}^1)$, we assume $E(u) = 0$ so that

$$m(x, \mathbf{Z}^1) = \int_{\epsilon} M(x, \mathbf{Z}^1, \epsilon) dF(\epsilon) \quad (1.24)$$

Estimation of (23) involves three steps. We first run a local linear regression of endogenous variable, debt ratio, on lag debt ratio and all the other exogenous covariates to get the residuals $\hat{\epsilon}_{it} = x_{it} - \hat{g}(\mathbf{Z}_{it}^1, \mathbf{Z}_{it}^2)$, where $\hat{g}(\cdot, \cdot)$ is the function estimate by local linear estimator in (11) and $\mathbf{Z}_{it} = [\mathbf{Z}_{it}^1, \mathbf{Z}_{it}^2]$. We denote H_1 as bandwidth used in the first stage.

We then estimate (24) using marginal integration by Linton and Nielsen (1995) by replacing it with its sample analogue:

$$\hat{m}(x, \mathbf{Z}^1) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{M}(x, \mathbf{Z}^1, \hat{\epsilon}_{it}), \quad (1.25)$$

where $\hat{M}(x, \mathbf{Z}^1, \hat{\epsilon})$ is the local constant estimator:

$$\hat{M}(x, \mathbf{Z}^1, \hat{\epsilon}) = \frac{\sum_{i=1}^N \sum_{t=1}^T K_H(W_{it}, W) y_{it}}{\sum_{i=1}^N \sum_{t=1}^T K_H(W_{it}, W)} \quad (1.26)$$

with $W_{it} = [x_{it}, \mathbf{Z}_{it}^1, \hat{\epsilon}_{it}]$, and $W = [x, \mathbf{Z}^1, \hat{\epsilon}]$ is evaluation set of points. Our interested gradients estimate is similar as in (25), except that we replace function estimate with its first-order gradient estimate in local constant estimator:

$$\hat{m}'(x, \mathbf{Z}^1) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{M}'(x, \mathbf{Z}^1, \hat{\epsilon}_{it}) \quad (1.27)$$

and $\hat{M}'(x, \mathbf{Z}^1, \hat{\epsilon})$ is estimated through local constant estimator similar to (26).¹² We purposely choose local linear and local constant estimator in stage 1 and 2 to fulfill the theoretical requirement in Su and Ullah (2008) for consistent estimators of both unknown function and gradient in (22). Please refer to the appendix for its detailed discussion.

Estimation results are reported in Table (1.9), and we only report results for groups whose threshold effects should be estimated nonparametrically based on test results. We see that all qualified groups have a significant nonlinear effect of debt ratio in either regime. In general, the partial effects of debt on TFP growth from instrumental variable estimation are close to the ones from nonparametric regression in Table (1.4), (1.5) and (1.6), except for the declined magnitude and cases where the signs switch. We conclude that for each group, the median impact of debt on firm's productivity growth shrinks a large extent after passing the threshold value. Firms in the western region even exhibit negative partial effect in the second regime with significance, implying excessive borrowing that hurt firms' growth.

¹²For the explicit formula of $\hat{M}'(x, \mathbf{Z}^1, \hat{\epsilon})$ in local constant regression, see Henderson et al. (2015) p. 127.

Table 1.8: Threshold Regression Result by Hansen (2000)

variables	State	Foreign	Northeastern
constant	0.221*** (0.0210)	0.5781*** (0.0033)	0.528*** (0.0041)
Ln(age)	-0.093*** (0.0110)	-0.074* (0.0390)	-0.076*** (0.0320)
exp	-2.339*** (0.0940)	-0.184*** (0.0610)	0.491*** (0.0970)
cfratio	2.214*** (0.0980)	0.477 (0.9610)	-0.139* (0.0770)
ln(asset)	-0.028 (0.0560)	0.2130*** (0.1090)	-0.024 (0.0220)
<i>Threshold Variable:</i>			
Debt Ratio			
Regime 1	0.096*** (0.0341)	-0.392 (0.0830)	-0.172** (0.0810)
Regime 2	-0.122*** (0.0580)	-0.419 (0.1140)	0.221 (0.1730)
Estimated threshold($\hat{\gamma}$)	0.2784***	0.3284	0.4841
R^2	0.0481	0.029	0.0344
n	12,251	16,924	4,771

Table 1.9: Results of Nonparametric Partial Effects of Debt Ratio on TFP Growth by Su and Ullah (2008)

Variable	Whole $\hat{\gamma} = 0.5122$		Private $\hat{\gamma} = 0.5085$		Collective $\hat{\gamma} = 0.5154$		Eastern $\hat{\gamma} = 0.5137$		Western $\hat{\gamma} = 0.3945$		Central $\hat{\gamma} = 0.4837$		Manufacturing $\hat{\gamma} = 0.5723$	
	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
Debt ratio	0.309*** (0.069)	0.132*** (0.050)	0.259*** (0.062)	0.104* (0.056)	0.189*** (0.086)	0.176** (0.079)	0.325*** (0.051)	0.119** (0.051)	0.663*** (0.109)	-0.039* (0.022)	-0.446*** (0.091)	-0.055 (0.078)	0.152*** (0.048)	0.083* (0.049)
Observations	31,704	34,848	9,090	16,382	3,295	4,742	22,319	24,978	2,492	5,178	3,120	4,108	35,294	27,345
R-squared	0.16734	0.12444	0.1935	0.1599	0.221	0.1677	0.154	0.154	0.204	0.203	0.123	0.144	0.2788	0.2463

1.6.6 Debt growth

We have identified that debt is having threshold effect in Chinese industrial firms, with diminishing partial effects after the detected threshold values in most cases. In concern that sometimes debt ratio may not be a good indicator for policy implications (i.e., firms can purchase more assets and borrow even more to maintain certain debt ratios), we also use debt growth as an alternative regressor and implement Henderson et al. (2015).¹³ The estimation results are provided in Table (1.10).¹⁴

Threshold values of 0.1037 and 0.0883 are detected for the whole sample with and without northeastern region, accounting for the 59.3th and 56.7th of respective sample sets. Debt growth, just like debt ratio in previous estimation, exhibits significant threshold effect, posing a large drop in marginal effect after the detected threshold values. Similar threshold effects are also found for firms among different ownership. The respective threshold values of debt growth are: 0.0343 (49th percentile) for SOEs, 0.1035 (57th percentile) for private firms, 0.1137 (60th percentile) for foreign firms, and 0.0673 (57th percentile) for collective firms. Note that the partial effect of debt growth in regime 2 turns to negative for SOEs, suggesting an inverted-U relationship between debt growth and TFP growth and their threshold values are the smallest, consistent with what we found in the previous estimation. The respective threshold values among regions are: 0.1037 (58.3th percentile) for the eastern region, 0.1004 (58.6th percentile) for the western region, 0.1093 (62th percentile) for the central region, and 0.1256 (62th percentile) for the northeastern region. Partial effects diminish to half the magnitude for most regions. But central region shows an inverted-U relationship as the sign of regime 2 is negative. For industries, manufacturing has a threshold of 0.0838 (56th percentile) in debt growth that will result to diminishing partial effect, though the effect is not significant in regime 2. Public utility has a threshold of 0.1138 (61th percentile) with negative yet insignificant partial effect on regime 2. Mining sector detects a threshold of 0.0538 (54th percentile) and both effects are negative with similar magnitude.¹⁵

1.7 Conclusion

Productivity and financial development have become two increasingly important topics in macroeconomics, and their relationship is of particular importance for a rapidly developing economy like China's. In order to examine the impact of financing on growth, this paper uses a panel of 8,319 Chinese industrial firms over the period of 1998-2007 to study the impact of firms' external financing, proxied by debt ratio, on total productivity growth. Based on theories of the finance-growth nexus and empirical findings from

¹³Please refer to appendix for a summary of debt growth statistics across firms.

¹⁴To save space only the partial effects of debt growth on TFP growth are reported.

¹⁵The results should be interpreted with caution, since the functional form and threshold tests are not implemented with this alternative regressor to save computation time.

Table 1.10: Nonparametric Threshold Regression Results: Debt Growth

Threshold estimate	<i>By ownership</i>									
	State		Private		Foreign		Collective			
	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
Variable	r=0.0343		r=0.1035		r=0.1137		r=0.0673			
Debt gr	0.0628** (0.0284)	-0.0720*** (0.0190)	0.0381*** (0.0014)	0.0250*** (0.0018)	0.1694*** (0.0271)	0.0419* (0.0216)	0.025*** (0.0063)	0.0082*** (0.0024)	0.025*** (0.0063)	0.0082*** (0.0024)
Observations	4818	5022	14701	11151	6068	3992	4807	3626	4807	3626
R-squared	0.1722	0.1334	0.2016	0.1344	0.2841	0.2783	0.1038	0.0874	0.1038	0.0874

Threshold estimate	<i>By region</i>					
	Eastern		Western		Central	
	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
Variable	r=0.1037		r=0.1004		r=0.1093	
Debt gr	0.0539*** (0.0036)	0.0209** (0.0091)	0.0621* (0.0320)	0.0327** (0.0146)	0.0157** (0.0069)	-0.0312* (0.0180)
Observations	27541	19713	4483	3137	4497	2714
R-squared	0.1832	0.1637	0.2361	0.2014	0.1937	0.1733
					0.1537	0.1631

Threshold estimate	<i>By Industry</i>			
	Manufacturing		Mining	
	Regime 1	Regime 2	Regime 1	Regime 2
Variable	r=0.0838		r=0.0538	
Debt gr	0.0498** (0.0203)	0.0036 (0.0026)	0.0485** (0.0206)	-0.0699* (0.0421)
Observations	35120	27519	1312	881
R-squared	0.2468	0.2003	0.1073	0.0981
			0.1346	0.0763

Threshold estimate	<i>Whole sample</i>			
	With northeast		Without northeast	
	Regime 1	Regime 2	Regime 1	Regime 2
Variable	r=0.1037		r=0.0883	
Debt gr	0.1051*** (0.0201)	0.0376* (0.0208)	0.0932*** (0.0204)	0.0492** (0.0241)
Observations	39070	27336	35253	26920
R-squared	0.2934	0.2635	0.2045	0.2568

the current literature, we hypothesize a non-monotonic relationship between the firms' debt and TFP growth.

Using a nonparametric threshold regression by Henderson et al. (2015) which permits greater functional flexibility, we find significant threshold effects of firms' debt on TFP growth. The results show that in each sample group, a debt ratio lower than its threshold value exhibits a larger partial effect than that exceeding the threshold, supporting our hypothesis that debt has non-linear and diminishing partial effects on firms' TFP growth. Our results also suggest that the detected threshold values are above the mean and median values in the data, indicating that over half of the observations are borrowing higher than their thresholds. After controlling for endogeneity, firms in the western region show a negative partial effect of debt after the threshold value, indicating an inverted U-shaped relationship between debt and TFP growth and evidence of "excessive" debt. To justify the validity of our nonparametric estimate results, we also conduct tests regarding the functional form mis-specification, existence of a threshold, and specific threshold values. The results are in general consistent with our basic regression outcome, but they also suggest that SOEs, foreign firms, and firms in the northeastern region should use the parametric threshold regression by Hansen (2000) to obtain efficient estimates. The results from Hansen (2000) again confirm the existence of a threshold effect in debt ratio and an inverted-U relationship between debt and TFP growth in SOEs. When debt growth is used to replace debt ratio as an alternative regressor, in concern that firms can manipulate their debt ratios by purchasing more assets thus imposing weaker policy implication, threshold effects with diminishing partial effects are also found in most groups. SOEs and firms in the central region shows negative partial effect after their respective threshold values, implying an inverted-U relationship between debt growth and TFP growth.

While we do not find direct evidence of "excessive" debt except for the western region and SOEs (i.e., the partial effect of debt would become negative after exceeding the threshold value),¹⁶ our results still provide evidence that partial effects of debt on TFP growth are heterogeneous. Specifically, debt's effect is much larger before the threshold value than that after it, supporting the non-monotonic relationship between debt and productivity growth suggested in the literature. The detected threshold values would have policy implications for which level of debt or debt growth would be optimal for firms' TFP growth in China. The mechanism through which firms' debt impacts total factor productivity in China is also an interesting topic for future research.

¹⁶When debt growth is used, we find "excessive" debt for SOEs and firms in the central region.

1.8 Appendixes

Appendix 1 Simulation Results For Henderson et al. (2015)

Since the original paper lacks the simulation result for multivariate case, we in the following Monte-Carlo study demonstrate the consistency of their proposed estimator in both univariate and multivariate setting. We consider the following cross-sectional regression with $x = \{x_i\}_{i=1}^N$ as threshold variable:

$$y_i = \alpha(x_i) + \beta I(x_i > \gamma) + u_i \quad (1.28)$$

where for univariate regression, we let

$$\mathbf{DGP}_1 \equiv \alpha(x_i) = \begin{cases} m_1(x_i) = x_i + x_i^2, & \text{if } x_i \leq \gamma \\ m_2(x_i) = \sin(2\pi x_i) - \beta, & \text{if } x_i > \gamma \end{cases}$$

and

$$\mathbf{DGP}_2 \equiv \alpha(x_i) = \begin{cases} m_1(x_i) = 3e^{x_i} + x_i, & \text{if } x_i \leq \gamma \\ m_2(x_i) = x_i + x_i^3 - \beta, & \text{if } x_i > \gamma \end{cases}$$

For bivariate regression, we have

$$y_i = \alpha(x_i, z_{1i}) + \beta I(x_i > \gamma) + u_i \quad (1.29)$$

and we let

$$\mathbf{DGP}_3 \equiv \alpha(x_i, z_{1i}) = \begin{cases} m_1(x_i, z_{1i}) = x_i + x_i^2 + z_{1i} + z_{1i}^2, & \text{if } x_i \leq \gamma \\ m_2(x_i, z_{1i}) = \sin(2\pi x_i) + \cos(2\pi z_{1i}) - \beta, & \text{if } x_i > \gamma \end{cases}$$

and

$$\mathbf{DGP}_4 \equiv \alpha(x_i, z_{1i}) = \begin{cases} m_1(x_i, z_{1i}) = x_i^3 + x_i e^{z_{1i}}, & \text{if } x_i \leq \gamma \\ m_2(x_i, z_{1i}) = x_i \cos(2\pi z_{1i}) - \beta, & \text{if } x_i > \gamma \end{cases}$$

Finally, for trivariate regression we have

$$y_i = \alpha(x_i, z_{1i}, z_{2i}) + \beta I(x_i > \gamma) + u_i \quad (1.30)$$

and we let

$$\mathbf{DGP}_5 \equiv \alpha(x_i, z_{1i}, z_{2i}) = \begin{cases} m_1(x_i, z_{1i}, z_{2i}) = x_i + x_i^2 + z_{1i} + z_{1i}^2 + z_{2i} + z_{2i}^2, & \text{if } x_i \leq \gamma \\ m_2(x_i, z_{1i}, z_{2i}) = \sin(2\pi x_i) + \cos(2\pi z_{1i}) + \sin(\pi z_{2i}) - \beta, & \text{if } x_i > \gamma \end{cases}$$

and

$$\mathbf{DGP}_6 \equiv \alpha(x_i, z_{1i}, z_{2i}) = \begin{cases} m_1(x_i, z_{1i}, z_{2i}) = x_i^3 + x_i e^{z_{1i}} + z_{1i} e^{z_{2i}}, & \text{if } x_i \leq \gamma \\ m_2(x_i, z_{1i}, z_{2i}) = x_i \cos(2\pi z_{1i}) + \sin(\pi z_{2i}) - \beta, & \text{if } x_i > \gamma \end{cases}$$

and each random variable $w \in [x, z_1, z_2]$ are independently drawn from normal distribution with mean zero and variance one. Error term u is normally distributed with mean zero and variance 0.5. We set the true value of β_0 from $[0, 0.5, 1]$, and true value of threshold γ_0 from $[-1, 0, 1]$.

We search γ over a 50 equal space grid points in x (i.e., Len=50). Bandwidth h_j for variable w_j is chosen based on rule of thumb $h_j = 1.059 * \hat{\sigma}_{x_j}^2 n^{-1/(4+d)}$, where $\hat{\sigma}_{x_j}^2$ is the sample variance of its variable, and $d = [1, 2, 3]$ is the variable dimension. Sample size changes from 100, 200, to 400, and we repeat each simulation 500 times (i.e., Sim=500) to evaluate the average of the bias of β and γ , and the absolute squared error (ASE) for function estimate $\hat{m}_1(\cdot)$ and $\hat{m}_2(\cdot)$. The results are reported in the Table (1.11)-(1.16) .

The results clearly shows that at the ASE of both smoothing functions $m_1(\cdot)$ and $m_2(\cdot)$ increases when dimension is higher, indicating the curse of dimensionality. However, for all chosen value of threshold level γ , its bias decreases toward zero rapidly as sample size increases.

Table 1.11: Simulation Results for **Univariate**: Sim=500, len=50

DGP_1									
n	β_0	γ_0	Bias(β)	Bias(γ)	$m_1(\cdot)$ -ASE	$m_2(\cdot)$ -ASE	Total-ASE	β_0 -ASE	γ_0 -ASE
100	0.0	-1.0	0.6651	0.1097	0.0698	0.0272	0.1395	0.2652	0.1295
200	0.0	-1.0	0.3769	0.1061	0.0572	0.0259	0.1267	0.2574	0.1259
400	0.0	-1.0	0.0593	0.0800	0.0462	0.0230	0.1011	0.2449	0.1184
100	0.5	-1.0	0.1923	0.1100	0.0690	0.0529	0.1423	0.2233	0.1298
200	0.5	-1.0	0.1453	0.1085	0.0567	0.0507	0.1251	0.2104	0.1235
400	0.5	-1.0	0.1070	0.1002	0.0460	0.0479	0.1026	0.1803	0.1212
100	1.0	-1.0	-0.3095	0.1094	0.0707	0.1279	0.1412	0.2262	0.1294
200	1.0	-1.0	-0.0610	0.1076	0.0561	0.1259	0.1256	0.1812	0.1205
400	1.0	-1.0	-0.0252	0.1005	0.0461	0.1226	0.1016	0.1321	0.1168
100	0.0	0.0	-1.0291	0.0561	0.0321	0.0214	0.0232	0.2547	0.0352
200	0.0	0.0	-0.8588	0.0540	0.0247	0.0207	0.0155	0.2482	0.0311
400	0.0	0.0	-0.7864	0.0501	0.0178	0.0187	0.0021	0.2218	0.0291
100	0.5	0.0	-1.5214	0.0560	0.0323	0.0393	0.0231	0.3750	0.0348
200	0.5	0.0	-1.2873	0.0532	0.0244	0.0367	0.0141	0.3370	0.0300
400	0.5	0.0	-1.0525	0.0416	0.0182	0.0305	0.0027	0.3110	0.0275
100	1.0	0.0	-2.1665	0.0541	0.0319	0.1082	0.0224	0.5782	0.0321
200	1.0	0.0	-1.2009	0.0345	0.0246	0.1073	0.0165	0.5199	0.0300
400	1.0	0.0	-0.9965	0.0124	0.0180	0.1069	0.0031	0.4988	0.0287
100	0.0	1.0	-1.2154	0.0115	0.0337	0.0222	-0.0371	0.2073	0.0089
200	0.0	1.0	-1.0393	0.0041	0.0264	0.0207	-0.0215	0.1662	0.0064
400	0.0	1.0	-1.0021	0.0027	0.0192	0.0195	-0.0208	0.1489	0.0037
100	0.5	1.0	-1.7118	0.0116	0.0351	0.0469	-0.0362	0.3478	0.0085
200	0.5	1.0	-1.1223	0.0112	0.0256	0.0456	-0.0329	0.3138	0.0062
400	0.5	1.0	-0.9115	0.0051	0.0185	0.0424	-0.0282	0.2951	0.0028
100	1.0	1.0	-2.2111	0.0105	0.0336	0.1203	-0.0380	0.5432	0.0081
200	1.0	1.0	-2.0297	0.0095	0.0266	0.1183	-0.0227	0.5176	0.0068
400	1.0	1.0	-1.9071	0.0059	0.0191	0.1141	-0.0200	0.4902	0.0032

Table 1.12: Simulation Results for **Univariate**: Sim=500, len=50

<i>DGP₂</i>									
n	β_0	γ_0	Bias(β)	Bias(γ)	$m_1(\cdot)$ -ASE	$m_2(\cdot)$ -ASE	Total-ASE	β_0 -ASE	γ_0 -ASE
100	0.0	-1.0	-0.1612	-0.0151	0.0408	0.0740	0.0669	0.2683	0.0211
200	0.0	-1.0	-0.1302	-0.0080	0.0405	0.0521	0.0500	0.2566	0.0011
400	0.0	-1.0	-0.1272	-0.0037	0.0320	0.0384	0.0386	0.2234	0.0006
100	0.5	-1.0	-0.2124	-0.0147	0.0417	0.4604	0.0669	0.4621	0.0220
200	0.5	-1.0	-0.1836	-0.0078	0.0398	0.4220	0.0513	0.4604	0.0009
400	0.5	-1.0	-0.1161	-0.0043	0.0318	0.3872	0.0388	0.4503	0.0006
100	1.0	-1.0	-0.2614	-0.0122	0.0428	1.3468	0.0679	0.7176	0.0295
200	1.0	-1.0	-0.2550	-0.0080	0.0381	1.2869	0.0497	0.7062	0.0009
400	1.0	-1.0	-0.2472	-0.0030	0.0345	1.2338	0.0383	0.6931	0.0006
100	0.0	0.0	-0.2640	0.0058	0.0569	0.1498	0.0995	0.7050	0.0010
200	0.0	0.0	-0.2512	0.0038	0.0503	0.1120	0.0847	0.6877	0.0006
400	0.0	0.0	-0.2402	0.0026	0.0498	0.0923	0.0744	0.6820	0.0005
100	0.5	0.0	-0.3129	-0.0189	0.0579	0.6259	0.0956	0.9858	0.0010
200	0.5	0.0	-0.3105	0.0014	0.0568	0.5745	0.0866	0.9701	0.0006
400	0.5	0.0	-0.2879	0.0012	0.0500	0.5334	0.0766	0.9589	0.0005
100	1.0	0.0	-0.3635	-0.0011	0.0566	1.6122	0.0938	1.3287	0.0012
200	1.0	0.0	-0.3113	0.0009	0.0520	1.5272	0.0850	1.1113	0.0006
400	1.0	0.0	-0.2993	0.0005	0.0457	1.4625	0.0748	0.8959	0.0005
100	0.0	1.0	-0.6407	-0.0129	0.1094	0.5688	0.1540	1.1777	0.0022
200	0.0	1.0	-0.5425	0.0014	0.0890	0.5102	0.1473	1.1052	0.0007
400	0.0	1.0	-0.4322	0.0012	0.0774	0.4697	0.1116	1.0365	0.0006
100	0.5	1.0	-0.6915	-0.0140	0.0950	1.2056	0.1554	1.8540	0.0053
200	0.5	1.0	-0.5861	0.0033	0.0888	1.0829	0.1474	1.7508	0.0007
400	0.5	1.0	-0.4869	0.0005	0.0729	1.0817	0.1334	1.0543	0.0006
100	1.0	1.0	-0.5397	-0.0104	0.1139	2.3504	0.1566	1.5288	0.0019
200	1.0	1.0	-0.4418	0.0036	0.0832	2.3008	0.1486	1.0462	0.0008
400	1.0	1.0	-0.3305	0.0004	0.0710	2.2498	0.1302	0.9789	0.0005

Table 1.13: Simulation Results for **Bivariate**: Sim=500, len=50

DGP_3									
n	β_0	γ_0	Bias(β)	Bias(γ)	$m_1(\cdot)$ -ASE	$m_2(\cdot)$ -ASE	Total-ASE	β_0 -ASE	γ_0 -ASE
100	0.0	-1.0	2.4470	0.0596	0.0961	0.1674	0.0876	0.6520	0.0035
200	0.0	-1.0	2.2138	0.0217	0.0884	0.1667	0.0816	0.5141	0.0005
400	0.0	-1.0	2.0374	0.0156	0.0489	0.1379	0.0806	0.4337	0.0002
100	0.5	-1.0	2.0093	0.0546	0.0790	0.3659	0.0857	0.4680	0.0030
200	0.5	-1.0	1.7381	0.0177	0.0530	0.3490	0.0817	0.3339	0.0003
400	0.5	-1.0	1.5702	0.0159	0.0452	0.3265	0.0807	0.2662	0.0002
100	1.0	-1.0	1.5235	0.0530	0.1927	1.0537	0.0803	0.2847	0.0028
200	1.0	-1.0	1.2357	0.0235	0.1226	1.0345	0.0784	0.1785	0.0006
400	1.0	-1.0	1.0385	0.0189	0.0494	0.7076	0.0626	0.1263	0.0002
100	0.0	0.0	1.2135	1.2901	-0.0819	0.2339	0.0756	0.1855	1.6645
200	0.0	0.0	0.8593	0.3467	-0.0692	0.2205	0.0655	0.0867	0.1202
400	0.0	0.0	0.6154	0.0368	-0.0655	0.1813	0.0597	0.0434	0.0014
100	0.5	0.0	0.7036	1.2907	-0.0831	0.4350	0.0774	0.0843	1.6660
200	0.5	0.0	0.3907	0.3661	-0.0720	0.4071	0.0639	0.0304	0.1340
400	0.5	0.0	0.1065	0.0376	-0.0619	0.3589	0.0592	0.0060	0.0014
100	1.0	0.0	0.2830	1.3082	-0.0823	1.1277	0.0757	0.0585	1.7113
200	1.0	0.0	-0.1318	0.3371	-0.0739	1.0985	0.0631	0.0153	0.1136
400	1.0	0.0	-0.0899	0.0379	-0.0655	1.0337	0.0593	0.0106	0.0014
100	0.0	1.0	1.9279	0.1605	-0.1023	0.9062	0.0962	0.4341	0.0257
200	0.0	1.0	1.0613	0.0697	-0.0905	0.8619	0.0717	0.3833	0.0049
400	0.0	1.0	0.8409	0.0168	-0.0833	0.7706	0.0625	0.3589	0.0003
100	0.5	1.0	1.3814	0.1506	-0.1034	1.1284	0.0935	0.2602	0.0227
200	0.5	1.0	1.0099	0.0786	-0.0915	1.0731	0.0811	0.2373	0.0062
400	0.5	1.0	0.8307	0.0213	-0.0864	1.0227	0.0719	0.1965	0.0005
100	1.0	1.0	0.9600	0.1424	-0.1022	1.8204	0.0936	0.1555	0.0203
200	1.0	1.0	0.8889	0.0606	-0.0911	1.7342	0.0711	0.1154	0.0037
400	1.0	1.0	0.8237	0.0155	-0.0867	1.6696	0.0628	0.0879	0.0002

Table 1.14: Simulation Results for **Bivariate**: Sim=500, len=50

DGP_4									
n	β_0	γ_0	Bias(β)	Bias(γ)	$m_1(\cdot)$ -ASE	$m_2(\cdot)$ -ASE	Total-ASE	β_0 -ASE	γ_0 -ASE
100	0.0	-1.0	0.1020	0.1131	0.1036	0.3289	0.0518	0.3096	0.1428
200	0.0	-1.0	0.0858	0.1023	0.0936	0.2298	0.0454	0.2863	0.1351
400	0.0	-1.0	0.0818	0.0941	0.0867	0.1871	0.0251	0.2785	0.1248
100	0.5	-1.0	0.0623	0.1137	0.1036	0.6279	0.0521	0.2570	0.1438
200	0.5	-1.0	0.0426	0.1097	0.0980	0.5288	0.0453	0.2163	0.1304
400	0.5	-1.0	0.0252	0.0706	0.0920	0.2919	0.0254	0.1851	0.1188
100	1.0	-1.0	0.0105	0.1101	0.1057	1.4239	0.0519	0.2423	0.1333
200	1.0	-1.0	0.0025	0.1070	0.0991	1.0261	0.0453	0.1837	0.1308
400	1.0	-1.0	0.0010	0.0815	0.0555	0.7956	0.0155	0.1576	0.1217
100	0.0	0.0	0.0384	0.0736	0.0803	0.3070	0.0332	0.3114	0.0627
200	0.0	0.0	0.0272	0.0612	0.0676	0.2922	0.0255	0.3016	0.0426
400	0.0	0.0	0.0109	0.0199	0.0560	0.2359	0.0156	0.2844	0.0205
100	0.5	0.0	0.0079	0.0761	0.0840	0.5585	0.0335	0.3136	0.0672
200	0.5	0.0	-0.0049	0.0560	0.0668	0.5503	0.0255	0.2975	0.0517
400	0.5	0.0	-0.0007	0.0221	0.0547	0.5042	0.0155	0.2520	0.0333
100	1.0	0.0	-0.0494	0.0744	0.0782	1.3118	0.0331	0.3699	0.0635
200	1.0	0.0	-0.0309	0.0500	0.0686	1.3025	0.0255	0.3208	0.0111
400	1.0	0.0	-0.0221	0.0315	0.0560	1.2460	0.0154	0.2868	0.0026
100	0.0	1.0	-0.1473	0.0121	0.0582	0.3542	0.1171	0.3180	0.0074
200	0.0	1.0	-0.1101	0.0017	0.0459	0.3022	0.0763	0.2623	0.0024
400	0.0	1.0	-0.1080	0.0003	0.0339	0.1965	0.0243	0.2276	0.0003
100	0.5	1.0	-0.2001	0.0113	0.0608	0.6585	0.1173	0.5173	0.0059
200	0.5	1.0	-0.1983	0.0013	0.0454	0.6124	0.0802	0.4302	0.0021
400	0.5	1.0	-0.1033	0.0005	0.0352	0.5483	0.0116	0.4127	0.0004
100	1.0	1.0	-0.2474	0.0007	0.0618	1.4564	0.1175	0.7254	0.0070
200	1.0	1.0	-0.1504	0.0005	0.0469	1.4270	0.1007	0.6587	0.0019
400	1.0	1.0	-0.0399	0.0002	0.0335	1.3931	0.0105	0.6243	0.0003

Table 1.15: Simulation Results for **Trivariate**: Sim=500, len=50

DGP_5									
n	β_0	γ_0	Bias(β)	Bias(γ)	$m_1(\cdot)$ -ASE	$m_2(\cdot)$ -ASE	Total-ASE	β_0 -ASE	γ_0 -ASE
100	0.0	-1.0	1.3632	1.1219	0.1033	0.3753	0.6004	0.3106	0.1395
200	0.0	-1.0	1.2393	1.0597	0.0929	0.3223	0.5546	0.2756	0.1183
400	0.0	-1.0	1.0373	0.6349	0.0646	0.2371	0.4784	0.2683	0.1095
100	0.5	-1.0	0.7971	1.0882	0.1112	0.7100	0.6017	0.2384	0.1335
200	0.5	-1.0	0.6426	1.0612	0.0986	0.6635	0.5583	0.1777	0.1191
400	0.5	-1.0	0.4532	0.6443	0.0736	0.5666	0.4790	0.1597	0.1022
100	1.0	-1.0	0.3565	1.0695	0.1061	1.5436	0.6011	0.2257	0.1285
200	1.0	-1.0	0.1713	1.0620	0.0847	1.4918	0.5538	0.1285	0.1178
400	1.0	-1.0	-0.1188	0.7337	0.0745	1.3958	0.4794	0.0859	0.1094
100	0.0	0.0	1.5347	0.7853	0.1060	0.3756	0.4419	0.3252	0.0751
200	0.0	0.0	1.2110	0.6150	0.0977	0.3099	0.4037	0.3019	0.0680
400	0.0	0.0	0.6744	0.5758	0.0773	0.2072	0.3519	0.2964	0.0468
100	0.5	0.0	0.9781	0.7978	0.1020	0.7012	0.4405	0.2693	0.0742
200	0.5	0.0	0.5945	0.6005	0.0982	0.6299	0.4026	0.2090	0.0632
400	0.5	0.0	0.2929	0.4667	0.0478	0.5208	0.3440	0.1698	0.0553
100	1.0	0.0	0.4084	0.8246	0.1026	1.5190	0.4389	0.2474	0.0768
200	1.0	0.0	0.1178	0.7921	0.0919	1.4594	0.4053	0.1736	0.0543
400	1.0	0.0	-0.0995	0.4481	0.0803	1.3339	0.3440	0.1124	0.0323
100	0.0	1.0	-1.7045	-0.2420	0.0711	0.3751	0.2514	0.3667	0.0220
200	0.0	1.0	-1.4377	-0.0788	0.0697	0.2417	0.2372	0.3449	0.0050
400	0.0	1.0	-0.8550	-0.0102	0.0508	0.1144	0.2192	0.3189	0.0005
100	0.5	1.0	-2.1643	-0.2272	0.0699	0.7669	0.2502	0.5174	0.0224
200	0.5	1.0	-1.8749	-0.0929	0.0502	0.6551	0.2382	0.4947	0.0063
400	0.5	1.0	-1.4745	-0.0143	0.0400	0.5231	0.2179	0.4030	0.0008
100	1.0	1.0	-2.7463	-0.2360	0.0692	1.6324	0.2489	0.6646	0.0230
200	1.0	1.0	-2.4484	-0.0781	0.0589	1.5436	0.2393	0.6351	0.0049
400	1.0	1.0	-1.9106	-0.0081	0.0448	1.4481	0.2212	0.6033	0.0005

Table 1.16: Simulation Results for **Trivariate**: Sim=500, len=50

<i>DGP₆</i>									
n	β_0	γ_0	Bias(β)	Bias(γ)	$m_1(\cdot)$ -ASE	$m_2(\cdot)$ -ASE	Total-ASE	β_0 -ASE	γ_0 -ASE
100	0.0	-1.0	2.9962	-0.0488	0.2309	0.0466	0.3684	1.0031	0.1412
200	0.0	-1.0	2.7158	-0.0105	0.0640	0.0246	0.3550	0.7678	0.0346
400	0.0	-1.0	2.5212	-0.0076	0.0324	-0.0049	0.3264	0.6486	0.0007
100	0.5	-1.0	2.5000	-0.0343	0.2089	0.2252	0.3706	0.7436	0.1062
200	0.5	-1.0	2.1725	-0.0100	0.0612	0.2073	0.3576	0.5068	0.0203
400	0.5	-1.0	2.0383	-0.0056	0.0426	0.1757	0.3233	0.4283	0.0007
100	1.0	-1.0	2.0331	-0.0395	0.2201	0.1052	0.3702	0.5247	0.1433
200	1.0	-1.0	1.6857	-0.0164	0.0691	0.0800	0.3562	0.3085	0.0229
400	1.0	-1.0	1.5366	-0.0103	0.0425	-0.0572	0.3225	0.2466	0.0007
100	0.0	0.0	1.1015	-0.4872	-0.2074	0.1360	0.3451	0.5040	1.6261
200	0.0	0.0	0.8359	-0.3724	-0.0971	0.0912	0.3218	0.2887	0.6966
400	0.0	0.0	0.5682	-0.2386	-0.0364	0.0203	0.2810	0.1388	0.1773
100	0.5	0.0	0.6150	-0.4075	-0.1168	0.3265	0.3413	0.4519	1.6124
200	0.5	0.0	0.3708	-0.3740	-0.0968	0.2712	0.3162	0.2298	0.7102
400	0.5	0.0	0.1026	-0.1463	-0.0479	0.2095	0.2784	0.1159	0.1019
100	1.0	0.0	0.1284	-0.4563	-0.1141	0.1243	0.3413	0.4262	0.6540
200	1.0	0.0	-0.1071	-0.2562	-0.0968	0.0718	0.3198	0.2064	0.4551
400	1.0	0.0	-0.0603	-0.1831	-0.0534	-0.0645	0.2781	0.1297	0.0124
100	0.0	1.0	-1.5469	-0.3376	-0.0693	0.6099	0.3007	0.6882	1.0484
200	0.0	1.0	-1.3511	-0.0546	-0.0472	0.3928	0.2636	0.4699	0.2238
400	0.0	1.0	-1.1494	0.0179	-0.0195	0.1768	0.2296	0.4043	0.0020
100	0.5	1.0	-1.9687	-0.3717	-0.0744	0.8757	0.2961	0.8870	1.0938
200	0.5	1.0	-1.4194	-0.0612	-0.0378	0.6849	0.2653	0.6901	0.2535
400	0.5	1.0	-0.9760	0.0180	-0.0160	0.4801	0.2273	0.5869	0.0018
100	1.0	1.0	-2.4281	-0.3288	-0.0705	1.6488	0.2990	1.0864	0.9938
200	1.0	1.0	-1.9864	-0.0152	-0.0492	1.4678	0.2681	0.9748	0.1423
400	1.0	1.0	-0.9332	0.0105	-0.0102	1.2798	0.2305	0.8860	0.0024

Appendix. 2.1 Bootstrap procedure for model specification test by Zheng (1996)

1. Let $\mathbf{Z} = [L, \mathbf{X}]$, and given the original sample $\{y_{it}, \mathbf{Z}_{it}\}_{i=1, t=1}^{N, T}$, calculate $\hat{J}$.
2. Generate centered bootstrapped residual $\{\hat{u}_{it}^*\}_{i, t=1}^{N, T}$, where for each observation $i = 1, \dots, N$ and $t = 1, \dots, T$, $\hat{u}_{it}^* = \frac{1-\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$ and $\hat{u}_{it}^* = \frac{1+\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1-\sqrt{5}}{2\sqrt{5}}$, where $\bar{\hat{u}}$ refers to the mean of $\hat{u}$.
3. Compute $\hat{J}^*$ using the bootstrap sample $\{y_{it}^*, \mathbf{Z}_{it}\}_{i=1, t=1}^{N, T}$, where $y_{it}^* = \mathbf{Z}_{it}\hat{\beta}_{OLS} + \hat{u}_{it}^*$, and $\hat{\beta}_{OLS}$ are the OLS gradient estimates to impose the null.
4. Repeat the step 2-3 B_1 times to obtain $\hat{J}_1^*, \hat{J}_2^*, \dots, \hat{J}_{B_1}^*$. Reject H_0 if $\hat{J} > \hat{J}_{1-\alpha}^*$, where $\hat{J}_{1-\alpha}^*$ is the upper $(1 - \alpha)$ percentile value of its empirical distribution with α the significance level.

We set $\alpha = 0.05$ and $B_1 = 399$, and implement the adaptive rule-of-thumb bandwidth selection (Silverman, 1986) to reduce computation burden.

Appendix. 2.2 Bootstrap procedure for testing parametric threshold function model specification by Ullah (1985)

1. Given the original sample $\{y_{it}, L_{it}, \mathbf{X}_{it}\}_{i=1, t=1}^{N, T}$, obtain RSS_0 from Hansen parametric threshold regression model and RSS_1 from nonparametric threshold regression model. Then calculate $\hat{T}$.
2. Generate centered bootstrapped residual $\{\hat{u}_{it}^*\}_{i, t=1}^{N, T}$, where for each observation $i = 1, \dots, N$ and $t = 1, \dots, T$, $\hat{u}_{it}^* = \frac{1-\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$ and $\hat{u}_{it}^* = \frac{1+\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1-\sqrt{5}}{2\sqrt{5}}$, where $\bar{\hat{u}}$ refers to the mean of $\hat{u}$. Then construct the bootstrapped independent variable $y_{it}^* = \hat{\alpha}_1 L_{it} I(L_{it} \leq \hat{\gamma}) + \hat{\alpha}_2 L_{it} I(L_{it} > \hat{\gamma}) + \mathbf{X}_{it}\hat{\beta} + u_{it}^*$, where the estimated parameters $[\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}]$ are from estimation in (1). Call $\{y_{it}^*, L_{it}, \mathbf{X}_{it}\}_{i, t=1}^{N, T}$ the bootstrapped sample.
3. Compute $\hat{T}^*$ in the same way as in step 1, except we replace the original sample $\{y_{it}, L_{it}, \mathbf{X}_{it}\}_{i=1, t=1}^{N, T}$ with the bootstrapped sample $\{y_{it}^*, L_{it}, \mathbf{X}_{it}\}_{i, t=1}^{N, T}$.
4. Repeat the step 2-3 B_1 times to obtain $\hat{T}_1^*, \hat{T}_2^*, \dots, \hat{T}_{B_1}^*$. Reject H_0 if $\hat{T} > \hat{T}_{1-\alpha}^*$, where $\hat{T}_{1-\alpha}^*$ is the upper $(1 - \alpha)$ percentile value of its empirical distribution with α the significance level.

We set $\alpha = 0.05$ and $B_1 = 399$, and implement the adaptive rule-of-thumb bandwidth selection to reduce computation burden.

Appendix. 2.3 Bootstrap procedure for testing the existence of threshold by Ullah (1985)

1. Let $\mathbf{Z} = [L, \mathbf{X}]$, and given the original sample $\{y_{it}, \mathbf{Z}_{it}\}_{i=1, t=1}^{N, T}$ to obtain RSS_0 from nonparametric regression model without threshold, and using the original sample $\{y_{it}, L_{it}, \mathbf{Z}_{it}\}_{i=1, t=1}^{N, T}$ to obtain RSS_1 from nonparametric threshold regression model. Then calculate $\hat{T}$.
2. Generate centered bootstrapped residual $\{\hat{u}_{it}^*\}_{i, t=1}^{N, T}$, where for each observation $i = 1, \dots, N$ and $t = 1, \dots, T$, $\hat{u}_{it}^* = \frac{1-\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$ and $\hat{u}_{it}^* = \frac{1+\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1-\sqrt{5}}{2\sqrt{5}}$, where $\bar{\hat{u}}$ refers to the mean of $\hat{u}$. Then construct the bootstrapped independent variable $y_{it}^* = \hat{m}(\mathbf{Z}_{it}) + u_{it}^*$, where the function estimate $\hat{m}(\cdot)$ estimated by local linear estimator in (11). Call $\{y_{it}^*, \mathbf{Z}_{it}\}_{i, t=1}^{N, T}$ the bootstrapped sample.
3. Compute $\hat{T}^*$ in the same way as in step 1, except we replace the original sample $\{y_{it}, \mathbf{Z}_{it}\}_{i=1, t=1}^{N, T}$ with the bootstrapped sample $\{y_{it}^*, \mathbf{Z}_{it}\}_{i, t=1}^{N, T}$.
4. Repeat the step 2-3 B_1 times to obtain $\hat{T}_1^*, \hat{T}_2^*, \dots, \hat{T}_{B_1}^*$. Reject H_0 if $\hat{T} > \hat{T}_{1-\alpha}^*$, where $\hat{T}_{1-\alpha}^*$ is the upper $(1 - \alpha)$ percentile value of its empirical distribution with α the significance level.

We set $\alpha = 0.05$ and $B_1 = 399$, and implement the adaptive rule-of-thumb bandwidth selection to reduce computation burden.

Appendix. 2.4 Bootstrap procedure for testing specific value of the threshold by Ullah (1985)

1. Let $\mathbf{Z} = [L, \mathbf{X}]$, and given the original sample $\{y_{it}, L_{it}, \mathbf{Z}_{it}\}_{i=1, t=1}^{N, T}$ to obtain RSS_0 from nonparametric regression model with threshold equal to the median of debt ratio, and to obtain RSS_1 from nonparametric threshold regression model with threshold equal to the estimated value. Then calculate $\hat{T}$.
2. Generate centered bootstrapped residual $\{\hat{u}_{it}^*\}_{i, t=1}^{N, T}$, where for each observation $i = 1, \dots, N$ and $t = 1, \dots, T$, $\hat{u}_{it}^* = \frac{1-\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$ and $\hat{u}_{it}^* = \frac{1+\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1-\sqrt{5}}{2\sqrt{5}}$, where $\bar{\hat{u}}$ refers to the mean of $\hat{u}$. Then construct the bootstrapped independent variable $y_{it}^* = \hat{\alpha}(\mathbf{Z}_{it}) + \hat{\beta} \cdot I(\mathbf{Z}_{it} > \hat{\gamma}_m) + u_{it}^*$, where all estimates $[\hat{\alpha}(\cdot), \hat{\beta}]$ are from nonparametric regression model where $\hat{\gamma} = \hat{\gamma}_m$. Call $\{y_{it}^*, \mathbf{Z}_{it}\}_{i, t=1}^{N, T}$ the bootstrapped sample.
3. Compute $\hat{T}^*$ in the same way as in step 1, except we replace the original sample $\{y_{it}, \mathbf{Z}_{it}\}_{i=1, t=1}^{N, T}$ with the bootstrapped sample $\{y_{it}^*, \mathbf{Z}_{it}\}_{i, t=1}^{N, T}$.
4. Repeat the step 2-3 B_1 times to obtain $\hat{T}_1^*, \hat{T}_2^*, \dots, \hat{T}_{B_1}^*$. Reject H_0 if $\hat{T} > \hat{T}_{1-\alpha}^*$, where $\hat{T}_{1-\alpha}^*$ is the upper $(1 - \alpha)$ percentile value of its empirical distribution with α the significance level.

We set $\alpha = 0.05$ and $B_1 = 399$, and implement the adaptive rule-of-thumb bandwidth selection to reduce computation burden.

Appendix. 3 Theoretical Justification for the Order of Polynomial Regression in Su and Ullah (2008)

To illustrate our choice of the order of polynomial regression in the two-stage procedure in section 6.5, consider the following triangular system equation:

$$y = m(x, Z_1) + u \quad (1.31)$$

$$x = g(Z_1, Z_2) + \epsilon \quad (1.32)$$

where $x \in \mathbb{R}^1$ is the endogenous regressor so that $E(u|x) \neq 0$, $Z_1 \in \mathbb{R}^{q_1}$ is a vector of exogenous regressors, and $Z_2 \in \mathbb{R}^{q_2}$ is a vector of instrumental variables. Denote $h_1 \in \mathbb{R}^{q_x+q_1}$ and $h_2 \in \mathbb{R}^{q_1+q_2}$ be a set of bandwidth used in the first and second stage, respectively, and let p_1 and p_2 be the order of polynomial regression in the first and second stage, respectively. We note that local constant or local linear is equivalent to the polynomial regression of order zero or one.

Assumption A5 and Remark 5 in Su and Ullah (2008) derives the specific rate of convergence of h_1 and h_2 that allows for consistent function and gradient estimates of $m(x, Z_1)$ to be obtained. Specifically, we have:

$$h_2 \propto n^{-\frac{1}{\gamma_2}} \quad (1.33)$$

$$h_1 \propto n^{-\frac{1}{\gamma_1}} \quad (1.34)$$

where $\gamma_2 = 2(p_2 + 1) + q_1 + 1$ and $\gamma_1 = 2(p_1 + 1) + (q_1 + q_2)$. Furthermore, let $\alpha = 1/\gamma_1$, they show that the consistent estimator also requires the choice of p_1 and p_2 to be such that $h_1 \propto n^{-\alpha}$, where

$$\underline{\alpha} < \alpha < \bar{\alpha} \quad (1.35)$$

with $\underline{\alpha} = \frac{1}{\gamma_2} \max\{\frac{p_2+1}{p_1+1}, \frac{p_2+3}{2(p_1+1)}\}$ and $\bar{\alpha} = \frac{1}{\gamma_2} \left(\frac{p_2+q_1}{q_1+q_2} \right)$. In other words, we need to choose p_1 and p_2 to ensure h_1 to converges at a rate bounded by the required interval. Since we have $q_x = 1, q_1 = 4, q_2 = 1$, our choice of $p_1 = 1$ (i.e., local linear estimator) and $p_2 = 0$ (i.e., local constant estimator) is theoretically justified. This can be seen from Figure 1, which plots the dimension of Z_1 and α . The green and blue dot lines represents $\underline{\alpha}$ and $\bar{\alpha}$, respectively. Clearly, our choice of p_1 and p_2 is valid so long as $q_1 > 3$, which is true in our case.

Figure 1.1: The rate of convergence h_1 given $p_1 = 1$ and $p_2 = 0$

Note: q_1 is the dimension of exogenous variables in (39) and $\alpha = 1/\gamma_1$ defined in (42). The value of α given each value of q_1 is calculated by setting $p_1 = 1$ and $p_2 = 0$.

Appendix 4 Chinese Industrial Classification Codes

CICC	Industry Name
6	Exploration and absterion of coal and char industry
7	Exploration of oil and natural gas industry
8	Picking of ferrous metal mine industry
9	Picking of non-ferrous metal mine industry
10	Picking of nonmetal mine industry
11	Other mining industry
13	Agriculture and food processing industry
14	Foodstuff manufacturing industry
15	Soft drink manufacturing industry
16	Tobacco manufacturing industry
17	Textile industry
18	Waving costume, shoes and cap manufacturing industry
19	Leather, fur and feather manufacturing industry
20	Wood working, and wood,bamboo,bush rope,palm, and straw manufacturing industry
21	Furniture manufacturing industry
22	Paper making and paper products industry
23	Print and copy of record vehicle industry
24	Stationary and sporting goods manufacturing industry
25	Oil processing, coking and nuclear manufacturing industry
26	Chemical material and chemical product manufacturing industry
27	Medicine manufacturing industry
28	Chemical fiber manufacturing industry
29	Rubber product industry
30	Plastics product industry
31	Nonmetallic mineral product industry
32	Ferrous metal refining and calendaring processing industry
33	Non-ferrous metal refining and calendaring processing industry
34	Metal product industry
35	Universal equipment manufacturing industry
36	Task equipment manufacturing industry
37	Transport and communication facilities manufacturing industry
39	Electric machine and fittings manufacturing industry
40	Communication apparatus ,computer and other electric installation manufacturing industry
41	Instrument and meter, stationery machine manufacturing industry
42	Handicraft and other manufacturing industry
43	Removal and processing of obsolete resource and material industry
44	Electricity and thermal manufacturing and supplying industry
45	Combustion gas manufacturing and supplying industry
46	Water manufacturing and supplying industry

Source: NBSC

Appendix 5 Geographic-Based Region classification

Eastern Region	Beijing Zhejiang	Tianjin Shandong	Hebei Fujian	Shanghai Guangdong	Jiangsu Hainan
Northeastern Region	Liaoning	Jilin	Heilongjiang		
Central Region	Shanxi Hunan	Anhui	Jiangxi	Henan	Hubei
Western Region	Neimenggu Yunnan Ningxia	Guangxi Tibet Xinjiang	Chongqing Shanxi	Sichuan Gansu	Guizhou Qinghai

Note: Region Classification is on the basis of National Bureau of Statistics of China

Appendix 6 Debt growth summary

	10%	25%	50%	75%	90%	mean
whole sample	-0.3176	-0.112	0.0416	0.2825	0.7378	0.5041
<i>By ownership</i>						
state	-0.235	-0.071	0.0392	0.2109	0.5591	0.372
private	-0.2738	-0.0905	0.0543	0.2996	0.7349	0.3603
foregin	-0.3967	-0.1645	0.0344	0.3165	0.8393	0.7239
collective	-0.293	-0.0978	0.0263	0.229	0.6411	0.3393
<i>By region</i>						
eastern	-0.3173	-0.1165	0.0438	0.2866	0.7225	0.4377
western	-0.3256	-0.1033	0.0434	0.2869	0.781	0.9073
northeastern	-0.3501	-0.1316	0.0334	0.2839	0.8471	0.3173
central	-0.2887	-0.0813	0.031	0.248	0.7835	0.6228
<i>By industry</i>						
manufacturing	-0.3216	-0.1153	0.0413	0.2857	0.7415	0.5101
public utility	-0.2	-0.0523	0.0548	0.257	0.7378	0.3505
mining	-0.2515	-0.0812	0.0363	0.226	0.648	0.4765

Data Source: NBSC

Appendix 7 Percentage of firm ownership types across regions

	Eastern region	West region	Central region	Northeastern region
State firms	9%	30%	30%	20%
Private firms	40%	27%	40%	27%
Collective firms	13%	11%	18%	10%
Foreign firms	34%	28%	8%	40%

Data Source: NBSC

Bibliography

- Aghion, P., Howitt, P., Mayer-Foulkes, D., 2005. The effect of financial development on convergence: Theory and evidence. *The Quarterly Journal of Economics* 120 (1), 173–222.
- Arcand, J. L., Berkes, E., Panizza, U., 2015. Too much finance? *Journal of Economic Growth* 20 (2), 105–148.
- Bai, Y., Lu, D., Tian, X., 2016. Do financial frictions explain chinese firm saving and misallocation? working paper.
- Beck, T., Levine, R., Loayza, N., 2000. Finance and the sources of growth. *Journal of financial economics* 58 (1-2), 261–300.
- Butler, A. W., Cornaggia, J., 2011. Does access to external finance improve productivity? evidence from a natural experiment. *Journal of Financial Economics* 99 (1), 184–203.
- Cecchetti, S. G., Kharroubi, E., 2012. Reassessing the impact of finance on growth. working paper.
- Chen, M., Guariglia, A., 2013. Internal financial constraints and firm productivity in china: Do liquidity and export behavior make a difference? *Journal of Comparative Economics* 41 (4), 1123–1140.
- Chow, G. C., et al., 2004. Economic reform and growth in china. *ANNALS OF ECONOMICS AND FINANCE*. 5, 93–118.
- Coricelli, F., Driffield, N., Pal, S., Roland, I., 2012. When does leverage hurt productivity growth? a firm-level analysis. *Journal of international Money and Finance* 31 (6), 1674–1694.
- Curtis, C. C., 2013. Economic reforms and the evolution of chinas tfp in the state and the private sectors. University of Richmond.
- Demirgüç-Kunt, A., Maksimovic, V., 1998. Law, finance, and firm growth. *The Journal of Finance* 53 (6), 2107–2137.

- Ding, S., Guariglia, A., Harris, R., 2016a. The determinants of productivity in chinese large and medium-sized industrial firms, 1998–2007. *Journal of Productivity Analysis* 45 (2), 131–155.
- Ding, S., Knight, J., Zhang, X., 2016b. Does china overinvest? evidence from a panel of chinese firms. *The European Journal of Finance*, 1–23.
- Easterly, W., Levine, R., 2001. It's not factor accumulation: stylized facts and growth models. *World Bank Economic Review* 15 (2), 177–219.
- Fu, X., 2005. Exports, technical progress and productivity growth in a transition economy: a non-parametric approach for china. *Applied economics* 37 (7), 725–739.
- Gang, F., Hope, N. C., 2013. The role of state-owned enterprises in the chinese economy. China-United States Exchange Foundation, US-China 2022.
- Gatti, R., Love, I., 2008. Does access to credit improve productivity? evidence from bulgaria. *Economics of Transition* 16 (3), 445–465.
- Gennaioli, N., Shleifer, A., Vishny, R., 2012. Neglected risks, financial innovation, and financial fragility. *Journal of Financial Economics* 104 (3), 452–468.
- Gomis, R. M., Khatiwada, S., 2017. Debt and productivity: Evidence from firm-level data. Tech. rep., Graduate Institute of International and Development Studies Working Paper.
- Guariglia, A., Liu, X., Song, L., 2011. Internal finance and growth: Microeconometric evidence on chinese firms. *Journal of Development Economics* 96 (1), 79–94.
- Hale, G., Long, C., 2011. If you try, you'll get by: Chinese private firms' efficiency gains from overcoming financial constraints. working paper.
- Hall, R. E., Jones, C. I., 1999. Why do some countries produce so much more output per worker than others? *The quarterly journal of economics* 114 (1), 83–116.
- Hansen, B. E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica: Journal of the econometric society*, 413–430.
- Hansen, B. E., 1999. Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of econometrics* 93 (2), 345–368.
- Hansen, B. E., 2000. Sample splitting and threshold estimation. *Econometrica* 68 (3), 575–603.

- Harless, D. W., Peterson, S. P., 1998. Investor behavior and the persistence of poorly-performing mutual funds. *Journal of economic behavior & organization* 37 (3), 257–276.
- Henderson, D., Parmeter, C., 2013. Threshold estimation in nonparametric regression: Government debt and economic growth. In: 4th CREA Rimini Workshop on Developments in Macro and Growth, May. pp. 23–24.
- Henderson, D. J., Papageorgiou, C., Parmeter, C. F., 2013. Who benefits from financial development? new methods, new evidence. *European Economic Review* 63, 47–67.
- Henderson, D. J., Parmeter, C. F., Su, L., 2015. Nonparametric threshold regression: Estimation and inference. working paper.
- Hsieh, C.-T., Klenow, P. J., 2009. Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics* 124 (4), 1403–1448.
- Jayaratne, J., Strahan, P. E., 1996. The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics* 111 (3), 639–670.
- Jin, M., Zhao, S., Kumbhakar, S. C., 2015. Financial constraints and firm productivity: Evidence from chinese manufacturing. working paper.
- Ketteni, E., Mamuneas, T., Savvides, A., Stengos, T., 2007. Is the financial development and economic growth relationship nonlinear? *Economics Bulletin* 15 (14), 1–12.
- King, R. G., Levine, R., 1993. Finance, entrepreneurship and growth. *Journal of Monetary economics* 32 (3), 513–542.
- Kornai, J., Maskin, E., Roland, G., 2003. Understanding the soft budget constraint. *Journal of economic literature* 41 (4), 1095–1136.
- Lang, L., Ofek, E., Stulz, R., 1996. Leverage, investment, and firm growth. *Journal of financial Economics* 40 (1), 3–29.
- Levine, O., Warusawitharana, M., 2014. Finance and productivity growth: Firm-level evidence. working paper.
- Levine, R., 2005. Finance and growth: theory and evidence. *Handbook of economic growth* 1, 865–934.
- Levine, R., Zervos, S., 1998. Stock markets, banks, and economic growth. *American economic review*, 537–558.

- Levinsohn, J., Petrin, A., 2003. Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies* 70 (2), 317–341.
- Li, H., 2015. Leverage and productivity. Unpublished manuscript: Stanford University.
- Linton, O., Nielsen, J. P., 1995. A kernel method of estimating structured nonparametric regression based on marginal integration. *Biometrika*, 93–100.
- Ma, Y., Tang, H., Zhang, Y., 2014. Factor intensity, product switching, and productivity: Evidence from chinese exporters. *Journal of International Economics* 92 (2), 349–362.
- Maliszewski, W., Arslanalp, S., Caparusso, J., Garrido, J., Guo, S., Kang, J. S., Lam, R., Law, D., Liao, W., Rendak, N., Wingender, P., Yu, J., Zhang, L., 2016. Resolving China’s Corporate Debt Problem. International Monetary Fund.
- Matemilola, B., Karimi, M., Bany-Arifin, A., McGowan, C., 01 2014. Non-linear effect of debt on firm value: Dynamic panel threshold evidence 11, 39–48.
- Midrigan, V., Xu, D. Y., 2014. Finance and misallocation: Evidence from plant-level data. *American economic review* 104 (2), 422–58.
- Modigliani, F., Miller, M. H., 1958. The cost of capital, corporation finance and the theory of investment. *The American economic review* 48 (3), 261–297.
- Nucci, F., Pozzolo, A., Schivardi, F., 2005. Is firms productivity related to its financial structure? evidence from microeconomic data. *Rivista di politica economica* 95 (1), 269–290.
- Pagan, A., Ullah, A., 1999. *Nonparametric econometrics*. Cambridge university press.
- Petrin, A., Poi, B. P., Levinsohn, J., et al., 2004. Production function estimation in stata using inputs to control for unobservables. *Stata journal* 4, 113–123.
- Pushner, G. M., 1995. Equity ownership structure, leverage, and productivity: Empirical evidence from japan. *Pacific-Basin Finance Journal* 3 (2-3), 241–255.
- Rahaman, M. M., 2011. Access to financing and firm growth. *Journal of Banking & Finance* 35 (3), 709–723.
- Rajan, R. G., 2006. Has finance made the world riskier? *European Financial Management* 12 (4), 499–533.
- Rajan, R. G., Zingales, L., 1998. Financial dependence and growth. *The American Economic Review* 88 (3), 559–586.

- Rioja, F., Valev, N., 2004a. Does one size fit all?: a reexamination of the finance and growth relationship. *Journal of Development economics* 74 (2), 429–447.
- Rioja, F., Valev, N., 2004b. Finance and the sources of growth at various stages of economic development. *Economic Inquiry* 42 (1), 127–140.
- Song, Z., Storesletten, K., Zilibotti, F., 2011. Growing like china. *American Economic Review* 101 (1), 196–233.
- Su, L., Ullah, A., 2008. Local polynomial estimation of nonparametric simultaneous equations models. *Journal of Econometrics* 144 (1), 193–218.
- Ullah, A., 1985. Specification analysis of econometric models. *Journal of quantitative economics* 1, 187–209.
- Xiao, G., 2005. Nonperforming debts in chinese enterprises: Patterns, causes, and implications for banking reform. *Asian Economic Papers* 4 (3), 61–113.
- Zheng, J. X., 1996. A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics* 75 (2), 263–289.
- Zhou, Y., Hall, J., 2017. The impact of marketization on entrepreneurship in china: Recent evidence. Technical Rport 17-22, West Virginia University.
- Zhu, X., Chen, Y., Feng, C., 2018. Green total factor productivity of china’s mining and quarrying industry: A global data envelopment analysis. *Resources Policy*.

Chapter 2

Recasting the Trade Impact on Labor Share: A Fixed Effect Semiparametric Estimation Study

2.1 Introduction

Intensified global trade since the 1970s has been associated with an array of positive effects on national income (see A. Giles and Williams (2000); Giles and Williams (2000) for a survey), while the declining share of labor income relative to national income has raised economists' concerns (Blanchard et al. (1997)).¹ Figure 3.1 documents the conditional density of labor share across 94 countries during 1970-2009 using a kernel density estimator.² The average global labor share has declined from 0.70 in 1970 to 0.54 in 2009.

Declined labor share has been attributed to the impact of trade openness ($(\text{export}+\text{import})/\text{GDP}$) by several studies since the 2000s. Due to deteriorating labor bargaining power against capital use in the presence of trade-induced capital-biased technology progress and increasing trade-induced competition since the 1970s, early studies documented a negative impact of trade openness on labor share across large sample countries (Rodríguez and Ortega (2001)). The results hold after controlling for the effects of the exchange rate crisis (Harrison (2005)), capitalization (Piketty (2015)), labor union and employment protection (Guscina (2006); Jaumotte and Tytell (2007); Guerriero and Sen (2012); Young and Law-

¹The declining trend happens in developed countries (Stockhammer et al. (2008)) as well as developing countries (Charpe (2011)).

²See the data description in Section 3.

Figure 2.1: Cross-country Labor Share Joint Density: 1970-2009

Date Source: Penn World Table 8.1. *Method:* author's calculation based on kernel smoothing conditional joint density estimates.

son (2014)), government spending (Harrison (2005); Jayadev (2007)), FDI flows (Decreuse and Maarek (2017)), and social and political globalization (Young and Lawson (2014); Young and Tackett (2017)). Particularly for developed countries, the negative impact of trade on labor share is often used as evidence against trading with developing countries. For example, U.S. president Donald Trump continuously blamed trading with China for “stealing” American workers’ jobs when he ran his presidential campaign in 2016. It raises policy concerns as to whether or not a nation should put more restrictions on trading with other countries.

Nonetheless, two main problems surface regarding the estimated impact of trade openness on labor share. First, most cross-country empirical studies conceal the individual effect of exports or imports. Taylor (2013) points out that labor share is falling in both trade-surplus and trade-deficit countries. He argues that when the trend cuts across all countries, it is likely that the underlying cause also cuts across all countries. In this paper, we look at the “cause cutting across all countries” by arguing that the partial effect of exports may be different from that of imports in both magnitude and sign due to their underlying economic nature, irrespective of economic structure and trade pattern. On the one hand, an increase in imported goods with higher quality and greater variety may intensify competition with domestic products. The resulting higher competition would decrease mark-ups over wages and monopoly

power. Since a low degree of monopoly is attributed to the faster growth of real wages, higher import levels may improve labor wage and thus national labor share (Kalecki (1938); Azmat et al. (2012)). Higher competition would also increase the return on human capital, therefore increasing compensation for highly educated or skilled workers (Diwan (2001); Daudey and García-Peñalosa (2007)). On the other hand, a rise in exports may bring about capital-augmenting innovation and mechanization, resulting in labor-saving technology and thus a lower labor share (Bernard et al. (2003); Acemoglu (2003)). This effect may be more severe in countries with abundant, unskilled labor and weak labor bargaining power protection. However, no established theoretical framework exists in the literature that explicitly reconciles the impact of exports and imports on labor share, therefore leaving the empirical evidence unexplored.

Secondly, the existing literature possibly restricts the impact of trade openness in a pre-specified linear functional form, which yields constant partial effects. If trade variables affect labor share beyond a simple linear fashion, however, the empirical estimation would potentially incur the cost of model mis-specification and thus limit its ability to inform policy implication.

This paper addresses the above issues by contributing to the literature two-fold. First, we disclose the individual nonlinear effect of export and import share on labor share that may not be fully revealed in the literature. Buch et al. (2008) and Luo and Zhang (2010) separately test the roles of export and import using regional (provincial) level data in Germany and Italy (China). While their empirical results are generally consistent with ours, they did not discuss the potential economic reasons of the individual roles of exports and imports. We propose two hypotheses that provide potential economic mechanisms for how export and import may affect labor share differently and nonlinearly. We further split our sample into developed (OECD) countries and developing (non-OECD) countries to examine the heterogeneous impact of trade on labor share in countries with different economic development.

Secondly, it serves as the first paper in this literature to implement semiparametric regression to avoid mis-specifying the functional form of trade variables. The standard partially linear model (PLM) is used to take into account certain variables that are potentially nonlinear in the regression. While both export and import shares are the (potentially) nonlinear variables we are interested in, in the presence of limited sample size we nonparametrically estimate only one nonlinear variable in PLM to alleviate the curse of dimensionality.³ The linear coefficients are estimated via Robinson (1988), and the nonlinear component is estimated by a local linear estimator after taking into account the predicted linear values. However, the main concern with this model is its inability to estimate the fixed effect due to the identification problem.

To resolve this issue, we employ a new technique of fixed effect partially linear model (FEPLM)

³See the detailed discussion of the empirical methodology in Section 3.

developed by Qian and Wang (2012). Their model eliminates the fixed effect through first-order difference, and treats the unknown function as an additive model estimated via marginal integration by Linton and Nielsen (1995). Since the consistent estimators in FEPLM rely on large sample data, we provide FEPLM results in a whole sample study and a standard PLM in sub-sample studies. As we will see, both models disclose similar results and reveal the trade impact masked under the conventional linear model estimation.

The remaining of this paper is organized as follows. Section 2 presents a literature review that offers the motivation to study labor share and summarizes the existing papers in terms of their interested variables, implemented methodologies, and main findings. Section 3 discusses two hypotheses based on the existing literature that describe how exports and imports may affect labor share differently and nonlinearly. Section 4 in turn details data and methodologies used to empirically test the hypotheses, whose results are shown in Section 5 from the fixed effect panel OLS regression and semiparametric models (PLM and FEPLM). Section 6 concludes with outlines for future research.

2.2 Literature Review

It had been a widely-held belief that the capital and labor share of national income is fixed.⁴ As Keynes (1939) notes:

...I mean the stability of the proportion of the national dividend accruing to labour, ir-respectively apparently of the level of output as a whole and of the phase of the trade cycle. This is one of the most surprising, yet best-established, facts in the whole range of economic statistics...it is the stability of the ratio for each country which is chiefly remarkable, and this appears to be a long-run, and not merely a short-period, phenomenon (1939, p. 48).

Not until the late 1980s was the labor share issue revisited. Blanchard et al. (1997) documented a significant decline in labor share in Italy, Spain, France, and Germany since the early 1980s. Subsequent research was motivated to examine the underlying determinants for the declined labor share across countries. Table 2.1 summarizes the literature with highlights of their interested variables, implemented methodology, and key findings. We see that trade openness has been frequently tested to account for the declined labor share. Rodríguez and Ortega (2001) document a significantly *positive* effect of trade openness on *capital* share using 176 countries during 1960-1999. Harrison (2005) explicitly tests the trade openness impact on labor share among 100 countries during 1960-1997, while controlling for other effects

⁴We note that the functional distribution of income between labor and capital is the central economic element in the writing of classical economists (i.e., Smith (1776); Ricardo (1817); Marx and Engels (1902))

such as the nominal exchange rate crisis. She shows that a one percentage rise in trade openness is associated with a decrease in labor share of 0.071%- 0.087%, all else being equal.

The negative effect of trade openness is also found in OECD countries. Guscina (2006) argues the shift from labor-biased technology towards capital-biased technology as a significant factor that lowers labor share in industrial countries. By using compensation share and employment share as two different proxies for labor share, he found that the IT revolution in the 1980s significantly boosted capital-augmenting technology, which gradually weakened the reliance of production on labor. This effect is also strengthened through increasing trade in industrial countries, resulting in deteriorating labor share. While the effects of labor market policies and unionization have endeavored to push income toward labor, they did not outweigh that of trade openness. The estimated coefficients of trade openness generally confirm the argument, ranging roughly from -0.128 to -0.140 before and after the IT revolution, respectively.⁵

Jaumotte and Tytell (2007) proxy the effect of trade using terms of trade in 18 OECD countries during 1982-2002 to show that labor share in most OECD countries is negatively associated with an increase in exports price and a decrease in imports price. The result again implies a negative impact of trade openness. Finally, Decreuse and Maarek (2017) find that wage rigidities play an essential role. As opposed to free-wage markets, countries that implement wage rigidities would allocate more resources towards capital-intensive sectors, rather than labor-intensive sectors. Examining eight OECD countries that engaged in trade with China during 1970-2005, they show that labor share is significantly lower in countries with wage-rigidity, but not in free-wage countries, such as the US and UK.

Nonetheless, trade openness presents an opposite impact in later studies with more effects controlled and new datasets. Guerriero and Sen (2012) show that the control of real income per capita GDP in an empirical model predicts a positive and significant effect of trade openness on labor share during 1970-2009 among, at most, 89 countries. A similar result is found in Decreuse and Maarek (2017), whose study focuses on the share of FDI net-flow to GDP. By controlling for the level of income per capita, they also find that trade openness significantly increases labor share among 93 countries during 1980-2000.

In short, the existing literature exhibits mixed empirical findings of the impact of trade openness along with two major caveats. First, the individual effects of exports and imports remain unclear. Table 2.1 shows that only Buch et al. (2008) and Luo and Zhang (2010) separately examine the role of export and import shares. Their findings show that exports and imports affect labor share differently in sign and magnitude, although their analysis is based on Germany and Italy (for the former), and China (for the latter). However, the underlying mechanisms of exports and imports effects on labor share have not been investigated. Also, trade openness, exports, and imports are assumed *a priori* to possess a

⁵However, their model omitted several important explanatory variables, such as capital schedule, population, and labor regulation. The effect of trade openness also varies both in magnitude or sign when different labor share proxies are used.

Table 2.1: Summary of Selected Existing Studies: Dependent variable is labor share

Author(s)	Year	Country	Period	Type	Key Regressors	Trade Var.	Method	Results
Blanchard et al. (1997)	1997	7	1963-96	developed	—	—	time series	labor share declined
Rodríguez and Ortega (2001)	2001	176	1960-99	developed/ing	labor growth; price growth; crisis; pop; gov; m1/gdp	openness	OLS	openness: neg
Diwan (2001)	2001	121-133	1975-90	developed/ing	—	—	OLS	crisis: neg
Gollin (2002)	2002	—	—	—	—	—	—	self-employment compensation needed
Bentolila and Saint-Paul (2003)	2003	14	1973-93	OECD	K/L; emp; output volatility	—	OLS	capital schedule: neg
Harrison (2005)	2005	100	1960-97	developed/ing	$\ln NER$; FDI; K; Y_{ij} ; P_{ij} /L crisis; K	openness	OLS; IV	openness: neg
Guscina (2006)	2006	18	1960-00	developed	L/Y; Y/L; FDI/Y; K/Y; union	openness	OLS	openness: neg
Daudey and García-Peñalosa (2007)	2007	39	1970-94	developed/ing	$\ln(\text{pgdp})$; civil liberties; HC	openness	OLS	openness: neg
Jayadev (2007)	2007	62-89	1973-76	developed/ing	$\log(Y)$; openness tax; crisis; gov;	capital account	OLS	openness: neg
Jaumotte and Tytell (2007)	2007	18	1982-02	OECD	L-reg; L/K; union	openness	DOLS	openness: neg
Buch et al. (2008)	2008	2	1970/91-05	developed	K; L; Y	P_{ex}/P_{im} exs; ims	VAR; DOLS	exs: neg/pos ims: neg
Rodriguez and Jayadev (2010)	2010	57-136	1950-05	developed/ing	—	—	—	Labor share declined
Decreuse and Maarek (2017)	2017	8	1970-05	OECD	K/Y; legal system	openness	OLS; IV	openness: neg
Aznat et al. (2012)	2012	18	1970-01	OECD	barriers; K/Y; TFP; conflict	—	OLS	privatization: neg
Lawless and Whelan (2011)	2011	15	1979-01	OECD	inflation	—	OLS; GMM	inflation:none
Guerriero and Sen (2012)	2012	89	1970-09	developed/ing	Y; Y^2 ; L-reg; K; HC; pop	openness	OLS; IV	openness: pos
Hutchinson and Persyn (2012)	2012	12	1970-2005	OECD	intermediate input price share of skill/unskilled Labor	openness	OLS	openness: mixed
Karabarbounis and Neiman (2013)	2013	59	1975-2012	developed/ing	unemployment; TFP; wage; mark-up; concentration ratio	—	Calibration	P_K/P_C : pos
Brada (2013)	2013	—	—	—	P_K/P_C	—	—	oil price: neg
Young and Lawson (2013)	2014	93	1970-09	developed/ing	Economic freedom index	—	OLS; IV	economic freedom index: pos
Luo and Zhang (2010)	2014	China	1993-07	developing	HC; Y/L; union; K/L; K/L; TFP; FDI; SOE; HC	exs, ims	OLS	exs: neg ims: pos
Young and Tackett (2017)	2017	125	1970-09	developed/ing	KOF social & political index	openness + (FDI) + (Portfolio)	OLS; IV	openness: pos/neg
Young and Zuleta (2015)	2017	US	1983-05	developed	union; K/L; L; conflicts	—	OLS; IV GMM	union: pos

Note: The results are based on authors' counting. Symbols for variables: Y_{ij} is GDP per capita of country i relative to country j; L is labor; K is capital stock; pgdp is per capita income; A is productivity level; P is price level; Y is national income; P_{ex} (P_{im}) is the price index of exports (imports); gov is government expenditure share; pop is population growth rate; $\ln NER$ is the nominal exchange rate in natural log; HC is human capital; L-reg is labor market regulation related policies; pos (neg) is positive (negative) significant correlation. Symbols for econometric techniques (skipped OLS, GMM and VAR): DOLS is dynamic OLS;

constant partial effect in linear regressions, which may be incorrectly specified and thus produces biased estimates. Hence, we discuss in the following that the effects of exports and imports on labor share 1) may be significantly different in sign, 2) may be nonlinear with their levels, and 3) may be different across countries with different economic development.

2.3 Potential Impact of Exports and Imports

Given the existing empirical and theoretical results in the literature, we argue that the individual and nonlinear effects of exports and imports are likely to be the result of the following four major mechanisms and their interrelated interaction: Stolper-Samuelson theorem (Ohlin (1952); Stolper and Samuelson (1941)),⁶ technical shift, increasing competition, and the change in labor's bargaining power.⁷

First, import may positively affect labor share mainly through increasing competition. One explanation is that rising domestic market competition against imported goods with relatively high quality and greater variety would constrain monopoly power in the goods markets. The rising competition in turn lowers the degree of monopoly by decreasing mark-ups, so that the real wage of labor may begin increasing (Kalecki (1938); Azmat et al. (2012);). For developing countries in particular, a massive amount of capital-intensive imported goods may also displace the domestic capital-intensive goods in manufacturing sectors, which is also in favor of labor income (Zhang and Lu (2014)). In addition, Goldberg et al. (2009) demonstrate that trade spurs import for both consumption goods and production inputs that were previously unavailable in domestic markets. As a result, the technology constraint is relaxed so that domestic productivity rises, resulting in more job creation and higher labor share (Azmat et al. (2012); Dinopoulos and Segerstrom (1999)). However, as import grows and domestic firms face more intense competition and lower mark-ups, they are likely to resort to cost-saving or labor-saving production to reduce costs and increase profits (Hung and Hammett (2016)). As a result, labor's bargaining power may be gradually weakened. Hence, the positive effect of imports would decline as imports increase. Therefore, we propose the first hypothesis regarding the nonlinearity of *import* effect:

Hypothesis 1 *The effect of imports on labor share is positive, but diminishes as imports increase.*

In contrast, export may negatively affect labor share mainly through technology shifts. On one hand, trade induces innovation and mechanization. Exporters are documented to be more productive, and more skill- and capital-intensive (Bernard et al. (2003); Melitz (2003)). Hence, technological progress in exporting industries would be capital-augmenting, which lowers the income share of labor (Acemoglu

⁶Given the theorem's strict assumptions, weak support by empirical evidence, and the fact that labor share is decreasing for both developed and developing countries, we argue that the effects are mainly from the latter three mechanisms.

⁷We refer readers to Guerriero and Sen (2012) for a detailed summary of each mechanism.

(2003)). This negative effect may be more intensified in developing countries endowed with labor-intensive production sectors and export-favored policies. During 1970-1990, for instance, export sectors in developing countries were populated by less-skilled and poorly-paid labor, facing less labor-protection policies, and producing non-sophisticated goods (Lawrence (2008)). On the other hand, more factor resources are allocated towards export sectors as export share increases, resulting in a transformation of trade specification that enables exporting industries to produce highly sophisticated exported goods (Boggio et al. (2010); Lawrence (2008); Keller (1998, 2002)). This trade-specification transformation would increase labor demand for high-skill workers, followed by higher bargaining power of workers and higher real wage as compensation. Since high-skill labor and capital are more complementary, an increase in export may generate a less negative (or even positive) effect on labor share. Therefore, we propose the second hypothesis regarding the nonlinearity of *exports* impact:

Hypothesis 2 *The effect of exports on labor share is negative, but diminishes as export increases.*

Despite the intuitive relevance on labor share, no well-established theoretical framework exists that formulates the above two effects on the functional income distribution. We thus proceed to test *Hypothesis 1* and *2* empirically, and offer new perspectives for whether the engagement in trade should always be perceived as a detrimental factor on labor share across countries.

2.4 Data and Methodology

2.4.1 Data and parametric model specification

We strive to estimate a robust measure of labor share. While UNIDO is the major dataset in the existing literature (Diwan (2001); Harrison (2005); Jayadev (2007); Guerriero and Sen (2012)), its inability to account for self-employment compensation would upwardly bias the results (Bentolila and Saint-Paul (2003)). We employ the newest measure of country-specific labor share from the Penn World Table (PWT) 8.1. Based on United Nations Accounting Statistics, PWT 8.1 adjusts the unobserved labor compensation of self-employment using a “best estimate” that includes three different adjustment approaches. The dataset has been shown to have superior properties to perform a cross-country labor share comparison (Inklaar and Timmer (2013)).

Even so, the “best estimated” labor share is used by substituting the missing values with interpolated values during the early part of the series in several developing/poor countries. This issue is presented in Figure 3.2. We therefore perform a robustness check by estimating our benchmark panel model that excludes those countries with interpolated values more than five years. The results reported in Appendix 5 are generally consistent with our findings of trade variables, both in sign and the level of significance,

Figure 2.2: Constant labor share issue in PWT 8.1

Source: Penn World Table 8.1.

Figure 2.3: Pooled OLS fitted line in export and import share against labor share

Source: Penn World Table 8.1.

indicating that empirical results are not sensitive to interpolated data in the superior estimation of labor share.⁸

First, we estimate the following benchmark fixed effect linear panel model (i.e., the parametric model), with the export or import share of GDP as our primary interested variables:

$$\ln(LS)_{it} = \alpha_1 + \alpha_2(Exs)_{it} + \alpha_3(Im s)_{it} + X_{it}\beta + \gamma_i + \epsilon_{it} \quad (2.1)$$

where $\ln(LS)$ is the natural-log labor share, calculated as the employment compensation share of GDP of country i at time period t . Variables Exs and $Im s$ stand for, respectively, the merchandise exports and imports share of GDP. Variable γ_i gives cross-section fixed effect, and ϵ_{it} is the mean-zero error term. To empirically test *Hypothesis 1 and 2*, we consider estimating the effect of exports share and imports share parametrically in (3.1) and nonparametrically in (2.2) and (2.9). Figure 3.3 serves as simple evidence for our hypotheses that the sign and magnitude of α_2 and α_3 may be different due to their own mechanisms of export and import on labor share across 94 countries during 1970-2009.

Vector X_{it} contains control variables. To follow related studies (see Table 2.1 and Section 2), we include the effect of capital schedule using a capital-output ratio in natural log form ($\ln k/y$), log real per capita GDP ($\ln PGDP$), population growth rate ($PopG$), FDI inflows ($FDI-ins$) and FDI outflows ($FDI-$

⁸We also checked the results for OECD and Non-OECD countries. Results remain quantitatively similar to what we reported in this paper, and are available upon request from the corresponding author.

Table 2.2: Descriptive data summary

Data	Description	Source	Mean	Median	Max	Min	S.D.
<i>LS</i>	Share of labor compensation in GDP at current estimation is based on “best estimate” that incorporates three different methods to re-correct the unobserved self-employment compensation. Missing values are interpolated.	PWT 8.1	0.533	0.545	0.852	0.175	0.133
<i>Exs</i>	Share of merchandise exports at current PPPs.	PWT 8.1	0.245	0.171	1.935	0.001	0.244
<i>Im</i> s	Share of merchandise imports at current PPPs.	PWT 8.1	0.284	0.220	2.192	0.001	0.252
<i>Openness</i>	The sum of <i>Exs</i> and <i>Im</i> s.	PWT 8.1	0.529	0.410	0.002	4.066	0.469
<i>EconG : barrier</i>	Index on import barriers, mean tariff rate, taxes on international trade and capital account restrictions.	KOF index	0.502	0.490	0.957	0.046	0.243
<i>SocialG</i>	Total social globalization index on the following subindices.	KOF index	0.435	0.405	0.926	0.069	0.219
<i>SocialG : culture</i>	Index with telephone traffic, transfer, international tourism, foreign population, and international letters.	KOF index	0.296	0.103	0.966	0.010	0.302
<i>SocialG : information</i>	Index on information flows: Internet users, television, and trade in newspapers.	KOF index	0.505	0.500	0.970	0.044	0.220
<i>SocialG : contact</i>	Index on cultural proximity: # of McDonalds and Ikeas, and trade in books (% in GDP).	KOF index	0.501	0.495	0.919	0.091	0.223
<i>PoliticalG</i>	Index on political globalization: embassies in country, membership in international organizations and international treaties, and participation in U.N. Security Council Missions.	KOF index	0.605	0.598	0.976	0.119	0.218
<i>FDI-ins</i>	FDI inflows as share of GDP.	UNCTAD	1.190	0.415	31.23	-10.01	2.360
<i>FDI-outs</i>	FDI outflows as share of GDP.	UNCTAD	0.726	0.040	44.99	-1.219	2.451
<i>ln invp</i>	Natural log price ratio of capital over household consumption.	PWT 8.1	0.215	0.129	1.907	-1.097	0.422
<i>ln k/y</i>	Natural log ratio of capital stock over GDP.	PWT 8.1	0.746	0.838	1.782	-1.468	0.509
<i>ln oilr</i>	Natural log real oil price in dollars.	St. Louis Fed	3.742	3.592	4.402	3.183	0.434
<i>PopG</i>	Population growth rate.	PWT 8.1	1.814	1.780	17.47	-4.506	1.503
<i>ln PGDP</i>	Natural log real per capita GDP.	PWT 8.1	8.708	8.812	11.65	5.593	1.176
<i>LaborReg</i>	Labor regulation index that includes the average of six survey-level rating for hiring market, hiring and firing, hours regulations, mandated cost, worker dismissal, and conscription	Fraser Institute	5.425	5.384	11.25	1.030	1.557
<i>HC</i>	Human capital index based on years of schooling.	PWT 8.1	2.232	2.249	3.592	1.066	0.585
Countries #	94; 90 with labor regulation; 84 with human capital						
Total obs	752; 720 with labor regulation; 672 with human capital						

outs) as share of GDP, log real oil price ($\ln oilr$) suggested by Brada (2013), and the log relative price ratio of investment to domestic consumption ($\ln invp$) that incorporates the argument by Karabarbounis and Neiman (2013). We also control the effect of trade transaction costs by economic restriction index ($EconG : Barrier$) from KOF index, an averaged index of imported barriers, mean tariff rate, taxes on international trade, and capital account restriction. It ranges from 0 to 100, with higher points assigned to countries with less economic trade restrictions.

For social and political factors, we follow the study of Young and Tackett (2017) by incorporating the social ($SocialG$) and political globalization ($PolitalG$) index in KOF dataset. Finally, we proxy labor market regulation using the index of labor market regulation ($laborReg$) from Gwartney et al. (2015) and control for human capital (HC). Note that the inclusion of these two variables reduces the sample size slightly due to missing values in the data. See Appendix 4 for the list of countries in the sample, as well as countries excluded due to missing values. See also Table 2.2 for a full data description and Table 2.3 for the correlation coefficients of regressors.

To maximize the total sample size while maintaining a balanced panel data with minimum of countries that have a constant labor share value, we examine at most 94 countries during 1970-2009. We follow the literature to minimize the impact of cyclical variation and maintain stationary properties by averaging all data within a 5-year interval: 1970-1974, 1975-1979, 1980-1984, 1985-1989, 1990-1994, 1995-1999, 2000-2004, 2004-2009. We note that social globalization contains three subindices that measure the globalization of personal contacts, information, and culture. Given their high correlation and our primary interests, only the total social and political globalization indices are used in this study.

2.4.2 Semiparametric model specification

Partially linear model (PLM)

Parametric estimation models require practitioners' *a priori* assumption for the underlying data generating process, thereby possibly concealing the potential nonlinearity of trade variables in the conditional mean function of labor share. Nonparametric regression is instead robust to model mis-specification as it permits a greater flexibility of the unknown conditional mean function to be estimated. However, a fully nonparametric approach is criticized due to its large sample size requirement for a consistent estimator. This is a well-known problem termed the *curse of dimensionality*.

To alleviate this problem, we first implement a partially linear model (PLM) that permits a finite number of variables to enter the regression nonlinearly:

$$\ln LS_{it} = X_{it}\beta + G(Z_{it}) + \epsilon_{it} \quad (2.2)$$

Table 2.3: Correlation coefficient of independent variables with labor regulation index and Human capital

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1)-Econ globalization:	1.00																	
Barriers	—																	
(2)-Econ globalization:		1.00																
Flow	(0.00)	—																
(3)-Social globalization			1.00															
Total	(0.00)	(0.00)	—															
(4)-Social globalization:				1.00														
Culture	(0.00)	(0.00)	(0.00)	—														
(5)-Social globalization:					1.00													
Information	(0.00)	(0.00)	(0.00)	(0.00)	—													
(6)-Social globalization:						1.00												
Personal contact	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	—												
(7)-Political globalization:							1.00											
Total	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	—											
(8)-Export shares								1.00										
(9)-Import shares									1.00									
(10)-FDI shares:										1.00								
Inflow	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	—								
(11)-FDI shares:											1.00							
Outflow	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	—							
(12)-Relative price												1.00						
(13)-Log(K/Y)													1.00					
(14)-Log(real oil price)														1.00				
(15)-Population growth															1.00			
(16)-Log(RGDP)																1.00		
(17)-Labor																	1.00	
regulation index																		1.00
(18)-Human Capital																		—
Total countries	84																	—
Total Obs	672																	—

Note: the significance of correlation is reported by Spearman p-value in parenthesis beneath the correlation coefficient.

where $i = 1, \dots, N$ and $t = 1, \dots, T$ represents our dataset with total N countries and T time period. ϵ_{it} is a mean-zero additive error and we allow for the presence of heteroskedasticity error process $E(\epsilon_{it}^2 | X_i, Z_i) = \sigma^2(X_i, Z_i)$ for fixed t . Regression (2.2) is similar to (3.1) except that we let Z contains d -dimensional variables that enter the model via an unknown function G with unrestricted functional form assumed. Therefore, we estimate the conditional mean function $G(\cdot)$ and its gradients with respect to Z nonparametrically. Given our limited observations, we restrict $d = 1$ to improve the efficiency of the PLM estimator in a mean squared error sense. In other words, we estimate the effect of one trade variable (either Exs or $Im s$) nonparametrically and the other one parametrically with quadratic form. However, since the standard PLM cannot estimate a constant due to the identification problem of $G(\cdot)$, we consider a fixed effect PLM in the next section.

The parametric component is estimated in two steps. We first follow Robinson (1988) to eliminate $G(\cdot)$ by taking expectation on both sides in (2.2) conditional on Z_{it} and subtract it by (2.2) again to obtain:

$$f(Z_{it}) [\ln LS_{it} - E(\ln LS_{it} | Z_{it})] = f(Z_{it}) [(X_{it} - E(X_{it} | Z_{it}))\beta + \epsilon_{it}] \quad (2.3)$$

where $f(Z_{it})$ is the density of Z_{it} that eliminates the random denominator from the Nadaraya-Watson (LCLS) estimator:

$$\hat{E}(\ln LS | Z) = \frac{\sum_{i=1}^N \sum_{t=1}^T k_h(Z_{it}, Z) \ln(LS)_{it}}{\sum_{i=1}^N \sum_{t=1}^T k_h(Z_{it}, Z)} \quad (2.4)$$

for $E(\ln LS_{it} | Z_{it})$, and

$$\hat{E}(X^s | Z) = \frac{\sum_{i=1}^N \sum_{t=1}^T k_h(Z_{it}, Z) X_{it}^s}{\sum_{i=1}^N \sum_{t=1}^T k_h(Z_{it}, Z)}, s = 1, \dots, p \quad (2.5)$$

for all dimension in $E(X_{it} | Z_{it})$, and $k_h(Z_{it}, Z)$ is a univariate kernel smoothing function $k_h(Z_{it}, Z) = (1/h)k(\frac{Z_{it}-Z}{h})$. Substitute (2.4) and (2.5) into (2.3) with the kernel density estimator evaluated at Z : $\hat{f}(Z) = (1/NT h) \sum_{i=1}^N \sum_{t=1}^T k_h(Z_{it}, Z)$, we obtain the $\sqrt{N}$ -consistent parametric estimator $\hat{\beta}_{PLM}$ that is asymptotically normally distributed:

$$\hat{\beta}_{PLM} = \left\{ \sum_{i=1}^N \sum_{t=1}^T S(X_{it}, Z_{it})^\top S(X_{it}, Z_{it}) \right\}^{-1} \left\{ \sum_{i=1}^N \sum_{t=1}^T S(X_{it}, Z_{it})^\top S(\ln LS_{it}, Z_{it}) \right\} \quad (2.6)$$

where $S(X_{it}, Z_{it}) = \hat{f}(Z_{it}) [X_{it} - \hat{E}(X_{it} | Z_{it})]$ and $S(\ln LS_{it}, Z_{it}) = \hat{f}(Z_{it}) [\ln LS_{it} - \hat{E}(\ln LS_{it} | Z_{it})]$.

The second step involves estimating the unknown function $G(\cdot)$ nonparametrically. Observe that $\tilde{Y}_{it} = G(Z_{it}) + \epsilon_{it}$, where $\tilde{Y}_{it} = (\ln LS_{it} - X_{it}\hat{\beta}_{PLM})$, we obtain both function estimator $\hat{G}(Z) \equiv \hat{\alpha}_0$ and

the first-order gradients estimates $\hat{G}'(Z) \equiv \hat{\alpha}_1$ evaluated at Z by a local linear estimator from:

$$(\hat{\alpha}_0, \hat{\alpha}_1) = \underset{\{\alpha_0, \alpha_1\}}{\operatorname{argmin}} \sum_{i=1}^N \sum_{t=1}^T \left[\tilde{Y}_{it} - \alpha_0 - (Z_{it} - Z)\alpha_1 \right]^2 k_h(Z_{it}, Z) \quad (2.7)$$

The intuition of the local linear estimator is that for each evaluation point we “locally” fit a linear line through Z based on its neighbor observation near to Z . The kernel function $k_h(Z_{it}, Z)$, such as a standard normal p.d.f., produces higher (lower) weights for points closer to (further away from) Z . Bandwidth h is used to control the range for which points around Z are considered as its neighbor observation. Therefore, a larger bandwidth implies more equal weight assigned to the neighbor points of Z , and vice versa. We then connect each of these linear lines to obtain the estimated unknown functions as well as its first-order gradients. In fact, it can be shown that the local linear estimator in (2.7) collapses to OLS estimator if the function $G(Z)$ is truly linear in Z (i.e., the local linear regression fits a linear line using all sample points as neighbor points of Z , implying that h diverges to positive infinity).

In this study, we choose $k_h(v) = \frac{1}{\sqrt{2\pi}} e^{-0.5v}$ as the Gaussian kernel, with its bandwidth selected based on cross-validation least square (CVLS) that minimizes the mean squared error of conditional mean function:

$$h_{cvls} = \underset{\{h>0\}}{\operatorname{argmin}} \sum_{i=1}^N \sum_{t=1}^T \left[\tilde{Y}_{it} - \hat{G}_{-i}(Z_{it}) \right]^2 \quad (2.8)$$

where $\hat{G}_{-i}(Z_{it})$ is similar to $\hat{G}(Z_{it})$ from (2.7), except that the i^{th} cross-section unit is removed when we evaluate at Z_{it} to ensure the validity of this optimization algorithm.⁹

To statistically support our empirical hypotheses, we formally test the significance of Z in $G(\cdot)$ and, if Z is potentially nonlinear, test the significance of its nonlinearity. To test the significance of Z in $G(\cdot)$, we implement the nonparametric test by Racine (1997) that examines whether the partial derivatives of $G(\cdot)$ are significantly different from zero over its entire support. Denoting λ as the test statistic and using aggregate L_2 norm measure, we test the following null $H_0 : \lambda = E \left[\frac{\partial E(\tilde{Y}|Z)^2}{\partial Z} \right] = 0$. See Appendix 1 for the estimation of λ along with its bootstrapping procedure.

To test the significant nonlinearity of Z in $G(\cdot)$, we employ a model specification test by Zheng (1996) to test whether the functional form of $G(\cdot)$ can be parametrically specified (i.e., $E(\tilde{Y}|Z) = \beta_0 + Z\beta_1$). Equivalently, we test the null hypothesis $H_0 : E(U|X) = 0$, where U is the regression error under the null. In this study, we consider testing the null functional form $G(\cdot)$ being linear (i.e., $H_{01} : G(Z) = \beta_0 + Z\beta_1$), quadratic (i.e., $H_{02} : G(Z) = \beta_0 + Z\beta_1 + Z^2\beta_2$), and cubic (i.e., $H_{03} : G(Z) = \beta_0 + Z\beta_1 + Z^2\beta_2 + Z^3\beta_3$). Therefore, the rejection of H_{01} indicates that Z enters the function $G(\cdot)$ nonlinearly, and the rejection of

⁹While the CVLS bandwidth selection given cross-section data is to leave-one-*observation* -out, we choose to leave one *cross-section-unit* out to follow the argument by Henderson and Parmeter (2015). The essential reason is to ensure that a country with its all observations as outliers can be entirely removed so as to prevent bandwidth from not being too small. As expected, this approach speeds up the optimization speed dramatically.

H_{02} and H_{03} indicates that the nonlinearity of Z is more complicated than a quadratic or cubic function. See Appendix 2 for the estimation of the test statistic, along with its bootstrapping procedure.

Fixed effect partially linear model (FEPLM)

We now implement the approach by Qian and Wang (2012) that allows for a PLM to be estimated with cross-sectional fixed effect. Similar to (2.2), we consider the following regression:

$$\ln LS_{it} = X_{it}\beta + G(Z_{it}) + \alpha_i + \epsilon_{it} \quad (2.9)$$

and the cross-country fixed effect α_i can be eliminated through first-order difference:

$$\Delta \ln LS_{it} = \Delta X_{it}\beta + G(Z_{it}) - G(Z_{i,t-1}) + \Delta \epsilon_{it} \quad (2.10)$$

We follow Qian and Wang (2012) by further ensuring that X_i , Z_i , and ϵ_i are *i.i.d.* for t fixed, and X_t is strictly stationary for i fixed. The covariance matrix of ϵ_{it} is allowed to be non-diagonal, so that both serial correlation and random walk are allowed in this approach.

As in PLM, the linear component in (2.10) is estimated from the following density weighted regression:

$$Y_{it}^* = X_{it}^*\beta + u_{it}^* \quad (2.11)$$

where $Y_{it}^* = \rho_{it} [\Delta \ln LS_{it} - E(\Delta \ln LS_{it} | Z_{it}, Z_{it-1})]$, $X_{it}^* = \rho_{it} [\Delta X_{it} - E(\Delta X_{it} | Z_{it}, Z_{it-1})]$, and $u_{it}^* = \rho_{it} \Delta \epsilon_{it}$, with $\rho_{it} = f(Z_{it}, Z_{it-1})$ the joint density of $[Z_{it}, Z_{it-1}]$. We obtain the estimator of β in (2.11), $\hat{\beta}_{FEPLM}$, in a similar fashion as in (2.6).

To estimate the nonlinear component $G(Z_{it}) - G(Z_{it-1})$ in (2.10), Qian and Wang (2012) treat it as an additive model $\tilde{Y}_{it} = G_1(Z_{it}) - G_2(Z_{it-1}) + \Delta \epsilon_{it}$, where $\tilde{Y}_{it} = (Y_{it} - X_{it}\hat{\beta}_{FEPLM})$. Define $\Omega(Z_1, Z_2) = G_1(Z_{it}) - G_2(Z_{it-1}) \equiv G_1(Z_1) - G_2(Z_2)$, our interest $G_1(Z_1)$ can be identified via marginal integration by Linton and Nielsen (1995):

$$G_1(Z_1) = \int \Omega(Z_1, Z_2) q(Z_2) dZ_2 \quad (2.12)$$

where $q(\cdot)$ is a pre-determined positive weighting function satisfying $\int q(Z_k) dZ_k = 1$ and $\int G_k(Z_k) q(Z_k) dZ_k = 0$ with $k = 1, 2$. Function $G_1(Z_1)$ is estimated by replacing (2.12) with its sample analogue:

$$\hat{G}_1(Z_1) = \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T \hat{\Omega}(Z_1, Z_{it-1}) \quad (2.13)$$

where $\hat{\Omega}(\cdot, \cdot) \equiv \hat{b}_0$ from

$$(\hat{b}_0, \hat{b}_1, \hat{b}_2) = \underset{\{b_0, b_1, b_2\}}{\operatorname{argmin}} \sum_{i=1}^N \sum_{t=2}^T \left[\tilde{Y}_{it} - b_0 - (Z_{it} - Z_1)b_1 - (Z_{it-1} - Z_2)b_2 \right]^2 K_H(Z_{it}, Z) \quad (2.14)$$

where $K_H(Z_{it}, Z) = \frac{1}{h_1 h_2} k_1 \left(\frac{Z_{it} - Z_1}{h_1} \right) k_2 \left(\frac{Z_{it-1} - Z_2}{h_2} \right)$ is a product kernel function. We choose the 2nd-order Gaussian kernel function and select bandwidths (h_1, h_2) based on the CVLS method as in (2.8). Similarly, we follow Severance-Lossin and Sperlich (1999) to estimate the first-order derivate $G'_1(Z_1) \equiv \frac{\partial G_1(Z_1)}{\partial Z_1}$ from

$$\hat{G}'_1(Z_1) = \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T \hat{\Omega}'(Z_1, Z_{it-1}) \quad (2.15)$$

where $\hat{\Omega}'(Z_1, Z_{it-1})$ is the estimator $\hat{b}_1$ from (2.14).¹⁰

2.5 Empirical Results

2.5.1 Parametric model results

Table 2.4 presents the whole-sample results from (3.1) with cross-country fixed effect.¹¹ We also estimate the effect of trade openness (i.e., $Exs + Imss$) to be consistent with the current literature. The linear (quadratic) functional form of trade openness, export share, and import share is examined in Column [1]-[6] (Column [7]-[12]). The most appealing result is that trade openness enters the regression significantly only in the *quadratic* function, with the average negative partial impact to labor share being 0.043, 20 times larger than in linear specification. Hence, a conventional quadratic parametric model gives statistical evidence that trade openness affects labor share beyond a linear fashion.

Results in Column [8], [10], and [12] show that export and import share enter the regression significantly and differently from each other. The effect of exports seems to be more “detrimental” than imports, since a one-percentage point increase in export (import) share on average decreases (increases) labor share by 0.216% (0.112%). However, the estimated quadratic coefficients indicate that both partial effects diminish as export or import share increases. The above results are consistent with hypothesis 1 and 2, since exports and imports can potentially influence labor share in a nonlinear fashion due to their own underlying relation to the economy, such as change in market competition (i.e., due to import expansion) and in trade specification transformation (i.e., due to export expansion).

¹⁰Note that the exogeneity condition $E(\epsilon|X, Z) = 0$ are implicitly assumed for the use of all regression model we considered. Hence, we test the null hypothesis where the our interested variables Exs or $Imss$ is exogenous via a control function approach (See Amsler et al. (2016) for its discussion). Using one- and two-year-lag variables as our instruments, we are fail to reject the null under each case. Hence, we are more confident regarding the issue of identification. Results are available upon request.

¹¹Standard errors are clustered by countries. Constant terms are omitted to save space.

Table 2.4: Parametric Regression Results of Trade Variable(s) Impact on Ln(Labor Share)

	Whole Sample											
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Openness	-0.003 (0.007)		-0.002 (0.005)		-0.002 (0.004)		-0.046*** (0.012)		-0.040*** (0.010)		-0.042* (0.008)	
Openness ²							0.015*** (0.002)		0.012*** (0.002)		0.014** (0.003)	
<i>Exs</i>		-0.048*** (0.017)		-0.044** (0.020)		-0.066*** (0.023)		-0.220*** (0.058)		-0.192*** (0.059)		-0.235*** (0.054)
<i>Exs</i> ²								0.130*** (0.039)		0.108*** (0.038)		0.145*** (0.038)
<i>Im</i> s		0.046** (0.016)		0.041*** (0.016)		0.062*** (0.017)		0.104* (0.051)		0.092* (0.053)		0.139*** (0.054)
<i>Im</i> s ²								-0.068** (0.043)		-0.053* (0.041)		-0.069* (0.041)
FDI-ins	-0.001*** (0.000)	-0.002*** (0.000)	-0.003*** (0.000)	-0.002** (0.000)	-0.001 (0.001)	-0.001 (0.001)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.001)	-0.002*** (0.001)	-0.003 (0.001)	-0.001 (0.001)
FDI-outs	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.001 (0.001)	0.003 (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.004 (0.001)	0.003* (0.001)
SocialG	0.020 (0.018)	0.014 (0.016)	0.012 (0.016)	0.006 (0.014)	0.015 (0.014)	0.006 (0.015)	0.024 (0.016)	0.013 (0.018)	0.012 (0.015)	0.003 (0.018)	0.020* (0.011)	0.006 (0.015)
PoliticalG	-0.110*** (0.008)	-0.111*** (0.008)	-0.101*** (0.010)	-0.102*** (0.013)	-0.041*** (0.008)	-0.043*** (0.007)	-0.103*** (0.008)	-0.108*** (0.010)	-0.094*** (0.010)	-0.099*** (0.014)	-0.037*** (0.007)	-0.044*** (0.008)
EconG: Barrier	-0.092*** (0.023)	-0.089*** (0.020)	-0.107*** (0.022)	-0.103*** (0.019)	-0.089*** (0.024)	-0.091*** (0.020)	-0.070** (0.025)	-0.066** (0.021)	-0.085*** (0.024)	-0.076*** (0.022)	-0.063*** (0.023)	-0.059*** (0.021)
lnky	0.031*** (0.009)	0.027*** (0.008)	0.027*** (0.009)	0.021*** (0.008)	0.047*** (0.012)	0.043*** (0.011)	0.031*** (0.009)	0.027*** (0.009)	0.027*** (0.010)	0.022** (0.009)	0.045*** (0.012)	0.043*** (0.011)
lninvp	0.042*** (0.008)	0.037*** (0.007)	0.045*** (0.009)	0.040*** (0.009)	0.052*** (0.011)	0.049*** (0.011)	0.042*** (0.009)	0.038*** (0.008)	0.044*** (0.010)	0.040*** (0.009)	0.050*** (0.012)	0.047*** (0.011)
Popg	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004* (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004** (0.002)	0.004*** (0.002)
lnRGDP	-0.030*** (0.007)	-0.027*** (0.007)	-0.024*** (0.005)	-0.021*** (0.005)	-0.0023* (0.005)	-0.002 (0.005)	-0.031*** (0.008)	-0.026*** (0.006)	-0.026*** (0.006)	-0.021** (0.004)	-0.009** (0.005)	-0.004 (0.005)
lnoilr	-0.007 (0.005)	-0.009 (0.005)	-0.008** (0.005)	-0.009* (0.005)	-0.010* (0.006)	-0.012** (0.006)	-0.006 (0.005)	-0.007*** (0.005)	-0.006 (0.005)	-0.008 (0.005)	-0.008 (0.006)	-0.009 (0.006)
laborReg			-0.004* (0.002)	-0.004* (0.002)	-0.004** (0.002)	-0.004** (0.002)	-0.004** (0.002)	-0.004** (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004** (0.002)	-0.004** (0.002)
HC					-0.071*** (0.017)	-0.070*** (0.015)					-0.071*** (0.017)	-0.070*** (0.014)
Cross-section Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj R^2	0.948	0.948	0.949	0.949	0.950	0.950	0.948	0.948	0.949	0.949	0.950	0.950
F-Stats	132.15***	131.31***	132.43***	131.55***	132.51***	132.17***	131.57***	129.80***	131.78***	129.74***	131.95***	131.04***
Total country Obs	94	94	90	90	84	84	94	94	90	90	84	84
Total Observation	752	752	720	720	672	672	752	752	720	720	672	672

Note: Dependent variable is the natural log of labor share. Cross-section clustered errors are reported beneath the estimated coefficients. ***, **, * stands for the statistically significant level at 1%, 5% and 10%, respectively.

In addition, the share of FDI outflows (inflows) are in line with the previous studies' results, which are in general significantly and positively (negatively) correlated with labor share. The fact that FDI outflow contributes to a higher labor share is expected, since additional exported capital would facilitate the use of new technology, therefore boosting employment, productivity, and labor real wages. However, the effect of FDI inflows is ambiguous in the literature. For example, an increase in FDI inflows may be associated with more job opportunities and thus higher wages in emerging sectors (UNCTAD (1994)). However, higher FDI inflows may also decrease wage level as competition with domestic sectors becomes severe, resulting in exiting firms and job losses. It seems that the latter effect dominates the former in our study and is also consistent with previous findings (Harrison (2005); Guerriero and Sen (2012)).

The role of market policies also significantly affects labor share. Economic barriers, for example, decrease labor share on average by 0.058%. A similar significant effect is found in labor market regulation, which may infringe on economic freedom by “dismissal regulations, centralized wage setting, extension of union contracts to non participating parties, and conscription” (Gwartney et al. (2015), p.16). Finally, we note that the effect of social and political globalization is sensitive to model specification as shown in Young and Tackett (2017). However, we show in the following that social globalization significantly increases labor share once we consider a more flexible functional form, particularly when we consider OECD countries.

Overall, we argue that the above findings are consistent with the most current literature. However, the results from the parametric model should be interpreted with caution, as the inclusion of the quadratic function of openness does not necessarily imply a correctly specified model, given that a polynomial estimation faces a trade-off between local and global structure (Magee (1998)). Given the potential risk of model mis-specification, we prefer to let the trade reveal the effect itself in a more flexible functional form.

2.5.2 Semiparametric model results

FEPLM results: Whole sample

Table 2.5 reports the results of FEPLM estimation in (2.9) by first assuming trade openness in the unknown function. Since the majority of observations in our full sample have trade openness less than 100%, we report the estimated gradients conditional on the 25%, 50%, 75%, and 100% of trade openness level. Similar to the results in parametric regression, the FEPLM clearly exhibits a diminishing return of trade openness on labor share, which is likely to be a result of the joint effects of export and import. However, the effect of trade openness in FEPLM is about a half of the effect in the parametric model. Comparing the results in column [3] of Table 2.5 with column [7] of Table 2.4, for instance, as the

Table 2.5: FEPLM Regression Results of Trade Variables Impact on Ln(Labor Share)

	Whole Sample					
	[1]		[2]		[3]	
	Nonlinear	Linear	Nonlinear	Linear	Nonlinear	Linear
	Openness		Openness		Openness	
25%	-0.0217*** (0.001)		-0.0184*** (0.0001)		-0.0275*** (-0.0095)	
50%	-0.0206*** (0.002)		-0.0168*** (0.001)		-0.0249*** (0.0021)	
75%	-0.0194*** (0.003)		-0.0151*** (0.002)		-0.0222*** 0.0023	
100%	-0.0181* (0.0102)		-0.0133* (0.0075)		-0.0193* (0.0107)	
FDI-ins		-0.016*** (0.006)		-0.021*** (0.007)		-0.026*** (0.008)
FDI-outs		0.024*** (0.005)		0.029*** (0.006)		0.029*** (0.007)
SocialG		0.057 (0.102)		0.062 (0.103)		0.064 (0.118)
PoliticalG		-0.166*** (0.073)		-0.157** (0.075)		-0.125* (0.076)
EconG: Barrier		-0.125* (0.075)		-0.087 (0.076)		-0.010 (0.087)
lnk/y		0.005 (0.031)		0.026 (0.032)		0.047 (0.034)
lninvp		0.129*** (0.025)		0.134*** (0.035)		0.158*** (0.037)
Popg		0.061*** (0.007)		-0.061*** (0.008)		-0.056*** (0.008)
lnRGDP		-0.019*** (0.007)		-0.016** (0.008)		-0.010 (0.008)
lnoilr		-0.007 (0.022)		-0.012 (0.022)		-0.011 (0.023)
laborReg				-0.021*** (0.006)		-0.023*** (0.006)
HC						0.034 (0.029)
Cross-country Fixed Effect	Yes		Yes		Yes	
Bandwidth	0.689		0.706		0.856	
Bandwidth Selection	CVLS	CVLS	CVLS	CVLS	CVLS	CVLS
Kernel Estimator	LLS	LCLS	LLS	LCLS	LLS	LCLS
R ²	0.641	0.312	0.677	0.334	0.660	0.349
Total country Obs		94		90		84
Total Observation		658		630		588

Note: Dependent variable is the natural log of labor share. Wild bootstrapped standard errors are reported in parenthesis for nonlinear component. ***, **, * stands for the statistically significant level at 1%, 5% and 10%, respectively.

Table 2.6: FEPLM Regression Results of Trade Variable(s) Impact on Ln(Labor Share)

Variable	Whole Sample					
	[1]		[2]		[3]	
	Nonlinear	Linear	Nonlinear	Linear	Nonlinear	Linear
	Exs		Exs		Exs	
5%	-0.1673*** (0.014)		-0.1271*** (0.010)		-0.1368*** (0.045)	
25%	-0.1541*** (0.014)		-0.1157*** (0.013)		-0.1218*** (0.042)	
50%	-0.0713*** (0.014)		-0.0737*** (0.016)		-0.0906*** (0.037)	
75%	-0.0583*** (0.015)		-0.0212 (0.021)		-0.0534* (0.032)	
Exs						
Exs ²						
Ims						
Ims ²						
FDI-ins						
FDI-outs						
SocialG						
PoliticalG						
EconG: Barrier						
lnk/y						
lninvp						
Popg						
lnRGDP						
lnoilr						
laborReg						
HC						
Cross-country Fixed Effect	Yes		Yes		Yes	
Bandwidth	0.070		0.147		0.081	
Bandwidth Selection	CVLS		CVLS		CVLS	
Kernel Estimator	LLS		LLS		LLS	
R ²	0.529		0.677		0.606	
Total country Obs	94		90		84	
Total Observation	658		630		588	

Figure 2.4: Estimated FEPLM for Trade Openness, Export Share, and Import Share

Note: Graph in row 1, 2 and 3 stands for the FEPLM function and gradient estimates for trade openness, export share, and import share, respectively, in column [3] of Table 2.5 and 2.6. The estimated OLS partial effects are from their corresponding values in column [5] and [6] of Table 2.4. All results are obtained by controlling for Labor Regulation and Human Capital.

trade openness increases from 25% to 100%, its corresponding partial effect declines from -0.028 to -0.019 relative to -0.043 from the parametric estimation. To access the significance of the nonlinearity of trade openness, we report the p-values of the tests for variable significance and three different model specifications of $G(\cdot)$ in Panel A of Table 2.8, and we only report the results that include labor regulation index and human capital. The rejection of the four null hypothesis H_0, H_{01}, H_{02} , and H_{03} at the 1% level under “Whole Sample” indicates that the nonlinearity of trade openness in the function of labor share is significant and cannot be simply described by quadratic or cubic function specification.¹²

We proceed to estimate the effect of export share (or import share) nonparametrically in $G(\cdot)$ and import share (or export share) quadratically in parametric function to alleviate the curse of dimensionality. The results reported in Table 2.6 clearly reveal a diminishing effect of both variables: in column [3], for example, a one percentage change in export (import) share is associated with a 0.122% (0.115%) decrease (increase) in labor share at the 25% level, the partial effect gradually decreases down to 0.053 (0.056) as export share (import share) rises to 75%. The significance of the nonlinearity of both variables is confirmed by the p-values in Panel A of Table 2.8.

To have a clear picture of the estimated function $\hat{G}(Z)$ in (2.13) and partial effect $\hat{G}'(Z)$ in (2.15) evaluated at points $\{Z_{it}\}_{i,t=1}^{N,T}$, we present graphical results along with their wild-bootstrapped 95% confidence intervals in Figure 2.4. We only report the results with all control variables to save space. Graphs in rows 1, 2, and 3, respectively, report the estimated function and gradients for trade openness, export share, and import share with their corresponding numerical results in column [3] of Table 2.5 and 2.6. In the second column graphs, we plot the estimated gradients of openness, export, and import in FEPLM (the solid line) against the parametric ones (the dotted-dash line) from column [5]-[6] in Table 2.4. Clearly, most of the OLS estimates are outside the 95% confidence interval constructed for the FEPLM estimates, indicating a potential misspecification of a linearly restricted regression model in (3.1). Overall, the FEPLM results support our hypotheses by showing different nonlinear effects of imports and exports, which go beyond a quadratic or cubic function according to the four-digit zero p-value from the test for H_{02} - H_{03} in Table 2.8.

PLM results: Group of countries

Estimating the impact of trade variables semiparametrically for countries in different economic development stages, namely OECD and Non-OECD countries, may reveal their heterogeneous impact on labor share. However, we face the challenge of obtaining a consistent estimator of FEPLM given the limited

¹²We also checked the sensitiveness of the p-values to bandwidths using different bandwidth selection, such as least-square cross-validation and AIC-based cross-validation. The results remain fairly the same and are available upon request from the corresponding author.

Table 2.7: Partial Linear Model Results of Trade Variables' Impact on Ln(Labor Share)

	OECD			Non-OECD		
	Nonlinear	linear	Nonlinear	linear	Nonlinear	linear
5%	EXS -0.349*** (0.027)		IMS 0.329* (0.198)	EXS -0.811*** (0.091)	IMS 0.867*** (0.300)	
25%	-0.238*** (0.108)		0.276* (0.167)	-0.809*** (0.077)	0.777*** (0.149)	
50%	-0.152* (0.092)		-0.052 (0.204)	-0.748*** (0.061)	0.633** (0.258)	
75%	-0.107 (0.157)		-0.083 (0.356)	0.532*** (0.061)	0.385* (0.232)	
Exs				-0.02 (0.183)		-0.995*** (0.146)
Exs^2				0.037 (0.126)		0.237* (0.143)
Imns		0.389** (0.167)			0.779*** (0.215)	
$Imns^2$		-0.104 (0.206)			-0.234** (0.141)	
FDI-ins		-0.009 (0.009)		-0.008 (0.007)	-0.008 (0.101)	-0.0003 (0.010)
FDI-outs		0.011* (0.006)		0.009* (0.005)	0.001 (0.019)	0.004 (0.020)
SocialG		0.298*** (0.108)		0.264** (0.104)	-0.161 (0.149)	-0.252* (0.146)
PoliticalG		-0.137 (0.093)		-0.108 (0.086)	-0.077 (0.096)	-0.094 (0.096)
EconG; Barrier		-0.356** (0.151)		-0.303** (0.131)	-0.168 (0.106)	-0.162 (0.102)
lnk/y		0.341*** (0.069)		0.367*** (0.065)	-0.032 (0.037)	-0.039* (0.031)
lninvp		0.512*** (0.086)		0.540*** (0.078)	0.078** (0.038)	0.068* (0.035)
Popg		-0.079*** (0.026)		-0.066*** (0.025)	-0.023** (0.009)	-0.026*** (0.009)
lnRGDP		0.006 (0.009)		0.003 (0.009)	0.016* (0.009)	0.019** (0.010)
lnoilr		-0.042** (0.021)		-0.043** (0.021)	-0.004 (0.028)	-0.004 (0.027)
laborReg		-0.004 (0.008)		0.005 (0.007)	-0.021*** (0.007)	-0.023*** (0.007)
HC		-0.064** (0.033)		-0.077** (0.032)	0.066* (0.035)	0.063* (0.035)
Function significant Test	0.088		0.008		0.000	0.000
Bandwidth	0.084		0.053		0.343	0.172
Bandwidth Selection	CVLS	CVLS	CVLS	CVLS	CVLS	CVLS
Kernel Estimator	LLLS	LLLS	LLLS	LLLS	LLLS	LLLS
R^2	0.466	0.53	0.437	0.57	0.567	0.504
Total country Obs	22	22	22	22	62	62
Total Observation	176	176	176	176	496	496

Figure 2.5: Gradients comparison across OECD and Non-OECD countries

Note: Reported gradients above each bar are the estimation of $G'(Z)$ in (2.2) conditional on the level of exports and imports in OECD and Trade-oriented countries. Bandwidth used is h_{cvls} in (2.8).

Table 2.8: Nonparametric Test Results for Variable Significance, Model Specification, and Density Equality

Panel. A		Whole Sample			OECD		Non-OECD	
		Openness	Exs		ImS	Exs	ImS	
Variable Significance Test		0.0000	0.0000	0.0000	0.0875	0.0075	0.0000	0.0000
$H_0 : G'(Z) = 0$		(0.0000)	(0.0000)	(0.0000)				
Model Specification Test 1		0.0175	0.0250	0.0775	0.0200	0.0450	0.0600	0.0850
$H_{01} : G(Z) = \beta_0 + Z\beta_1$								
Model Specification Test 2		0.0075	0.0150	0.0875	0.0150	0.0425	0.0575	0.0625
$H_{02} : G(Z) = \beta_0 + Z\beta_1 + Z^2\beta_2$								
Model Specification Test 3		0.0800	0.0325	0.0400	0.0900	0.0775	0.0525	0.0350
$H_{03} : G(Z) = \beta_0 + Z\beta_1 + Z^2\beta_2 + Z^3\beta_3$								
B (B_1)		399 (1000)						
Kernel function		2nd-order Gaussian						
Bandwidth		Adaptive Rule-of-Thumb						
Sample Size		672	672	672	176	176	496	496
Panel. B								
Density Equality Test ($H_{04} : q_1(\mathbf{g}^1) = q_2(\mathbf{g}^2)$)								
		Export share			Import share			
Compared groups		$\hat{J}$	$\hat{T}$	P-value	$\hat{J}$	$\hat{T}$	P-value	
Non-OECD & OECD		91.26	7.74	0.0000	73.83	32.89	0.0000	
B		399						
Kernel function		2nd-order Gaussian						
Bandwidth		Adaptive Rule-of-Thumb						
Sample size		OECD	176					
		Non-OECD	496					

Note: Panel. A presents the p-value of nonparametric tests for variable's significance given H_0 and model specification given (H_{01} - H_{03}) by Racine (1997) and Zheng (1996), respectively. For testing H_0 , we report the p-value in parenthesis for FEPLM. Panel. B shows the p-value of nonparametric density equality test by Li (1996). B (B_1) is the number of bootstrap replication for model specification test (significance of variable test), and bandwidth selection is adaptive rule-of-thumb in (2.8). All results are obtained after controlling for Labor Regulation Index and Human Capital.

sample size in OECD countries. For the purpose of comparison, we estimate the effect of export and import in OECD and Non-OECD countries via PLM in (2.2) and report the results in Table 2.7.

We see that while the nonlinear impact of export or import continues to be significant, the effect varies significantly in magnitude across groups. Labor share in OECD countries is far less affected at all trade levels compared to that in non-OECD countries. For instance, while the export share at 25% level decreases labor share on average by 0.24% in OECD countries, it decreases 0.81% in non-OECD countries with relatively more labor-intensive technology and non-sophisticated trade specification. This is consistent with hypothesis 2. Similar differences are also observed in the import effect, suggesting higher domestic competition in developing countries and thus the larger increase in labor share, as stated in hypothesis 1. Finally, we do not find statistical evidence that labor share in OECD countries is significantly affected by export or import share higher than a 50% level. Indeed, the function significance test reveals both export and import enter the nonparametric regression more significantly in non-OECD countries than in OECD countries.

To clearly observe the difference in estimated partial effects across the two groups of countries, we plot the estimated gradients against export and/or import share level in Figure 2.5. We employ a consistent univariate density equality test by Li (1996) to test the null hypothesis where the distributions of export or imports effects from both OECD and Non-OECD countries are equal. See the estimation of test statistics, along with its bootstrap procedure in Appendix 3. Panel B of Table 2.8 reports the p-value of test statistics $\hat{T}$ and $\hat{J}$ that are normally and standard normally distributed, respectively. The 4-digit zero p-value indicates that the estimated impact of export or import share is significantly different in OECD and non-OECD countries.

In fact, the reason for such a difference can be explained by the difference between the estimated capital-schedule effect by $\ln(k/y)$ in OECD and non-OECD countries in Table 2.7. It is possible that as export becomes a dominant factor in terms of a country's GDP (i.e., OECD countries), export sectors are likely to expand, creating more job opportunities that produce sophisticated goods by skilled labor (Boggio et al. (2010); Lawrence (2008)). This change in trade structure would result in a positive elasticity of substitution between capital and labor in OECD countries, supported by a complementary use of capital and labor (i.e., the partial effects of "capital schedule" in OECD countries are positive, ranging from 0.341-0.367). In contrast, non-OECD countries not experiencing such changes are likely to have substitutable labor given cheaper capital price (i.e., the partial effects of "capital schedule" in non-OECD countries are negative, ranging from (-0.032)-(-0.039)). Hence, we expect that labor in non-OECD countries would be more complementary with capital when the trade sector is promoted enough to enhance sophisticated goods production, increasing the demand for high-skilled labor and thus real wage. The faster decline in

the partial effects of export in non-OECD countries may suggest that demand for high skilled labor is surging as trade structure transforms towards a capital-intensive one, as is the case in China (Wang and Wei (2010)). This may explain why only non-OECD countries show a positive and significant effect of human capital.

2.6 Conclusion

The impact of trade is of concern in the literature of functional distribution in national income between capital and labor. While the cross-country decline of labor share was attributed to raising trade openness in earlier studies, the opposite effect is also found in later studies. Given that a theoretical framework that offers an unambiguous implication may not exist, the empirical evidence of the individual effect of exports and imports on labor share awaits further empirical analysis to look for the causes applicable for all countries.

We move into this direction by revealing the nonlinear effects of export share and import share with the support of the newest labor share estimation from PWT 8.1 and a partially linear model with or without a cross-country fixed effect. Our empirical results are fairly in line with our hypotheses, indicating that the significantly negative (positive) partial effect of export share (import share) diminishes as its level rises. While a quadratic parametric regression also detects the nonlinearity, the rejection in the model specification test indicates that a function of trade variable without model restriction should be estimated. The revealed nonlinearity implies that the effect of export or import is far from constant due to the joint effect of the Stolper-Samuelson theorem, technical shift, increasing competition, and the change in labor's bargaining power.

Three caveats are listed below for future research. The first lies in the difficulties of FEPLM, which cannot take into account the fixed effect given a limited sample size. In other words, PLM can be estimated with relatively small observations but in the absence of fixed effect, while FEPLM allows for the presence of fixed effect but requires larger observations to achieve consistent estimator. The trade-off between sample size and fixed effect in semiparametric panel regression is posited with challenges. In addition, the nonlinear effects of export and import, though being significant, still lack a formally theoretical foundation that justifies the underlying nonlinear causality channel. Finally, the use of country-aggregate data restricts our ability to identify the underlying causes of how export and import affect labor share differently in countries with different levels of economic development. Future research strives to pin down these questions, particularly by focusing on the heterogeneous and nonlinear effects of export and import across developing and developed countries with the support of micro-level data.

2.7 Appendixes

Appendix 1 Following Racine (1997), we evaluate the relevance of Z (i.e., *openness*, *Exs*, or *Ims*) in the conditional mean function $G(\cdot)$ by testing the following null and alternative hypothesis:

$$\begin{cases} H_0 : \lambda = E \left[\frac{\partial E(\tilde{Y}|Z)^2}{\partial Z} \right] = 0 \\ H_1 : \lambda = E \left[\frac{\partial E(\tilde{Y}|Z)^2}{\partial Z} \right] > 0 \end{cases}$$

To estimate λ , we replace the unknown gradients $G'(Z) \equiv \frac{\partial G(Z)}{\partial Z}$ with its nonparametric estimator $\hat{G}'(Z) \equiv \hat{\alpha}_1$ obtained in (2.7) to construct our feasible pivoted test statistic:

$$\hat{t} = \frac{\hat{\lambda}}{sd(\hat{\lambda})}$$

where

$$\hat{\lambda} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left[\frac{\hat{G}'(Z_{it})}{sd(\hat{G}'(Z_{it}))} \right]^2$$

and $sd(\cdot)$ is the sample standard deviation. We bootstrap the distribution of $\hat{t}$ and reject H_0 if $\hat{t} > \hat{t}_{0.95}^*$, where $\hat{t}_{0.95}^*$ is the 0.95 percentile of the bootstrapped test statistic $\hat{t}^*$ under the null. The following outlines its bootstrap procedure:

1. Given the original sample $\{\tilde{Y}_{it}, Z_{it}\}_{i=1, t=1}^{N, T}$, calculate $\hat{\lambda}$. Re-sample with replacement the original sample to obtain a new dataset $\{\tilde{Y}_{it}^*, Z_{it}^*\}_{i=1, t=1}^{N, T}$ to compute $\hat{\lambda}$. Repeat this step B_2 times to obtain re-sampled test statistics $\hat{\lambda}_1^*, \hat{\lambda}_2^*, \dots, \hat{\lambda}_{B_2}^*$, which is used to compute the standard error $sd(\hat{\lambda})$ test statistic $\hat{t}$.
2. Re-sampling with replacement the sample residual $\{\hat{u}_{it}\}_{i=1, t=1}^{N, T}$ to get resampled residual $\{\hat{u}_{it}^*\}_{i=1, t=1}^{N, T}$ and center it, where $\hat{u}_{it} = Y_{it} - \bar{Y}$ with $\bar{Y}$ the sample mean to impose the null where Z does not show up in $G(\cdot)$ so that $\beta(Z_{it}) = 0 \forall i, t$.
3. Obtain bootstrap sample $\{\tilde{Y}_{it}^*, Z_{it}\}_{i=1, t=1}^{N, T}$, where $\tilde{Y}_{it}^* = \bar{Y} + u_{it}^*$. Compute $\hat{\lambda}$, $sd(\lambda)$, and $\hat{t}^*$ as in step 1 using the bootstrap sample just obtained.
4. Repeat the step 2-3 B_1 times to obtain $\hat{t}_1^*, \hat{t}_2^*, \dots, \hat{t}_{B_1}^*$. We reject H_0 if $\hat{t} > \hat{t}_{1-\alpha}^*$, where $\hat{t}_{1-\alpha}^*$ is the upper $(1 - \alpha)$ percentile value of its empirical distribution, and α is the significance level.

We set $\alpha = 0.05$ and set $B_1 = 1000$, and $B_2 = 100$ in line with Racine (1997). We choose the 2nd-order Gaussian kernel function, and follow Dehnad (1987) to select bandwidth based on $h_{AROT} = 1.059A(NT)^{-1/5}$ where $A = \min\{\hat{\sigma}, IQR_z/1.34\}$ and IQR_z is the inter-quantile range of the variable Z .

Appendix 2 We employ a model specification test by Zheng (1996) to test the significant nonlinearity of Z in $G(\cdot)$, whose null hypothesis states whether the functional form of $G(\cdot)$ can be parametrically specified, such as linear or quadratic function. Equivalently, we test whether $E(U|Z) = 0$ under the null. By the law of iteration, we have the following infeasible test statistic under the null and the alternative:

$$\begin{cases} H_0 : I = E(UE(U|Z)f(Z)) = 0 \\ H_1 : I = E(UE(U|Z)f(Z)) > 0 \end{cases}$$

where $f(\cdot)$ is the density of its argument, which can be consistently estimated by its kernel estimator $\hat{f}(Z) = (1/NT) \sum_{i=1}^N \sum_{t=1}^T K_h(Z_{it}, Z)$. We estimate I by replacing U with the residual from the null regression (i.e., the OLS residuals), and replacing the conditional mean with the local constant estimator of U on Z :

$$\hat{E}(U|Z) = \frac{\sum_{i=1}^N \sum_{t=1}^T K_h(Z_{it}, Z) U_{it}}{\sum_{i=1}^N \sum_{t=1}^T K_h(Z_{it}, Z)}$$

By leaving out the diagonal terms to lower its finite-sample bias, we construct the feasible test statistic

$$\hat{I} = \frac{1}{N^2 h} \sum_{i=1}^N \sum_{t=1}^T \sum_{\substack{j=1 \\ \{j,k\} \neq \{i,t\}}}^N \sum_{k=1}^T \hat{u}_{it} \hat{u}_{jk} K_h(Z_{it}, Z_{jk})$$

with its variance:

$$\hat{\sigma}^2 = \frac{2}{N^2 h} \sum_{i=1}^N \sum_{t=1}^T \sum_{\substack{j=1 \\ \{j,k\} \neq \{i,t\}}}^N \sum_{k=1}^T \hat{u}_{it}^2 \hat{u}_{jk}^2 K_h^2(Z_{it}, Z_{jk})$$

so we finally construct the standardized version of test statistic:

$$\hat{J} = (NT\sqrt{h}) \frac{\hat{I}}{\hat{\sigma}} \quad (2.16)$$

which converges to the standard normal distribution under the null. The following illustrates its bootstrap procedure.

1. Given the original sample $\{\tilde{Y}_{it}, Z_{it}\}_{i=1, t=1}^{N, T}$, calculate $\hat{J}$.
2. Generate centered bootstrapped residual $\{\hat{u}_{it}^*\}_{i, t=1}^{N, T}$, where for each observation $i = 1, \dots, N$ and $t = 1, \dots, T$, $\hat{u}_{it}^* = \frac{1-\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$ and $\hat{u}_{it}^* = \frac{1+\sqrt{5}}{2}(\hat{u}_{it} - \bar{\hat{u}})$ with probability $\frac{1-\sqrt{5}}{2\sqrt{5}}$, where $\bar{\hat{u}}$ refers to the mean of $\hat{u}$.
3. Compute $\hat{J}^*$ using the bootstrap sample $\{\tilde{Y}_{it}^*, Z_{it}\}_{i=1, t=1}^{N, T}$, where $\tilde{Y}_{it}^* = m(Z_{it}; \beta) + \hat{u}_{it}^*$, with $m(Z_{it}; \beta)$ the parametric functional form of our choice, including linear, quadratic, or cubic function, that impose the null.

4. Repeat the step 2-3 B times to obtain $\hat{J}_1^*, \hat{J}_2^*, \dots, \hat{J}_B^*$. Reject H_0 if $\hat{J} > \hat{J}_{1-\alpha}^*$, where $\hat{J}_{1-\alpha}^*$ is the upper $(1 - \alpha)$ percentile value of its empirical distribution α is the significance level.

We set $\alpha = 0.05$ and $B = 399$, and implement the adaptive rule-of-thumb bandwidth selection to reduce computation burden. We choose the 2nd-order Gaussian kernel function, and we follow Dehnad (1987) to select bandwidth based on $h_{AROT} = 1.059A(NT)^{-1/5}$ where $A = \min\{\hat{\sigma}, IQR_z/1.34\}$ and IQR_z is the inter-quantile range of the variable Z .

Appendix 3 We employ a nonparametric test by Li (1996) for the distribution equality of estimated gradients of exports and imports between OECD and Non-OECD countries. Let $\mathbf{g}^1$ and $\mathbf{g}^2$ be vectors that contain the first-order gradient of export (or import) $\{g_{it}^1\}_{i,t=1}^{N_1,T}$ in one group of countries with observations N_1 and $\{g_{it}^2\}_{i,t=1}^{N_2,T}$ in another one with observations N_2 . We also assume that g^m is i.i.d. distributed for fixed t and strictly stationary for fixed i for $m = 1, 2$. Define their associated density function $q_1(\cdot)$ and $q_2(\cdot)$, we test the null hypothesis $H_{04} : q_1(\mathbf{g}^1) = q_2(\mathbf{g}^2)$ for almost all continuous variable $\mathbf{g}^1$. A test statistic with non-zero center term is constructed on the basis of integrated squared error between two different densities evaluates at the same random variable:

$$T = \int_{\mathbf{g}^1} [q_1(\mathbf{g}^1) - q_2(\mathbf{g}^1)]^2 d\mathbf{g}^1$$

which can be extended to:

$$T = \int_{\mathbf{g}^1} [q_1(\mathbf{g}^1)dQ_1(\mathbf{g}^1) + q_2(\mathbf{g}^1)dQ_2(\mathbf{g}^1) - q_1(\mathbf{g}^1)dQ_2(\mathbf{g}^1) - q_2(\mathbf{g}^1)dQ_1(\mathbf{g}^1)]d\mathbf{g}^1 \quad (2.17)$$

where $Q_1(\cdot)$ and $Q_2(\cdot)$ are the c.d.f. of $\mathbf{g}^1$ and $\mathbf{g}^2$, respectively, and T converges to zero if and only if the null is true. We replace the density $q_1(\cdot)$ and $q_2(\cdot)$ with their univariate kernel density estimator:

$$\hat{q}_1(g^1) = \frac{1}{N_1 T h_1} \sum_{i=1}^{N_1} \sum_{t=1}^T k_h(\hat{g}_{it}^1, \hat{g}^1) \quad (2.18)$$

$$\hat{q}_2(g^2) = \frac{1}{N_2 T h_2} \sum_{i=1}^{N_2} \sum_{t=1}^T k_h(\hat{g}_{it}^2, \hat{g}^2) \quad (2.19)$$

and select the bandwidth h_1 and h_2 based on adaptive rule-of-thumb h_{AROT} . Replacing (2.18) and

(2.19) into (2.17) we have the following feasible center-free test statistic:

$$\begin{aligned}\hat{T} &= \frac{1}{N_1^2|H|} \sum_{i=1}^{N_1} \sum_{t=1}^T \sum_{\substack{j=1 \\ j \neq i}}^{N_1} \sum_{\substack{k=1 \\ k \neq t}}^T k_h(\hat{g}_{it}^1, \hat{g}_{jk}^1) + \frac{1}{N_2^2|H|} \sum_{i=1}^{N_2} \sum_{t=1}^T \sum_{\substack{j=1 \\ j \neq i}}^{N_2} \sum_{\substack{k=1 \\ k \neq t}}^T k_h(\hat{g}_{it}^2, \hat{g}_{jk}^2) \\ &- \frac{1}{N_1 N_2 |H|} \sum_{i=1}^{N_1} \sum_{t=1}^T \sum_{\substack{j=1 \\ j \neq i}}^{N_2} \sum_{\substack{k=1 \\ k \neq t}}^T k_h(\hat{g}_{it}^1, \hat{g}_{jk}^2) - \frac{1}{N_2 N_1 |H|} \sum_{i=1}^{N_2} \sum_{t=1}^T \sum_{\substack{j=1 \\ j \neq i}}^{N_1} \sum_{\substack{k=1 \\ k \neq t}}^T k_h(\hat{g}_{it}^2, \hat{g}_{jk}^1)\end{aligned}$$

where $H = h_1 h_2$. The standardized version of T is given by:

$$\hat{J} = \sqrt{N_1 N_2 |H|} \left(\frac{\hat{T}}{\hat{\sigma}} \right) \xrightarrow{d} N(0, 1) \quad (2.20)$$

where

$$\begin{aligned}\hat{\sigma}^2 &= \frac{2}{N_1^2|H|} \sum_{i=1}^{N_1} \sum_{t=1}^T \sum_{\substack{j=1 \\ j \neq i}}^{N_1} \sum_{\substack{k=1 \\ k \neq t}}^T k_h^2(\hat{g}_{it}^1, \hat{g}_{jk}^1) + \frac{2}{N_2^2|H|} \sum_{i=1}^{N_2} \sum_{t=1}^T \sum_{\substack{j=1 \\ j \neq i}}^{N_2} \sum_{\substack{k=1 \\ k \neq t}}^T k_h^2(\hat{g}_{it}^2, \hat{g}_{jk}^2) \\ &- \frac{2}{N_1 N_2 |H|} \sum_{i=1}^{N_1} \sum_{t=1}^T \sum_{\substack{j=1 \\ j \neq i}}^{N_2} \sum_{\substack{k=1 \\ k \neq t}}^T k_h^2(\hat{g}_{it}^1, \hat{g}_{jk}^2) - \frac{2}{N_2 N_1 |H|} \sum_{i=1}^{N_2} \sum_{t=1}^T \sum_{\substack{j=1 \\ j \neq i}}^{N_1} \sum_{\substack{k=1 \\ k \neq t}}^T k_h^2(\hat{g}_{it}^2, \hat{g}_{jk}^1)\end{aligned}$$

The following provides procedure for a density equality test by Li (1996):

1. Given the original sample of gradients estimated in two different samples $\{\hat{g}_{it}^1, \hat{g}_{jt}^2\}_{i,j,t=1}^{N_1, N_2, T}$ obtained by the local linear estimator of $G(\cdot)$, compute $\hat{J}$.
2. To impose the null where two densities are identical in an L_2 norm sense, pool the original sample into $\Theta = \{\hat{g}_{11}^1, \dots, \hat{g}_{N_1 T}^1, \hat{g}_{11}^2, \dots, \hat{g}_{N_2 T}^2\}$, call it pool sample. Randomly select with replacement N_1 observations from the pool sample Θ for the estimation of $q_1(\cdot)$, and also randomly select with replacement N_2 observations from the pool sample Θ for the estimation of $q_2(\cdot)$. Call these samples $\{\hat{g}_{it}^{1*}\}_{i,t=1}^{N_1, T}$ and $\{\hat{g}_{it}^{2*}\}_{i,t=1}^{N_1, T}$, respectively.
3. Compute $\hat{J}^*$ in the same way as in step 1, except that the original sample $\{\hat{g}_{it}^1, \hat{g}_{jt}^2\}_{i,j,t=1}^{N_1, N_2, T}$ is replaced with the sample $\{\hat{g}_{it}^{1*}\}_{i,t=1}^{N_1, T}$ and $\{\hat{g}_{it}^{2*}\}_{i,t=1}^{N_2, T}$ in step 2.
4. Repeat the step 2-3 B times to obtain $\hat{J}_1^*, \hat{J}_2^*, \dots, \hat{J}_B^*$. Reject H_0 if $\hat{J} > \hat{J}_{1-\alpha}^*$, where $\hat{J}_{1-\alpha}^*$ is the upper $(1 - \alpha)$ percentile value of its empirical distribution α is the significance level.

We set $\alpha = 0.05$ and $B_1 = 399$, and implement the adaptive rule-of-thumb bandwidth selection to reduce computation burden.

Appendix 4 Selected Sample Countries: Whole sample

Country code	Country	Country code	Country
AUS*	Australia*	JPN*	Japan*
AUT*	Austria*	KEN	Kenya
BDI	Burundi	KOR	Republic of Korea
BEL*	Belgium*	KWT	Kuwait
BEN	Benin	LBN	Lebanon
BFA	Burkina Faso	LSO	Lesotho
BGR	Bulgaria	MAR	Morocco
BHR	Bahrain	MEX	Mexico
BHS	Bahamas	<u>MLT</u>	<u>Malta</u>
BOL	Bolivia	MNG	Mongolia
BRA	Brazil	MOZ	Mozambique
BRB	Barbados	MRT	Mauritania
BWA	Botswana	MUS	Mauritius
CAF	Central African Republic	MYS	Malaysia
CAN*	Canada*	NER	Niger
CHE*	Switzerland*	NGA	Nigeria
CHL	Chile	NLD*	Netherlands*
CHN	China	NOR*	Norway*
CIV	Cte d'Ivoire	NZL*	New Zealand*
CMR	Cameroon	OMN	Oman
COL	Colombia	PAN	Panama
CRI	Costa Rica	PER	Peru
CYP	Cyprus	PHL	Philippines
DEU*	Germany*	POL	Poland
DNK*	Denmark*	PRT*	Portugal*
DOM	Dominican Republic	PRY	Paraguay
ECU	Ecuador	QAT	Qatar
EGY	Egypt	ROU	Romania
ESP*	Spain*	<u>RWA</u>	<u>Rwanda</u>
FIN*	Finland*	SAU	Saudi Arabia
FJI	Fiji	SEN	Senegal
FRA*	France*	SGP	Singapore
GAB	Gabon	<u>SLE</u>	<u>Sierra Leone</u>
GBR*	United Kingdom*	<u>SWE*</u>	<u>Sweden*</u>
GRC*	Greece*	SWZ	Swaziland
GTM	Guatemala	TCD	Chad
HND	Honduras	TGO	Togo
HUN	Hungary	THA	Thailand
IDN	Indonesia	TTO	Trinidad and Tobago
IND	India	TUN	Tunisia
IRL*	Ireland*	TUR*	Turkey*
IRN	Iran (Islamic Republic of)	TZA	Tanzania: Mainland
ISL*	Iceland*	URY	Uruguay
ISR	Israel	USA*	United States*
ITA*	Italy*	VEN	Venezuela
JAM	Jamaica	ZAF	South Africa
JOR	Jordan	ZWE	Zimbabwe
Total	94		

Note: Countries that are starred, underlined, and bolded are , respectively, initial OECD countries, countries with missing values for labor regulation, and countries with missing values for human capital.

Appendix 5 Robustness Check: Whole sample

	Whole Sample											
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Openness	-0.020 (0.025)		-0.025 (0.021)		-0.025 (0.017)		-0.127*** (0.037)		-0.133*** (0.030)		-0.130*** (0.033)	
Openness ²							0.029*** (0.008)		0.027*** (0.008)		0.025*** (0.009)	
Ems		-0.097*** (0.026)		-0.123*** (0.031)		-0.192*** (0.027)		-0.388*** (0.058)		-0.310*** (0.064)		-0.533*** (0.070)
Ems ²								0.182*** (0.054)		0.127*** (0.064)		0.224*** (0.060)
Ims		0.038* (0.021)		0.054** (0.021)		0.089*** (0.023)		0.136* (0.082)		0.139* (0.085)		0.200* (0.105)
Ims ²								-0.053 (0.064)		-0.010 (0.073)		-0.077 (0.046)
FDI-ins		-0.003** (0.001)		-0.003 (0.003)		-0.002 (0.002)		-0.002*** (0.001)		-0.001 (0.002)		-0.002 (0.002)
FDI-outs		0.005*** (0.001)		0.001 (0.002)		0.001 (0.002)		0.003*** (0.001)		0.001 (0.001)		0.004*** (0.001)
SocialG		0.063 (0.046)		0.107*** (0.035)		0.061** (0.025)		0.060 (0.051)		0.070 (0.043)		0.013 (0.035)
PoliticalG		-0.051*** (0.017)		-0.049** (0.023)		-0.012 (0.029)		-0.073*** (0.018)		-0.027 (0.021)		-0.008 (0.032)
EconG: Barrier		-0.180*** (0.057)		-0.152*** (0.041)		-0.233*** (0.038)		-0.129*** (0.047)		-0.107 (0.040)		-0.171*** (0.033)
lnky		0.048*** (0.014)		0.044*** (0.017)		0.076*** (0.019)		0.053*** (0.014)		0.047*** (0.016)		0.079*** (0.021)
lninvp		0.092*** (0.016)		0.114*** (0.024)		0.142*** (0.025)		0.097*** (0.016)		0.117*** (0.023)		0.147*** (0.025)
Popg		0.008** (0.007)		0.021*** (0.005)		0.021*** (0.006)		0.006* (0.003)		0.021*** (0.006)		0.018*** (0.006)
lnRGDP		-0.032*** (0.007)		-0.019* (0.011)		-0.009 (0.012)		-0.027*** (0.007)		-0.019* (0.011)		0.006 (0.013)
lnoilr		-0.007 (0.006)		-0.009*** (0.003)		-0.001 (0.004)		-0.004 (0.007)		-0.014*** (0.003)		-0.007 (0.005)
laborReg				-0.019*** (0.004)		-0.018*** (0.005)		-0.007 (0.005)		-0.020*** (0.005)		-0.018*** (0.005)
HC						-0.028 (0.018)		-0.029* (0.016)				-0.041*** (0.012)
Cross-section Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
R ²	0.915	0.915	0.921	0.921	0.926	0.927	0.916	0.917	0.923	0.923	0.927	0.930
Adj-R ²	0.898	0.899	0.906	0.906	0.910	0.912	0.900	0.900	0.907	0.907	0.912	0.914
F-Stats	56.49***	55.47***	59.23***	57.94***	59.63***	59.44***	56.40***	54.19***	58.95***	56.27***	59.40***	59.00***
Total country Obs	37	37	35	35	32	32	37	37	35	35	32	32
Total Observation	296	296	280	280	256	256	296	296	280	280	256	256

Bibliography

- A. Giles, J., Williams, C. L., 2000. Export-led growth: a survey of the empirical literature and some non-causality results. part 1. *The Journal of International Trade & Economic Development* 9 (3), 261–337.
- Acemoglu, D., 2003. Labor-and capital-augmenting technical change. *Journal of the European Economic Association* 1 (1), 1–37.
- Amsler, C., Prokhorov, A., Schmidt, P., 2016. Endogeneity in stochastic frontier models. *Journal of Econometrics* 190 (2), 280–288.
- Azmat, G., Manning, A., Reenen, J. V., 2012. Privatization and the decline of labour’s share: international evidence from network industries. *Economica* 79 (315), 470–492.
- Bentolila, S., Saint-Paul, G., 2003. Explaining movements in the labor share. *Contributions in Macroeconomics* 3 (1).
- Bernard, A. B., Eaton, J., Jensen, J. B., Kortum, S., 2003. Plants and productivity in international trade. *American economic review* 93 (4), 1268–1290.
- Blanchard, O. J., Nordhaus, W. D., Phelps, E. S., 1997. The medium run. *Brookings Papers on Economic Activity* 1997 (2), 89–158.
- Boggio, L., Dall’Aglio, V., Magnani, M., 2010. On labour shares in recent decades: A survey. *Rivista Internazionale di Scienze Sociali*, 283–333.
- Brada, J. C., 2013. The distribution of income between labor and capital is not stable: But why is that so and why does it matter? *Economic Systems* 37 (3), 333–344.
- Buch, C., Monti, P., Toubal, F., 2008. Trade’s impact on the labor share: Evidence from german and italian regions. working paper, Institut für Angewandte Wirtschaftsforschung (IAW).

- Charpe, M., 2011. The labour share of income: Determinants and potential contribution to exiting the financial crisis. *World Employment and Social Outlook 2011* (1), 55–74.
- Daudey, E., García-Peñalosa, C., 2007. The personal and the factor distributions of income in a cross-section of countries. *The Journal of Development Studies* 43 (5), 812–829.
- Decreuse, B., Maarek, P., 2017. Can the hos model explain changes in labor shares? a tale of trade and wage rigidities. *Economic Systems* 41 (4), 472–491.
- Dehnad, K., 1987. *Density estimation for statistics and data analysis*.
- Dinopoulos, E., Segerstrom, P., 1999. A schumpeterian model of protection and relative wages. *American Economic Review* 89 (3), 450–472.
- Diwan, I., 2001. Debt as sweat: labor, financial crises, and the globalization of capital. mimeo, the World Bank.
- Elsby, M. W., Hobijn, B., Şahin, A., 2013. The decline of the us labor share. *Brookings Papers on Economic Activity* 2013 (2), 1–63.
- Giles, J. A., Williams, C. L., 2000. Export-led growth: a survey of the empirical literature and some non-causality results. part 2. *Journal of International Trade & Economic Development* 9 (4), 445–470.
- Goldberg, P., Khandelwal, A., Pavcnik, N., Topalova, P., 2009. Trade liberalization and new imported inputs. *American Economic Review* 99 (2), 494–500.
- Gollin, D., 2002. Getting income shares right. *Journal of political Economy* 110 (2), 458–474.
- Guerriero, M., Sen, K., 2012. What determines the share of labour in national income? a cross-country analysis. working paper.
- Guscina, A., 2006. Effects of globalization on labor's share in national income. *International Monetary Fund*.
- Gwartney, J., Lawson, R., Hall, J., 2015. *Economic freedom of the world 2015 annual report*. The Fraser Institute.
- Harrison, A., 2005. Has globalization eroded labors share? some cross-country evidence. working paper.
- Henderson, D. J., Parmeter, C. F., 2015. *Applied nonparametric econometrics*. Cambridge University Press.

- Hung, J. H., Hammett, P., 2016. Globalization and the labor share in the united states. *Eastern Economic Journal* 42 (2), 193–214.
- Hutchinson, J., Persyn, D., 2012. Globalisation, concentration and footloose firms: in search of the main cause of the declining labour share. *Review of World Economics* 148 (1), 17–43.
- Inklaar, R., Timmer, M., 2013. Capital, labor and tfp in pwt8. 0. University of Groningen (unpublished).
- Jaumotte, M. F., Tytell, I., 2007. How has the globalization of labor affected the labor income share in advanced countries? No. 7-298. International Monetary Fund.
- Jayadev, A., 2007. Capital account openness and the labour share of income. *Cambridge Journal of Economics* 31 (3), 423–443.
- Kalecki, M., 1938. The determinants of distribution of the national income. *Econometrica: Journal of the Econometric Society*, 97–112.
- Karabarbounis, L., Neiman, B., 2013. The global decline of the labor share. *The Quarterly Journal of Economics* 129 (1), 61–103.
- Keller, W., 1998. Are international r&d spillovers trade-related?: Analyzing spillovers among randomly matched trade partners. *European Economic Review* 42 (8), 1469–1481.
- Keller, W., 2002. Trade and the transmission of technology. *Journal of Economic growth* 7 (1), 5–24.
- Keynes, J. M., 1939. Relative movements of real wages and output. *The Economic Journal* 49 (193), 34–51.
- Lawless, M., Whelan, K. T., 2011. Understanding the dynamics of labor shares and inflation. *Journal of Macroeconomics* 33 (2), 121–136.
- Lawrence, R. Z., 2008. Blue collar blues: Is trade to blame for rising US income inequality? Vol. 85. Columbia University Press.
- Lawrence, R. Z., 2015. Recent declines in labor’s share in us income: a preliminary neoclassical account. Tech. rep., National Bureau of Economic Research.
- Li, Q., 1996. Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews* 15 (3), 261–274.
- Linton, O., Nielsen, J. P., 1995. A kernel method of estimating structured nonparametric regression based on marginal integration. *Biometrika*, 93–100.

- Luo, C., Zhang, J., 2010. Declining labor share: is china's case different? *China & World Economy* 18 (6), 1–18.
- Magee, L., 1998. Nonlocal behavior in polynomial regressions. *The American Statistician* 52 (1), 20–22.
- Marx, K., Engels, F., 1902. *Wage-labor and Capital*. New York Labor News Company.
- Melitz, M. J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71 (6), 1695–1725.
- Ohlin, B., 1952. *Interregional And International Trade*. Vol. 39. Harvard University Press.; Cambridge.
- Piketty, T., 2015. About capital in the twenty-first century. *American Economic Review* 105 (5), 48–53.
- Qian, J., Wang, L., 2012. Estimating semiparametric panel data models by marginal integration. *Journal of Econometrics* 167 (2), 483–493.
- Racine, J., 1997. Consistent significance testing for nonparametric regression. *Journal of Business & Economic Statistics* 15 (3), 369–378.
- Ricardo, D., 1817. *The works and correspondence of David Ricardo Vol. 1: On the principles of political economy and taxation*.
- Robinson, P. M., 1988. Root-n-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, 931–954.
- Rodriguez, F., Jayadev, A., 2010. The declining labor share of income. *Journal of Globalization and Development* 3 (2), 1–18.
- Rodríguez, F., Ortega, D., 2001. Openness and factor shares. documento presentado en el seminario La teoría del desarrollo en los albores del siglo XXI, Santiago de Chile 28.
- Rodrik, D., 1998. Has globalization gone too far? *Challenge* 41 (2), 81–94.
- Rodrik, D., 2006. What's so special about china's exports? *China & World Economy* 14 (5), 1–19.
- Severance-Lossin, E., Sperlich, S., 1999. Estimation of derivatives for additive separable models. *Statistics: A Journal of Theoretical and Applied Statistics* 33 (3), 241–265.
- Smith, A., 1776. *The wealth of nations* (new york: Modern library, 1937). Google Scholar, 423.
- Stockhammer, E., Onaran, Ö., Ederer, S., 2008. Functional income distribution and aggregate demand in the euro area. *Cambridge journal of Economics* 33 (1), 139–159.

- Stolper, W. F., Samuelson, P. A., 1941. Protection and real wages. *The Review of Economic Studies* 9 (1), 58–73.
- Taylor, T., 2013. Labor’s falling share, everywhere. [Blog Post]. Retrieved from <http://conversableeconomist.blogspot.com/2013/06/labors-falling-share-everywhere.html>.
- UNCTAD, 1994. World investment report 1994: Transnational corporations, employment and the workplace: an executive summary. United nations conference on trade and development.
- Wang, Z., Wei, S.-J., 2010. What accounts for the rising sophistication of china’s exports? In: *China’s growing role in world trade*. University of Chicago Press, pp. 63–104.
- Young, A. T., Lawson, R. A., 2013. Capitalism and Labor Shares : A Cross-Country Panel Study , 1970 to 2010 Capitalism and Labor Shares : A Cross-Country Panel Study , 1970 to 2010 (October), 1–31.
- Young, A. T., Lawson, R. A., 2014. Capitalism and labor shares: A cross-country panel study. *European Journal of Political Economy* 33, 20–36.
URL <http://dx.doi.org/10.1016/j.ejpoleco.2013.11.006>
- Young, A. T., Tackett, M. Y., 2017. Globalization and the decline in labor shares: Exploring the relationship beyond trade and financial flows. *European Journal of Political Economy*.
- Young, A. T., Zuleta, H., 2015. Do unions increase labor shares? evidence from us industry-level data. *Eastern Economic Journal*, 1–18.
- Zhang, X., Lu, Y., 2014. Effects of commodity trade structure variations on labor’s share of income in chinaan empirical study based on data of production industry sectors. *China Economist* 4, 009.
- Zheng, J. X., 1996. A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics* 75 (2), 263–289.

Chapter 3

Exporting Behavior and Labor Share in Chinese Manufacturing Industries: A Semiparametric Approach

3.1 Introduction

China has been actively engaged in global economic activities in the past four decades. As of 2016, China exported 2,098 billion dollars, making it the largest export economy and the “world’s factory” in the world. While China’s export has dramatically contributed to its outstanding economic growth, little is known about exports’ impact on labor in China behind the logo of “made in China”.

It is well documented that labor’s slice of pie has shrunk worldwide since the 1980s (?), with its associated reason commonly attributed to international trade (Harrison (2005);Jaumotte and Tytell (2007);Daudey and García-Peñalosa (2007)). This argument is often used by developed countries as evidence against trading with developing countries. For example, during his presidential campaign in 2016, Donald Trump criticized China’s export for “stealing” American workers’ jobs. Meanwhile, however, it is overlooked that during 1998-2007, labor share in China has experienced a drop of 12.45 percentage points (Bai and Qian (2010)), and in Chinese manufacturing industries a drop of 3 percentage points, the most tradeable sector in the nation in our sample (see Figure 3.1).¹

Among others, export in international trade plays a key role in explaining labor share. In fact, the current literature suggests that the effect of export can be decomposed into its neutral effect (i.e., effect

¹The detailed discussion of data is given in Section 3.

Figure 3.1: Labor Share in Chinese Manufacturing Industries: 1998-2007

Date Source: Chinese Annual Surveys of Industrial Production.

that depends only on exports) and non-neutral effect (i.e., effect that depends on other variables). On one hand, the neutral effect stems from the neoclassical trade theory of Heckscher-Ohlin model and Stolper-Samuelson theorem (Ohlin (1952); Stolper and Samuelson (1941)). It states that owners of capital gain higher return in capital-abundant countries (usually developed countries) that export capital-intensive goods, whereas labor gain higher returns in labor-abundant countries (usually developing countries) that export labor-intensive products. On the other hand, the non-neutral effect has been investigated by a wave of subsequent studies. We summarize that the main channels through which export may non-neutrally affect labor share include capital intensity (Piketty and Zucman (2014); Piketty et al. (2017); Lawrence (2015); Elsby et al. (2013)), monopoly power in imperfect competition (Azmat et al. (2012); Bentolila and Saint-Paul (2003)), and capital-augmenting technologies (Acemoglu (2003); Bentolila and Saint-Paul (2003)). Hence, the joint effect of export should be taken into account in studying Chinese labor share to avoid biased results.

Nonetheless, most studies in the Chinese labor share literature have paid attention only to the neutral effect of export (i.e., the partial effect of export is a constant), leaving the existing argument of its non-neutral effects unexplored (see, for example, Zhang and Lu (2014)). Since the neutral effect of export is expected to be positive in China with abundant labor, omitting the non-neutral effect prevents us from revealing whether the non-neutral effect further strengthens or weakens the neutral effect of export, thereby limiting the ability to infer policy implication.

In this paper, we contribute to the literature three-fold by empirically testing the effect of export on

labor's share of income in Chinese manufacturing industries. First, we employ a panel of 4,767 Chinese manufacturing firms from 1998 to 2007. It is a period of extensive trade that covers the nation's entry to the World Trade Organization (WTO). Using large firm-level data is likely to uncover the micro-level foundation of the relationship between trade and labor's income share. Therefore, we are able to examine the proportion of the slice of pie obtained by recent millions of workers in China by understanding its underlying determinants.

Second, we follow the aforementioned literature to decompose the total effect of export into its neutral effect and three non-neutral effects: capital intensity (proxied by capital-labor relative ratio), monopoly power in imperfect competition (proxied by markup), and capital-augmented technology (proxied by total factor productivity growth). Figure (3.2) plots the three variables during 1998-2007 for exporters and non-exporters in Chinese manufacturing firms, whose magnitude clearly depends on whether firms have exporting behavior. Hence, it serves as simple evidence that the three variables have potential impact on Chinese labor share through the level of exports (i.e., the non-neutral effects of export), which is thoroughly discussed and empirically tested in our study.²

Finally, we do not restrict the labor share regression to be a linear function of export to reduce the risk of model mis-specification that yields biased estimates. Instead, we implement a semiparametric varying coefficient model with fixed effect (VCM-FE) that allows for the coefficients of our explanatory variables (proxies of non-neutral effects) to depend on an unknown function of exports, which nests the linear OLS panel regression model as a special case. Inspired by Wang and Yang (2009) and Sun and Malikov (2018), we estimate VCM-FE using a spline-backfitted kernel estimator that reduces the finite sample bias and variance simultaneously, with their associated smoothing parameters chosen based on sound statistical theories. Another advantage of estimating VCM-FE is its ability to obtain firm-specific estimates of neutral and non-neutral export effect, therefore revealing the potential heterogeneity of export impact across firms.

We note that understanding the economic determinants of labor share is particularly important for China with export-led economic growth strategy, since it reveals the proportion of economics benefits that has been shared by labor along its participation in the global supply chain.³ Given China's high degree of international trade engagement, labor participation rate, and heterogeneity across firms' characteristics, the Chinese economy is an ideal laboratory to study the relationship between trade and labor's share of income. It also has policy implication to Chinese labor market, since the magnitude of Chinese labor share decline is important to understand China's income inequality (Zhou (2015)).

Following the introduction, Section 2 presents a brief literature summary, with the main emphasize on

²See the detailed discussion of the data in Section 3.

³We also refer to Atkinson (2009) for three key reasons for studying factor shares.

Figure 3.2: Exporters vs Non-exporters in capital intensity, total factor productivity, and markup.

Date Source: Chinese Annual Surveys of Industrial Production. *Method:* Authors' own calculation.

bridging the literature of trade and labor share to see how exports potentially affect labor share neutrally through its own channel and non-neutrally through three channels commonly discussed in labor share literature. Section 3 details the data implemented in our analysis, followed by the empirical methodologies in Section 4. Section 5 in turn presents and discusses empirical results, and Section 6 concludes with future research outlines.

3.2 Literature Background

3.2.1 Labor share and export in China

Labor's share of income is important for contemplating economic growth itself, since the extent to which the growth benefits of a nation has been enjoyed by labor and/or capital remains unclear in general. In political economics, labor share is always considered a proxy of labor's welfare, so the declining trend of labor share is often used by unions as an evidence that is against labor (Young and Zuleta (2015)). It may also serve to understand income inequality, given that capital is more likely to be concentrated in high-income households (Zhou (2015)).

Nonetheless, the constancy of labor share - the ratio of labor compensation to domestic output - has seemed to reach a consensus to economists until the late of 1980s when a declining trend of labor share is documented by Blanchard et al. (1997). This economic phenomenon occurs in developed countries as well as in traditionally "labor-intensive" developing countries (Boggio et al. (2010)).

China as a labor-intensive developing country achieves the growth miracle partially due to its export-

led growth strategy. A wave of studies that have focused on the performance and determinants of China's export-led growth (Amiti and Freund (2010); Guo and N'Diaye (2009); Girma et al. (2009); Jiang (2008)). Surprisingly, only a handful studies address the questions of how export may potentially affect labor share, albeit millions of workers participating in export sector (Luo and Zhang (2010)).

Bai and Qian (2010) employ the firm-level dataset similar in this paper, and find that it is the state-owned enterprises (SOEs) reconstruction and the monopoly power enhancement that significantly lower labor share in Chinese industries during 1998 to 2007. However, they do not consider the role of export in their paper, which is highly likely an omitted variable. Luo and Zhang (2010) perceive the increasing share of foreign firms in export industries and trade transformation towards capital-intensive products as the main reasons to explain the decline of labor share. While expecting a positive effect of export on China's labor share, they find that the positive effect is insignificant by implementing provincial-level data. Finally, Zhou (2015) argues that the change of industrial structure accounts for a minor part of reasons for the decline of labor share in industrial sector, although the study concentrates on the effect of import penetration in the time of export-import processing trade in Pearl River delta and export-led processing trade in Yangtze River, making the sample size to be restrictive.

To our best knowledge, none of the existing studies in China's labor share literature have explicitly study the effect of export via a micro-level foundation using large firm-level data. Also, the fact that the effect of export should be decomposed into its neutral and non-neutral effect is neglected by current studies. In this paper, we bridge the gap between the theory in trade and labor share by discussing in the following how exports influence labor share 1) through itself and 2) through non-trade variables commonly discussed in the labor share literature.

3.2.2 Links between trade and labor share

The literature of trade and labor share suggests that explanatory variables of labor share can be grouped into non-trade variables and trade variables. On one hand, non-trade variables are highly related to production process and mainly include capital intensity, capital-augmented technology, and the degree of monopoly power in non-perfect competition. First, capital intensity is argued in a theoretical model by Elsby et al. (2013), stating that labor's share of income is essentially related to the capital-labor ratio (i.e., capital-labor ratio): higher capital intensity would result in higher (lower) labor share provided a complementary (substitutable) relationship between labor and capital. Since capital and labor are found to be substitutable in developing countries (Tian and Wang (2018)), higher capital intensity is likely to decrease China's labor share.

Second, given imperfect competition, higher market power of firms would result in higher markup

that widens the gap between marginal product of labor (MPL) and real wage, which exerts downward pressure on labor share through labor bargaining power reduction (Bentolila and Saint-Paul (2003); Bai and Qian (2010); Rodrik (1998); Kalecki (1938)). Finally, Acemoglu (2003) argues that capital-augmenting technology may worsen labor share in the presence of trade, since labor’s bargaining power against the use of capital is likely to be deteriorated in the presence of “trade-induced” capital-biased (or labor-saving) technology.

On the other hand, trade variables, particularly exports, link to labor share through its *neutral* and *non-neutral* effect. The *neutral* effect stems from the classical Heckscher-Ohlin framework and Stolper-Samuelson Theorem (Ohlin (1952); Stolper and Samuelson (1941)), which predicts that a trade-induced change in product prices alters the real return for the factor intensively used in the production. Thus, owners of capital (labor) receive higher return (real wage) in capital-intensive (labor intensive) countries. Nonetheless, in this paper we argue that trade also possesses *non-neutral* effect that works through the three non-trade variables aforementioned, which are discussed in turn in the following.

First, exports and capital intensity may be positively related. A general consensus in the trade literature is that exporters are more capital intensive than non-exporters (Bernard and Jensen (1997, 1999)), which is consistent with our findings in Panel (a) of Figure (3.2). Given that capital intensity is likely to decrease China’s labor share given a substitutable relationship between capital and labor, an increase in export may further intensify the effect of capital intensity. We test the presence of this mechanism as the first non-neutral channel by estimating a general version of interaction between capital intensity and export.

In addition, trade may consolidate firms’ monopoly power under imperfect competition. Exporters usually charge higher markups than non-exporters (De Loecker and Warzynski (2012)), whereas Lu (2010) argues Chinese exporters sell less of products in domestic market than non-exporters. Hung and Hammett (2016) argue that globalization may induce special “know-how” firms to expand firm size and charge higher markup, since their products don’t have close substitutes. De Loecker and Warzynski (2012) also argues that exporter usually charge higher markup. This view is also consistent with the implication from “super-star” firm model by Dorn et al. (2017), which is used to explain the declining labor share in the US industries. Hence, trade is likely to enlarge the negative effect of firm’s monopoly power on labor share by eroding labor’s bargaining power. Therefore, we test this second possible non-neutral channel by interacting exports with a proxy for firms monopoly power (Bentolila and Saint-Paul (2003); Stockhammer (2013)).

Finally, the channels through which trade affects labor share via capital-augmented technology is less unambiguous. One side of the argument is that international trade has its ability to relax production

constraints and enhance productivity in general (i.e. technological spillovers), thereby resulting in a substantial increase in national income that betters off workers (Keller (1998); Keller (2002); Zhu and Jeon (2007)). However, another part of the literature perceives trade-induced technological innovation and mechanization as capital-augmenting rather than labor-augmenting technology, which lowers labor share when capital and labor are substitutable (Bentolila and Saint-Paul (2003); Karabarbounis and Neiman (2013); Dinopoulos and Segerstrom (1999)). Exporters are usually more productive than non-exporters (Aw et al. (2011); Bernard et al. (2003); Bernard and Jensen (1997)). In the case of China, it is also the case (Ma et al. (2014)) and the latter argument implies that international trade may amplify the negative effect of capital-augmenting technology. We test these two arguments as the last non-neutral channel by interacting exports and the proxy for capital-augmented technology in a general manner.

In sum, the current empirical literature between trade and labor share suggests their potential connection but awaits for its empirical evidence, particularity in China with the largest export sector in the world. Most cross-country studies in the literature only test the neutral effect of trade restricted in a linear or a quadratic functional form, making its non-neutral effects remain silent (Harrison (2005); Guscina (2006); Guerriero and Sen (2012)). Only a few studies have moved into this direction by examining the underlying interactive relationship between trade and labor share from a microeconomic foundation (Böckerman and Maliranta (2011); Berkowitz et al. (2016, 2017); Perugini et al. (2017)), although China is not considered in their sample. Hence, this paper fills the gap by providing a thorough empirical evidence of how firms' exports influence its workers and therefore render labor market policy implications for labors in the Chinese manufacturing industries.

3.3 Data

We employ data from Chinese Annual Surveys of Industrial Production (ASIP). It contains balance-sheet information on all state-owned enterprises (SOEs) and “above-scale” enterprises with annual sales of 5 million Chinese yuan (RMB) (611,995 dollars).⁴ We only look at manufacturing industries and exclude mining and public utility sectors. A list of 2-digit Chinese Industrial Classification (CIC) codes and names are provided in Appendix 2. Our main variable labor share is defined as labor's compensation (*wage+benefit*) divided by firms' value added, where *benefit* are bonus paid to labor at the end of each year. Our trade variable is *Export intensity*, defined as the ratio of firms' export to total output. Our non-trade variables include *Capital intensity*, the natural logarithm of real capital per worker; Total factor productivity (TFP), calculated based on Levinsohn and Petrin (2003) and Petrin et al. (2004) that

⁴The exchange rate of 8.17 RMB per USD is taken as the average nominal exchange rate between 1998 and 2007 from the Federal Reserve Bank.

measures technological progress ⁵; and *Markup*, the difference between firms' sales and cost divided by cost that captures the degree of monopoly power in the product market.

In order to avoid the impact of firm entry and exit on labor share, we curtail the sample to be a balanced panel with incumbent firms from 1998 to 2007. We also drop firms having missing values in our main variables and the highest and lowest 1 percentile of all variables to mitigate outliers bias. The final data contains a balanced panel of 4,767 firms from 1998 to 2007, resulting in 47,670 firm-year observations.

Table (3.1) reports the sample mean and standard deviation (in parentheses) for the whole firm-year observations and other dimensions. We observe that when the sample is split based on whether a firm has non-zero export intensity, exporters have higher labor share than whole firm-year observations and non-exporters by nearly 3 and 5 percentage points, respectively. Also, exporters are more capital-intensive and productive than non-exporters, consistent with classical trade theory (Bernard and Jensen (1997, 1999)). The fact that they have lower markup and thus lower monopoly power is consistent with the findings in Lu (2010).

In addition, the difference between these two groups is also found when the whole sample is divided in terms of nationality. We follow the literature to define a firm as foreign firm if more than 50% of the firms' paid-in capital is from foreign countries. We found that foreign firms engaging intensively in export sector, with 70% of observations of foreigner firms are exporters.⁶ Domestic firms are defined similarly and classified into three types of ownership: state-owned enterprises (SOEs), private, and collective. A comparison between foreign and domestic firms reveals that the former are two times higher than the latter in export intensity and capital intensity. Foreign firms also exhibit slightly higher (lower) markup (TFP) than domestic firms.

Furthermore, splitting the sample into the eastern (coastal) and inland regions show that eastern regions have higher labor share and export intensity, but lower capital intensity, markup and TFP than the inland. The results should not be so surprising, given that "heavy industries", such as car and electronic manufacturing and equipment industries that are usually capital intensive and productive, have moved into inland regions in recent years, leaving "light industries", such as textile and leather manufacturing industries, in the eastern area (Ma (2018)).

Finally, since China joined in World Trade Organization in November of 2001, we split sample before and after the year of 2002. There are no obvious differences in our main variables, except that the after-2002 sample shows slightly higher (lower) export intensity, capital intensity and TFP (labor share and

⁵The details can be found in the appendix of Tian (2018), all price deflators are taken from National Bureau of Statistics of China.

⁶See Tian (2018) for more detailed discussions about firm ownership in China.

Table 3.1: Descriptive Data Summary

Variables	Whole	Exporters	Non-exporters	Foreign firms	Domestic firms	East	Inland	Before 2002	After 2002
Labor share	0.293 (0.195)	0.321 (0.201)	0.272 (0.187)	0.297 (0.200)	0.291 (0.193)	0.301 (0.193)	0.268 (0.196)	0.296 (0.194)	0.291 (0.195)
Export intensity	0.227 (0.442)	0.539 (0.544)	0.000 (0.000)	0.447 (0.489)	0.158 (0.402)	0.256 (0.474)	0.141 (0.321)	0.218 (0.357)	0.233 (0.490)
Capital intensity	0.435 (0.482)	0.468 (0.516)	0.413 (0.454)	0.643 (0.624)	0.370 (0.406)	0.431 (0.483)	0.448 (0.480)	0.424 (0.500)	0.443 (0.469)
Markup	0.348 (0.628)	0.314 (0.445)	0.373 (0.732)	0.351 (0.553)	0.347 (0.650)	0.328 (0.533)	0.409 (0.848)	0.413 (0.759)	0.305 (0.520)
Total factor productivity	2.328 (0.822)	2.472 (0.824)	2.224 (0.804)	2.541 (0.829)	2.262 (0.809)	2.301 (0.809)	2.408 (0.855)	2.160 (0.767)	2.441 (0.838)
Observations (% in whole sample)	47670 (100%)	20061 (42)%	27609 (58)%	11366 (24)%	36304 (76)%	35,639 (54)%	12,031 (46)%	19,068 (41)%	28,602 (59)%

Source: Chinese Annual Surveys of Industrial Production 1998-2007

markup).⁷

In sum, the statistics indicate that exporters and non-exporters exhibit different patterns in labor share, capital intensity, firms' markup, and TFP. This is also the case with foreign and domestic firms and firms in different regions. We now turn to investigate how those differences due to export intensity can help to explain labor share movement in China by empirically testing its neutral and non-neutral effects.

3.4 Empirical Methodologies

One common empirical model considered to estimate the neutral and non-neutral effect of exports on Chinese labor share is a fixed-effect linear regression model (OLS-FE) with interaction terms between exports and the rest regressors:

$$LS_{it} = exint_{it}\alpha_0 + k_{it}(\alpha_1 + exint_{it}\alpha_4) + markup_{it}(\alpha_2 + exint_{it}\alpha_5) + TFP_{it}(\alpha_3 + exint_{it}\alpha_6) + \mu_i + v_{it} \quad (3.1)$$

where for $i = 1, \dots, n$ and $t = 1, \dots, T$ being the index of firms and time, LS is labor share, $exint$, k , $markup$ and TFP are export intensity, capital intensity, markup, and TFP growth, respectively. Variable μ_i is the firm-specific unobserved heterogeneity that is potentially correlated with covariates, and v_{it} is a zero-mean error term independent with regressors. Hence, exports in model (3.1) affects labor share *neutrally* through its own function $exint\beta_0$ and *non-neutrally* through k , $markup$, and TFP .

However, the effect of $exint$ in (3.1) may be biased, since our pre-specified linear functional form of both neutral and non-neutral effect of export may suffer from potential model mis-specification that yields inconsistent OLS estimates. To alleviate such risk, we generalize (3.1) by estimating the following semiparametric fixed-effect varying coefficient model (VCM-FE):

$$LS_{it} = m_0(exint_{it}) + k_{it}m_1(exint_{it}) + markup_{it}m_2(exint_{it}) + TFP_{it}m_3(exint_{it}) + \mu_i + v_{it} \quad (3.2)$$

where the neutral effect of export now works through $m_0(\cdot)$ and non-neutral effect through $m_s(\cdot)$, $s = 1, 2, 3$, and all functions $m_j(\cdot)$ are potentially nonlinear with respect to $exint$ for $j = 0, \dots, 3$. In other words, the partial effect of capital intensity, markup, or TFP growth on labor share is allowed to nonlinearly vary with the magnitude of export intensity. We note that model (3.2) nests model (3.1) as a special case when $\{m_j(\cdot)\}_{j=0}^3$ are all linear functions of $exint$.

The varying coefficient model in (3.2) has gained its increasing popularity in various empirical works

⁷This can be due to that we only consider incumbent firms, excluding the effect of firm entry and exit.

due to its flexibility of interactive effect estimation, efficiency of estimators, and easiness of result interpretation (see Park et al. (2015) for a survey). Recently, Sun et al. (2009) estimate (3.2) by constructing a kernel-smoothed version of orthogonal projection matrix that concentrates out the fixed effect while consistently estimating $m_j(\cdot)$. However, their estimation is computational demanding, which is inappropriate in our large sample. Estimator in Sun and Malikov (2018) involves estimation of (3.2) by a series estimator that reduces computational cost but may increase estimation inefficiency. Inspired by Wang and Yang (2009) who consider estimating an additive model via a combination of both kernel and series estimator, we propose the use of the following spline-backfitted kernel estimator to simultaneously reduce inefficiency and bias of estimators for $m(\cdot)$ and $m'(\cdot)$ in (3.2).

For notation simplicity, we let $y_{it} \equiv LS_{it}$, $z_{it} \equiv exint_{it}$, and $\mathbf{x} = [x_{1,it}, x_{2,it}, x_{3,it}] \equiv [k_{it}, markup_{it}, TFP_{it}]$. Hence, model (3.2) is equivalent to

$$y_{it} = X_{it}m(z_{it}) + \mu_i + v_{it} \quad (3.3)$$

where $X_{it} = [1, x_{1,it}, x_{2,it}, x_{3,it}]$, and $m(z_{it}) = [m_0(z_{it}), m_1(z_{it}), m_2(z_{it}), m_3(z_{it})]^\top$.

We consistently estimate $m(\cdot)$ and its derivative $m'(\cdot)$ via two steps. First, we estimate functions $m(\cdot)$ by its series estimator. Let $\{\phi_k(\cdot)\}_{k=1}^{L_n}$ be a sequence of cubic B-spline basis functions. For a point of $z \in Z$ with $Z = \{z_{it}\}_{i=1, t=1}^{n, T}$, we approximate the unknown function $m_j(z)$ by its series estimator $\tilde{m}_j(z)$ from

$$\tilde{m}_j(z) = \sum_{k=1}^{L_n} \phi_k(z)\beta_k, \quad j = 0, \dots, 3$$

where $L_n = J_n + m$ is the number of basis function, with J_n the interior knots placed on the range of Z , and m is the polynomial order of the basis function. In other words, we split the function $m_s(\cdot)$ into a total of $J_n + 1$ subintervals defined on the support of Z , and fit functions in each subinterval by a polynomial function of order m . Hence, we estimate $m_s(\cdot)$ using piece-wise polynomial function. The use of $J_n + m$ subintervals is necessary due to the recursive nature of B-spline function.⁸ Finally, we follow the literature to choose $m = 3$ in our study and assume that $J_n \rightarrow \infty$ and $J_n/n \rightarrow 0$ as $n \rightarrow \infty$.

In matrix form, we rewrite (3.3) as

$$Y = Q(X, \Phi_n(z))\boldsymbol{\eta} + D\boldsymbol{\mu} + V \quad (3.4)$$

where X_{it} and $\Phi_n(z_{it}) = [\phi_1(z_{it}), \dots, \phi_{L_n}(z_{it})]$ is the it^{th} row of X and $\Phi_n(z)$, respectively, $\boldsymbol{\eta} = [\beta_{0,L_n}^\top, \dots, \beta_{3,L_n}^\top]^\top$ is a $(4L_n \times 1)$ vector with $\beta_{j,L_n} = [\beta_{j,1}, \dots, \beta_{j,L_n}]^\top$, and the it^{th} row of $Q(X, \Phi_n(z))$ is given by $Q(X_{it}, \Phi_n(z_{it})) = [\Phi_n(z_{it}), x_{1,it} \otimes \Phi_n(z_{it}), x_{2,it} \otimes \Phi_n(z_{it}), x_{3,it} \otimes \Phi_n(z_{it})]$. Finally, we follow Sun et al.

⁸We refer readers to see Eubank (1999) for an extensive review of series estimation.

(2009) to assume $\sum_{i=1}^n \mu_i = 0$ so that $D = [-\iota_{n-1}, I_{n-1}]^\top \otimes \iota_T$, where I_{n-1} is a $(n-1) \times (n-1)$ identity matrix, and ι_e is a $(e \times 1)$ vector of ones.

Define $M_D = I_{nT} - D(D^\top D)^{-1}D^\top$ and let $W \equiv Q(X, \Phi_n(z))$, we “swipe out” the fixed effect in (3.4) by pre-multiplying M_D on its both sides to have $M_D Y = M_D W \boldsymbol{\eta} + M_D V$, so we obtain our first stage series estimator $\tilde{\boldsymbol{\eta}}$ from

$$\tilde{\boldsymbol{\eta}} = \underset{\{\boldsymbol{\eta}\}}{\operatorname{argmin}} (Y - W \boldsymbol{\eta})^\top M_D (Y - W \boldsymbol{\eta}) \quad (3.5)$$

and we approximate function $m_j(z)$ by $\tilde{m}_j(z) = \Phi_n(z) \tilde{\boldsymbol{\beta}}_{j,L_n}$.

However, $\tilde{m}_j(z)$ is not efficient, since functions in each subinterval are fitted by relatively small observation. Hence, in the second step we improve the estimation efficiency via a one-step kernel smoothing backfitting. For $s = 1, 2, 3$, let $\tilde{m}_{-1s}(z)$ ($\tilde{m}_{1s}(z)$) be a vector of $\tilde{m}(z) = [\tilde{m}_0(z), \tilde{m}_1(z), \tilde{m}_2(z), \tilde{m}_3(z)]^\top$ with the 1^{st} and s^{th} elements $[\tilde{m}_0(z), \tilde{m}_s(z)]$ removed (maintained). Also, let $X_{-1s,it}$ and $X_{1s,it}$ defined similarly. We efficiently estimate the function $m_0(z)$ and $m_s(z)$ with their derivate $m'_0(z)$ and $m'_s(z)$ by their corresponding local linear estimator $\hat{\delta}_{1s}(z) = [\hat{m}_0(z), \hat{m}_s(z)]^\top \equiv [\hat{a}_0, \hat{a}_1]^\top \equiv \hat{\mathbf{a}}^\top$ and $\hat{\delta}'_{1s}(z) = [\hat{m}'_0(z), \hat{m}'_s(z)]^\top \equiv [\hat{b}_0, \hat{b}_1]^\top \equiv \hat{\mathbf{b}}^\top$ from

$$(\hat{\mathbf{a}}^\top, \hat{\mathbf{b}}^\top) = \underset{\{\mathbf{a}^\top, \mathbf{b}^\top\}}{\operatorname{argmin}} \sum_{i=1}^n \sum_{t=1}^T [\tilde{y}_{it} - \tilde{\mu}_i - X_{1s,it} \mathbf{a} - X_{-1s,it} \otimes (z_{it} - z) \mathbf{b}]^2 k_h(z_{it}, z) \quad (3.6)$$

where $\tilde{y}_{it} = y_{it} - X_{-1s,it} \tilde{m}_{-1s}(z_{it})$, $\tilde{\mu}_i$ is the i^{th} observation in $\tilde{\boldsymbol{\mu}} = [-\iota_{n-1}^\top \tilde{\boldsymbol{\mu}}_{-1}, \tilde{\boldsymbol{\mu}}_{-1}^\top]^\top$ with $\tilde{\boldsymbol{\mu}}_{-1} = (D^\top D)^{-1} D^\top (Y - W \tilde{\boldsymbol{\eta}})$, and $k_h(z_{it}, z) = K(\frac{z_{it} - z}{h})$ is the univariate kernel function. We choose Gaussian kernel in our study.

The intuition behind our two-step estimator is that in the first step, we “hammer down” the entire function $m_s(\cdot)$ into many pieces and then fit data within each piece using a polynomial function (i.e., the series estimator). However, doing so results in higher variance of the function estimates in each subinterval, which are obtained using the information provided by the data only within a small interval. Thus, we improve the estimation efficiency in the second step by “smoothing” the function’s shape from the first step using kernel function (i.e., the kernel estimator). We do so by first constructing a new dependent variable $\tilde{y}$ that can be only explained by x_s and z . We then estimate $[m_0(\cdot), m_s(\cdot)]^\top$ by a local linear estimator in (3.6). It is termed “local linear” because for each evaluation point $z \in Z$, we “locally” fit a “linear” line through z based on the its neighbor observation near to z . The nearness is controlled by the bandwidth h in the kernel function $k_h(z_{it}, z)$, which assigns higher (lower) weight for points closer to (further away from) z . Therefore, a larger bandwidth implies more equal weight assigned to the neighbor points of z , and vice versa. Finally, we connect each of these linear lines to obtain the estimated function as well as its derivative $[m'_0(\cdot), m'_s(\cdot)]^\top$. Note that local linear estimator includes OLS

estimator as a special case, since if $m_j(\cdot)$ is truly linear, the local linear estimator fits data using all sample points as neighbor points of z , which requires $h \rightarrow \infty$. Hence, our two-step estimator allows the estimated function $m_j(\cdot)$ to be properly splitted and smoothed in a sense that the mean squared error (MSE) of estimates are minimized.

We note that the most crucial factors in our estimators are J_n and h , which not only control for the range for which data we locally fit, but also for the finite sample bias and variance the estimator possesses: a higher (lower) J_n or h will decrease (increase) the variance but increase (decrease) the bias of series or local linear estimator. This is a well-known bias-variance trade-off in nonparametric regression estimator. Hence, the selection of J_n and h must rely on sound statistical theories so as to mitigate bias and variance in a MSE sense. We follow Craven and Wahba (1978) to choose J_{gcv} based on generalized cross-validation:

$$J_{gcv} = \underset{\{J_n\}}{\operatorname{argmin}} \frac{\frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T [y_{it} - X_{it} \tilde{m}(z_{it}) - \tilde{\mu}_i]^2}{(1 - J_n/nT)^2} \quad (3.7)$$

and choose h_{cvls} based on least-square cross validation:

$$h_{cvls} = \underset{\{h\}}{\operatorname{argmin}} \sum_{i=1}^n \sum_{t=1}^T [\tilde{y}_{it} - \tilde{\mu}_i - \hat{m}_{1s,-i}(z_{it}) - X_{1s,it} \hat{m}_{1s,-i}(z_{it})]^2 \quad (3.8)$$

where $\hat{m}_{1s,-i}(\cdot)$ is a “leave-one-firm-out” local linear estimator that assigns $k_h(z_{it}, z_{mg}) = 0$ in (3.6) whenever $i = m$.⁹ We now turn to present and discuss the estimation result.

3.5 Results Discussions

We highlight that due to the nature of our local estimator in VCM-FE model, our estimates are observation-specific as opposed to the global estimator in linear regression. Also, to clearly see the impact of export intensity in (3.2), we define terms in the following that are of our interest. For a particular point $exint$, we first define its *neutral effect* as the function value $m_0(exint)$, with its *neutrally partial effect* $\frac{\partial m_0(exint)}{\partial exint} \equiv m'_0(exint)$. Intuitively, the *neutrally partial effect* is the partial impact of $exint$ on labor share by assuming that the non-neutral mechanisms through k , *markup*, and *TFP* do not play a role in determining labor share. Second, we define *non-neutral effect* of export intensity as functions $m_j(\cdot)$, which are the partial effect of k , *markup*, and *TFP*. Finally, we define *totally partial effect* as $\frac{\partial LS}{\partial exint} = \frac{\partial m_0(exint)}{\partial exint} + k_{it} \frac{\partial m_1(exint)}{\partial exint} + markup_{it} \frac{\partial m_2(exint)}{\partial exint} + TFP_{it} \frac{\partial m_3(exint)}{\partial exint}$, which consists of both *neutrally partial effect* (i.e., $\frac{\partial m_0(exint)}{\partial exint}$) and *non-neutral partial effect* through the term

⁹The essential reason of using leave-one-firm-out cross validation is to ensure that a firm i with its observations $z_{i1}, \dots, z_{iT}$ as outliers can be entirely removed so as to prevent h from not being too small. See chapter 11 in Henderson et al. (2015) for the related discussion.

$k_{it} \frac{\partial m_1(exint)}{\partial exint} + markup_{it} \frac{\partial m_2(exint)}{\partial exint} + TFP_{it} \frac{\partial m_3(exint)}{\partial exint}$. Therefore, the variation of labor share is explained by the joint impact of export intensity that works neutrally through $m_0(\cdot)$ and non-neutrally through the other three channels $m_s(\cdot)$.

3.5.1 Whole sample

We first report the results for whole sample in Table (3.3) by VCM-FE. The estimated neutrally partial ($\hat{m}'_0(z)$) and non-neutrally effect ($\hat{m}(z)$) are reported in each columns, conditioning on the level of export intensity at three quantile (0.25, 0.5, 0.75) and two percentiles near to the boundary (0.1, 0.9). Neutrally partial effect of export intensity are significant at almost 1% except when $exint$ exceeds 75%. Also, it exhibits a diminishing return, as it declines by 66% as $exint$ rises from 0.1 to 0.9, with its mean of 0.1009. It indicates that when the effect of *capital intensity*, *TFP*, and *markup* are irreverent in the regression (i.e., non-neutral effects of $exint$ are zero), one-percentage point increase in export intensity on average leads to a 0.1 of a percentage-point increase in labor share. The results support in general the Stolper-Samuelson theorem behind its neutral effect, predicting that exports increase labor share in China with labor-abundant production function, but assuming that its potential non-neutral effects is absent.

However, all non-neutral effects are highly significant and uniformly negative at 1% level over the observations of $exint$. We found that the negative non-neutral effect through capital intensity ($\hat{m}_1(z)$) significantly lowers labor share with its mean effect of -0.0373. This is in line with our hypothesis regarding the non-neutral effect through capital intensity, suggesting that higher exports hurt workers' welfare through the channel of capital due to their substitutable relationship in Chinese manufacturing industries. The negative non-neutral effect $\hat{m}_3(z)$ indicates that export intensity amplifies the effect of TFP on labor's share of income with its mean effect of -0.1965. Given the ambiguous effect of capital-augmented technology (proxied by TFP) on labor share discussed in Section 2, our results supports its second argument, arguing that capital-biased technological progress induced by trade erodes labor's bargaining power and thus shrinks their slice of pie (Acemoglu (2003); Karabarbounis and Neiman (2013)). Finally, the negative non-neutral effect through markup ($\hat{m}_2(z)$) meets our expectation, with its mean effect of -0.0109 and its magnitude slightly decreasing as export intensity rises. It suggests that "know-how" firms in trade are likely to have higher market power in imperfect competition, thus charge higher markup that weakens labor share through widening the MPL and real wage difference.

Our results above indicate that the effect of exports is jointly explained by its neutrally partial and non-neutral effects. We thus reports the totally partial effect (TPE) of export intensity $\frac{\partial LS}{\partial exint}$ by VCM-FE and OLS-FE in the middle panel of Table (3.3) for comparison purpose. The estimated TPE by VCM-FE and by OLS-FE are -0.011 and -0.072, respectively, implying that the positive neutral effect of export is

Table 3.2: Neutrally Partial and Non-neutrally Effects from OLS-FE Estimation

Variable	Coef	Whole	Exporters	Domestic	Foreign	Before 2002	After 2002	East	Inland
<i>expint</i>	$\hat{\alpha}_0$ (<i>NP</i>)	0.030*** (0.005)	0.0005 (0.009)	0.092*** (0.009)	0.032*** (0.010)	0.071** (0.035)	0.033*** (0.006)	0.025*** (0.005)	0.066*** (0.014)
<i>k</i>	$\hat{\alpha}_1$	-0.034*** (0.001)	-0.047*** (0.003)	-0.027*** (0.001)	-0.065*** (0.003)	-0.131*** (0.006)	-0.045*** (0.001)	-0.038*** (0.001)	-0.023*** (0.002)
<i>markup</i>	$\hat{\alpha}_2$	-0.006*** (0.001)	-0.019*** (0.005)	-0.004*** (0.001)	-0.015*** (0.005)	-0.008** (0.004)	0.001 (0.002)	-0.010*** (0.001)	0.0005 (0.001)
<i>TFP</i>	$\hat{\alpha}_3$	-0.190*** (0.001)	-0.220*** (0.003)	-0.186*** (0.001)	-0.204*** (0.003)	-0.939*** (0.006)	-0.213*** (0.001)	-0.197*** (0.001)	-0.174*** (0.002)
<i>exint</i> $\times$ <i>k</i>	$\hat{\alpha}_4$	-0.011*** (0.002)	-0.009*** (0.003)	-0.010*** (0.002)	-0.008** (0.004)	0.013 (0.011)	-0.009*** (0.002)	-0.009*** (0.002)	-0.030*** (0.005)
<i>exint</i> $\times$ <i>markup</i>	$\hat{\alpha}_5$	-0.014*** (0.003)	-0.003 (0.006)	-0.024*** (0.007)	-0.009 (0.007)	0.012 (0.012)	-0.014*** (0.004)	-0.013*** (0.004)	-0.010*** (0.004)
<i>exint</i> $\times$ <i>TFP</i>	$\hat{\alpha}_6$	-0.025*** (0.002)	-0.010*** (0.003)	-0.039*** (0.003)	-0.023*** (0.004)	-0.002 (0.013)	-0.030*** (0.002)	-0.020*** (0.002)	-0.041*** (0.005)
Mean of NN(<i>k</i>)		-0.036	-0.052	-0.029	-0.068	-0.128	-0.047	-0.040	-0.026
Mean of NN(<i>markup</i>)		-0.008	-0.020	-0.007	-0.019	-0.006	-0.002	-0.013	-0.001
Mean of NN(<i>TFP</i>)		-0.195	-0.225	-0.190	-0.213	-0.940	-0.218	-0.201	-0.179
Fixed Effect		Y	Y	Y	Y	Y	Y	Y	Y
Firms		4,767	1,260	3,356	887	4,767	4,767	3,762	1,196
Observations		47,670	12,600	33,560	8,870	19,068	28,602	37,620	11,960
Adjusted R ²		0.474	0.488	0.475	0.493	0.621	0.492	0.482	0.466
F Statistic		6,828.26***	1,898.47***	4,819.61***	1,358.82***	5,141.70***	4,640.12***	5,546.05***	1,662.30***

Note: Results are estimate of OLS-FE from 3.1. NP is short for neutrally partial effect of export intensity. Mean of NNP stands for the average of neutrally partial effect of export and non-neutral effect through *k*, *markup*, and *TFP*. *p<0.1; **p<0.05; ***p<0.01

Table 3.3: Neutrally and Non-neutrally Partial Effect: Whole sample and exporters

<i>exint</i>	Whole sample (exp: 42.08%)				Exporters (exp: 100%)			
	NP	NN(<i>k</i> , <i>markup</i> , <i>TFP</i>)			NP	NN(<i>k</i> , <i>markup</i> , <i>TFP</i>)		
	$\hat{m}'_0(z)$	$\hat{m}_1(z)$	$\hat{m}_2(z)$	$\hat{m}_3(z)$	$\hat{m}'_0(z)$	$\hat{m}_1(z)$	$\hat{m}_2(z)$	$\hat{m}_3(z)$
0.1	0.1822*** (0.038)	-0.0380*** (0.003)	-0.0203*** (0.008)	-0.1922*** (0.003)	0.0221*** (0.009)	-0.0439*** (0.006)	-0.0271*** (0.009)	-0.2217*** (0.006)
0.25	0.1619*** (0.038)	-0.0397*** (0.003)	-0.0243*** (0.008)	-0.2002*** (0.004)	0.0148* (0.009)	-0.0527*** (0.005)	-0.0276*** (0.009)	-0.2167*** (0.006)
0.5	0.1214*** (0.038)	-0.0419*** (0.003)	-0.0226*** (0.009)	-0.2035*** (0.004)	0.0235*** (0.009)	-0.0524*** (0.005)	-0.0354*** (0.012)	-0.2172*** (0.006)
0.75	0.0833*** (0.039)	-0.0447*** (0.003)	-0.0227*** (0.006)	-0.2157*** (0.004)	0.0257*** (0.010)	-0.0486*** (0.004)	-0.0438*** (0.010)	-0.2256*** (0.005)
0.9	0.0617 (0.039)	-0.0490*** (0.003)	-0.0214*** (0.006)	-0.2238*** (0.004)	0.0137*** (0.029)	-0.0529*** (0.004)	-0.0313*** (0.008)	-0.2303*** (0.005)
Mean of NP/NN	0.1009	-0.0373	-0.0109	-0.1965	0.0203	-0.0501	-0.0330	-0.2177
<i>Total partial effect</i>								
VCM-FE	-0.0108				0.0148			
OLS-FE	-0.0714				-0.4949			
<i>N</i>	4,767				1,260			
Total obs:	47,670				12,600			
<i>R</i> ² :	0.3545				0.3497			
<i>h</i> _{<i>cvs</i>}	0.2617				0.2214			
<i>J</i> _{<i>gcv</i>} :	4				4			
Knots space	0.2712				0.2319			

Note: Results are estimate of VCM-FE from 3.2. Mean of NP and Mean of NNP stands for, respectively, the average of neutrally partial effect of export and non-neutral effect through *k*, *markup*, and *TFP*. First and second stage estimators are cubic B-spline series estimator and local linear kernel estimator, with their corresponding optimal knots *J*_{*gcv*} and bandwidth *h*_{*cvs*} reported on the lower panel. *p<0.1; **p<0.05; ***p<0.01

dominated by its non-neutral effect through the channels of capital intensity, markup, and technological progress. While the difference between the two estimated TPE seems large, it is likely due to the linear model mis-specification in (3.1). To clearly see this point, in Figure (3.3) we plot the neutral effect in VCM-FE by $\hat{m}_0(z)$ against that in OLS-FE by $exint_{it}\alpha_0$, and plot the non-neutral effects in VCM-FE by $\{\hat{m}_s(z)\}_{s=1}^3$ against that in OLS-FE by $\hat{\beta}_1 = \hat{\alpha}_1 + exint_{it}\hat{\alpha}_4$, $\hat{\beta}_2 = \hat{\alpha}_2 + exint_{it}\hat{\alpha}_5$, and $\hat{\beta}_3 = \hat{\alpha}_3 + exint_{it}\hat{\alpha}_6$ obtained in Table (3.2). We notice that the source of potential model-misspecification mainly stems from the estimation of the neutral effect, with its OLS-FE estimates (with the mean of 0.0056) highly deviates from its VCM-FE counterpart and close to zero. Hence, the neutral effect in OLS-FE is likely to be underestimated. Finally, the non-neutral effects are slightly underestimated by OLS-FE but fairly close to its VCM-FE results in terms of the mean of non-neutrally effects (i.e., mean of NN) reported in Table (3.2) and (3.3). Clearly, it is the non-neutral effect of exports through TFP growth that has the most detrimental effect on Chinese labor share in manufacturing industries.

3.5.2 Exporters only

Given that 58% of our observations are non-exporters with zero entry in export intensity, we examine the effect of exports on labor share in the sample with non-exporters excluded. The results are present in Table (3.3), with their corresponding plots against OLS-FE counterparts shown in Figure (3.4).¹⁰ Since

¹⁰Notice that our method only allows for balanced panel, so the observations for exporters drop to 12,600. However, this does not reduce our estimation efficiency, since the source of estimation inefficiency in VCM-FE arises for higher dimension

Figure 3.3: Neutral and Non-neutral effect: whole sample

the estimated neutrally partial $\hat{m}'(z)$ and non-neutral effect $\hat{m}_s(z)$ are essentially the upper portion of that from the result in whole sample, we only discuss the change of the estimates rather than their trend. First, the mean of $\hat{m}'_0(z)$ drops down from 0.1009 in whole sample to 0.0203, suggesting that exports may increase labor share by a large magnitude only when export intensity is relatively low. Non-neutral effect through the three variables are also enlarged. The non-neutral effect through capital intensity $\hat{m}_1(z)$ are monotonically negative, with its mean of -0.0501 larger than that in whole sample by a factor of two. Similar conclusion holds in the non-neutral effects through TFP $\hat{m}_3(z)$ (with mean of -0.2177) and through markup $\hat{m}_2(z)$ (with mean of -0.033). We note that the totally partial effect of export is positive and given by 0.0148. The reason can be seen from Figure (3.2) that exporters charge relatively low markup, partially offsetting its associated negative effect $\hat{m}_2(z)$ as opposed to the case in whole sample (recall the formula of totally partial effect in Section 4). However, the negative non-neutral effects on labor share in exporters are not offset by its positive neutral effect.

of z , which is clearly not the case in this paper. We do the same for other sub-groups as well.

Figure 3.4: Neutral and Non-neutral effect: exporters

Figure 3.5: VCM-FE Results comparison: domestic and foreign firms

Table 3.4: Varying coefficient estimation results: domestic and foreign firms

<i>exint</i>	Domestic firms (exp: 31.4%)				Foreign firms (exp: 72.16%)			
	NP	NN(<i>k, markup, TFP</i>)			NP	NN(<i>k, markup, TFP</i>)		
	$\hat{m}'_0(z)$	$\hat{m}_1(z)$	$\hat{m}_2(z)$	$\hat{m}_3(z)$	$\hat{m}'_0(z)$	$\hat{m}_1(z)$	$\hat{m}_2(z)$	$\hat{m}_3(z)$
0.1	0.1375*** (0.024)	-0.0308*** (0.003)	-0.0158** (0.007)	-0.1924*** (0.004)	0.1356*** (0.017)	-0.0652*** (0.004)	-0.0193 (0.015)	-0.2009*** (0.006)
0.25	0.0986*** (0.025)	-0.0317*** (0.004)	-0.0208** (0.009)	-0.1987*** (0.005)	0.1073*** (0.019)	-0.0705*** (0.007)	-0.0361* (0.020)	-0.2061*** (0.008)
0.5	0.0741*** (0.027)	-0.0294*** (0.004)	-0.0149* (0.009)	-0.2045*** (0.005)	0.0851*** (0.019)	-0.0689*** (0.006)	-0.0379*** (0.014)	-0.2105*** (0.007)
0.75	0.0322 (0.029)	-0.0329*** (0.004)	-0.0205** (0.009)	-0.2144*** (0.005)	0.0836*** (0.022)	-0.0685*** (0.006)	-0.0306*** (0.013)	-0.2223*** (0.006)
0.9	0.0548* (0.032)	-0.038*** (0.004)	-0.0296*** (0.011)	-0.2226*** (0.005)	0.0835*** (0.037)	-0.0705*** (0.006)	-0.0234*** (0.007)	-0.2331*** (0.006)
Mean of NP/NN	0.0834	-0.0312	-0.0201	-0.2041	0.0988	-0.0687	-0.0295	-0.2146
<i>Total partial effect</i>								
VCM-FE	-0.0192				0.0101			
OLS-FE	-0.2960				-0.4588			
<i>N</i>	3,356				1,260			
Total obs:	33,560				8,870			
<i>R</i> ²	0.3519				0.4975			
<i>h_{cvtls}</i>	0.4711				0.5982			
<i>J_{gcv}</i>	2				2			
Knots space	0.5563				0.6677			

3.5.3 Domestic vs foreign firms

It is known that foreign firms are more likely to engage in exporting behavior. In our sample, more than 70% of observations in foreign firms are exporters. We proceed to split the sample into foreign and domestic firms according to the share of paid-in capital, and report the result in Table (3.4).¹¹ For the purpose of comparison, we plot the estimated neutral effect $\hat{m}(z)$ and non-neutral $\hat{m}(z)$ against export intensity for both groups in Figure (3.5). One obvious difference lies on the neutrally partial effect, which is uniformly higher in foreign firms with its mean of 0.0988 compared to 0.0834 in domestic firms. This is also in line with the function of neutral effect $\hat{m}_0(z)$ plotted in the first panel of Figure (3.5). Hence, exports directly tilts relatively higher income share toward labor in foreign than in domestic firms. Another significant difference shows up in the non-neutral effect through capital intensity $\hat{m}_1(z)$ and can be seen from the upper-right panel of Figure (3.5). The mean of which is -0.0687 in foreign firms, but is shrunk by half to -0.0312 in domestic firms. Therefore, given that foreign firms are more capital intensive from Table (3.1), capital does more harm to worker's share of income in foreign firms through trade due to labor substitution. The TFP and markup in domestic firms generates slightly fewer negative impact on labor share through export intensity than foreign firms, with their average given by (-0.0201, -0.2041) in domestic and (-0.0295, -0.2146) in foreign, respectively. As a result, We do not find distinguished difference for the non-neutral effect of TFP and markup between the two types of firms. Lastly, the neutral effect of export that are highly significant and positive is largely offset by its corresponding negative non-neutral effects, pulling down the total partial effect to 0.0101 in foreign firms

¹¹Please refer to Tian (2018) for details regarding to ownership types.

Table 3.5: Varying coefficient estimation results: east and inland

	East (exp: 46.61%)				Inland (exp: 30.53%)			
	NP	NN($k, markup, TFP$)			NP	NN($k, markup, TFP$)		
	$\hat{m}_0(z)$	$\hat{m}_1(z)$	$\hat{m}_2(z)$	$\hat{m}_3(z)$	$\hat{m}_0(z)$	$\hat{m}_1(z)$	$\hat{m}_2(z)$	$\hat{m}_3(z)$
$exint$								
0.1	0.2916*** (0.034)	-0.0370*** (0.003)	-0.0276*** (0.008)	-0.2010*** (0.004)	0.0930*** (0.023)	-0.0365*** (0.005)	-0.0048 (0.010)	-0.1765*** (0.006)
0.25	0.2044*** (0.035)	-0.0396*** (0.003)	-0.0324*** (0.008)	-0.2041*** (0.004)	0.0144 (0.025)	-0.043*** (0.006)	-0.0069 (0.013)	-0.1817*** (0.008)
0.5	0.1764*** (0.035)	-0.0435*** (0.004)	-0.0275*** (0.010)	-0.2068*** (0.005)	0.0856*** (0.029)	-0.0400*** (0.007)	-0.005 (0.020)	-0.1934*** (0.010)
0.75	0.1351*** (0.042)	-0.0449*** (0.003)	-0.0369*** (0.009)	-0.2185*** (0.004)	-0.1425*** (0.032)	-0.0483*** (0.007)	-0.0082 (0.006)	-0.2025*** (0.010)
0.9	0.1009* (0.058)	-0.0489*** (0.003)	-0.0381*** (0.010)	-0.2252*** (0.004)	0.2507*** (0.048)	-0.0519*** (0.007)	-0.0101 (0.009)	-0.2112*** (0.009)
Mean of NP/NN	0.1703	-0.0415	-0.0306	-0.2111	0.0864	-0.0439	-0.007	-0.1931
Total partial effect								
VCMFE-SBK	0.0720				0.0419			
OLSFE	-0.3890				-0.3899			
N	4,767				4,767			
Total obs:	35,480				11,960			
R^2	0.3650				0.3212			
J_{gcv}	3				5			
h_{cvls}	0.3420				0.3784			
Knots space	0.412				0.4795			

and even to -0.0192 in domestic firms.

Figure 3.6: Gradients comparison: east and inland

Table 3.6: Varying coefficient estimation results: before and after 2002

	Before 2002 (exp: 40.19%)				After 2002 (exp: 43.35%)			
	NP	NN(k , markup, TFP)			NP	NN(k , markup, TFP)		
	$\hat{m}'_0(z)$	$\hat{m}_1(z)$	$\hat{m}_2(z)$	$\hat{m}_3(z)$	$\hat{m}'_0(z)$	$\hat{m}_1(z)$	$\hat{m}_2(z)$	$\hat{m}_3(z)$
$exint$								
0.1	0.1881*** (0.025)	-0.0237*** (0.004)	0.0086 (0.009)	-0.2348*** (0.006)	0.2050*** (0.021)	-0.0468*** (0.003)	-0.0072 (0.007)	-0.2223*** (0.004)
0.25	0.1404*** (0.025)	-0.0242*** (0.004)	-0.0053 (0.010)	-0.2393*** (0.006)	0.1589*** (0.022)	-0.0516*** (0.004)	0.0004 (0.007)	-0.2314*** (0.004)
0.5	0.0923*** (0.026)	-0.0277*** (0.005)	-0.014 (0.010)	-0.2376*** (0.007)	0.1162*** (0.025)	-0.057*** (0.004)	0.0051 (0.007)	-0.2383*** (0.005)
0.75	0.0778*** (0.027)	-0.0231*** (0.005)	0.0095 (0.013)	-0.2447*** (0.007)	0.0940*** (0.030)	-0.0626*** (0.004)	-0.0196*** (0.009)	-0.2523*** (0.005)
0.9	0.0872*** (0.036)	-0.0275*** (0.004)	0.0041 (0.007)	-0.2550*** (0.006)	0.0738*** (0.046)	-0.0653*** (0.004)	-0.0216*** (0.008)	-0.2576*** (0.005)
Mean of NP/NN	0.1174	-0.0299	0.0033	-0.2319	0.1293	-0.0535	-0.0068	-0.2355
Total partial effect								
VCM-FE	0.0248				0.0035			
OLS-FE	-1.8411				-0.4400			
N	4,767				4,767			
Total obs:	19,068				28,602			
R^2	0.3282				0.3738			
J_{gcv}	3				4			
h_{cvls}	0.3142				0.3766			
Knots space	0.4001				0.4335			

Figure 3.7: Gradients comparison: before and after 2002

3.5.4 Eastern and inland Region

We then take into account for the potential heterogeneity in regions. In China, most exporters in *light* manufacturing firms, such as textile and leather industries (See CICC code 17 and 19 in the appendix), are labor-dependent and thus labor-intensive industries with their strong preference to reside in eastern (coastal) region due to the convenience of transportation. By contrast, *heavy* industries, such as metal product industries (CICC 34) and car manufacturing industry (CICC 37) have moved into inland region with less population density and wide ground for construction. We therefore split the sample into eastern and inland region on the basis of regional code by National Bureau of Statistics of China (NBSC) and is reported in Appendix 3. Results are reported in Table (3.5) with its plot on Figure (3.6)

The neutral effect of export moderately benefits labors in manufacturing firms in eastern region, with its mean of 0.1703 compared to that of 0.0864 in inland region. Unlike our previous findings, exports does not lower labor share through markup as its partial effect $\hat{m}_2(z)$ is indistinguishable from zero. In contrast, firms residing in eastern area have the markup effects significantly negative, decreasing labor share on average by a percentage point of 0.306 for every one unit it increases. Labor in inland region are also less affected by exports through capital intensity, as its non-neutral effect through capital $\hat{m}_1(z)$ in inland is uniformly more negative than that in eastern area, suggesting different labor-capital elasticity of substitution across firms in inland area and coastal line. Finally, the difference of the impact of technological progress through trade $\hat{m}_3(z)$ is subtle between the two regions, implying that the negative impact of capital-biased technology on labor share is independent with regions.

3.5.5 Before vs after 2002

China's membership of the World Trade Organization (WTO) in November, 2001 motives many studies to examine the effect of its entry on its financial market construction and economic growth (Allen et al. (2005); Drysdale et al. (2000); Blancher and Rumbaugh (2004)). In the context of Chinese labor share, however, the effect of its entry is void but non-trivial. The fact that Chinese labor are substitutable from capital and that capital-biased technologies are largely implemented during trade would exert downward pressure on the share of labor's income. Thus, an interesting and important question surface as to whether exports shrink labor's share of income in a higher magnitude through capital due to China's entry of WTO.

We seek to the answer by splitting the sample based on the year of 2002, and report the estimation results in table 3.6. Results show that the neutral effect of exports after 2002 slightly increases as the neoclassical trade theory suggests, and the non-neutral effect through markup and TFP are surprisingly unchanged.

As we conjectured, the most (and perhaps the only one) notable difference lies on the non-neutral effect through capital intensity $\hat{m}_1(z)$, which can be seen from the upper-right panel of Figure (3.7). Before China joined WTO, our results show that exports affect labor share negatively through capital intensity, since the effect of capital intensity is fairly stable and not changing dramatically across the level of export intensity. In the year after 2002, the effect of capital intensity through exports not only emerges with highly negative value, but also diverge away from its level before 2002; that is, the effects in two time period do not overlap each other within their 95% confidence intervals. Conditioning on 10% of export intensity, for instance, a one percentage increase in capital relative to labor after 2002 would decrease labor share through exports by 0.023% higher than that in the year before 2002. This difference rises up to 0.027%, 0.029%, and 0.071% as export intensity is conditional on 25%, 50% and 75%, respectively. Evidently, China's participation in international trade since 2002 has unintentionally shifted the share of national income owned by labor toward the owners of capital by making Chinese workers in manufacturing industry to be more substitutable in the presence of cheaper capital price.

3.6 Conclusion

In this paper, we bridge the literature of trade and labor share to propose a hypothesis that export is likely to alter labor's share of income in a nation through its neutral effect and non-neutral effects. The neutral effect depends on export itself and is well supported by the neoclassical trade theory. The non-neutral effects depend on capital intensity, the degree of monopoly power, and the capital-augmenting technology commonly discussed in labor share literature. We highlight the channels through which the aforementioned three factors may help to explain the movement of labor share through export, which is empirically tested using firm-level data in Chinese manufacturing industries during 1998-2007, given the nation's intensively engagement in trade and millions of labor participation.

To alleviate potential model mis-specification, we implement a semiparametric varying coefficient model with fixed effect with a spline-backfitted kernel estimator that improves estimation efficiency and reduces computational time. Our results are fairly consistent with our hypothesis. On one hand, exports have relatively high and positive neutral effect on Chinese labor share, assuming no role is played by its non-neutral effects. However, non-neutral effects are consistently negative, with its effect through TFP the most severe one, followed by capital intensity and by markup. Since the totally partial effect of exports that depends on both neutral and non-neutral effect, the whole sample result shows that the positive neutral effect predicted by trade theory is largely pull down by its non-neutral impact, resulting in a mean of -0.0108. We attribute its underlying reasons partially to China's membership of WTO since

2002, which greatly eases the elasticity of substitution between labor and capital that drags down the overall effect of export on labor share.

In sum, we found that it is the capital-augmented technology and capital-intensity that mostly erode labor's gain from the *growth pie* by making labor to be easily substitutable in the face of the surging capital. Most importantly, we do not find evidence showing that the negative non-neutral effects through TFP and capital intensity diminish as exports increase. This should be of an important concern for policy makers, as the labor's overall benefit in the current stage seems not be well protected. Given the absence of strong labor unions or other institutions that are in favor of labor's benefits, the labor share in Chinese manufacturing industries is not likely to be stopped shrinking. A careful set of policies are therefore needed at least in manufacturing industries that potentially prevent labor's share of income from being shifted away.

3.7 Appendixes

Appendix 1 Testing for the endogeneity of export intensity

Endogeneity is a problem in estimating the effect of export on labor share. In this paper we make the assumption that firms make decisions of using labor in reaction to exporting behavior. However, there are cases where firms are inclined to export due to cheaper labor cost. That is, a reverse causality may exist in our regression that causes endogeneity. For the purpose of the identification, we test the null hypothesis where export is endogenous. Given the fact that the semiparametric varying coefficient model is a general case of a linear model, we test the null from the following fixed-effect linear model due to its computational ease:

$$\ln(ls)_{it} = \alpha_1 export_{it} + \alpha_2 capital_{it} + \alpha_3 TFP_{it} + \alpha_4 markup_{it} + u_i + \gamma_t + e_{it} \quad (3.9)$$

One of the traditional ways to test for exogeneity is to use a control function (Amsler et al. (2016)). This is equivalent to two-stage least square estimation (2SLS). In principle, we first regress a reduced-form regression of the (potential) endogenous variable (i.e. export) on instrumental variables and other exogenous variables in the original structural model. After we get predicted residuals, say $\hat{v}$, we plug $\hat{v}$ back to the original structural equation. The significant (insignificant) coefficient on $\hat{v}$ indicates that export is endogenous (exogenous). In absence of valid external instrumental variables, we employ one-year and two-year lags of export as instruments. When using one-year lag as instrument, the coefficient of $\hat{v}$ is -0.079 with standard error of 0.185. When both one-year and two-year lags are employed, the coefficient of $\hat{v}$ becomes -0.225 with standard error of 0.044. The insignificance of both coefficients under two instrumental sets indicates that export is exogenous. When we use the same instruments and add interaction terms, the coefficients of $\hat{v}$ still remain insignificant, indicating the exogeneity of export, and the coefficients of variables are quantitatively similar as the ones in our benchmark model. So this paper will proceed under the exogeneity of export.

Appendix 2 Chinese Industrial Classification Codes (manufacturing)

CICC	Industry Name
13	Agriculture and food processing industry
14	Foodstuff manufacturing industry
15	Soft drink manufacturing industry
16	Tobacco manufacturing industry
17	Textile industry
18	Waving costume, shoes and cap manufacturing industry
19	Leather, fur and feather manufacturing industry
20	Wood working, and wood, bamboo, bush rope, palm, and straw manufacturing industry
21	Furniture manufacturing industry
22	Paper making and paper products industry
23	Print and copy of record vehicle industry
24	Stationary and sporting goods manufacturing industry
25	Oil processing, coking and nuclear manufacturing industry
26	Chemical material and chemical product manufacturing industry
27	Medicine manufacturing industry
28	Chemical fiber manufacturing industry
29	Rubber product industry
30	Plastics product industry
31	Nonmetallic mineral product industry
32	Ferrous metal refining and calendaring processing industry
33	Non-ferrous metal refining and calendaring processing industry
34	Metal product industry
35	Universal equipment manufacturing industry
36	Task equipment manufacturing industry
37	Transport and communication facilities manufacturing industry
39	Electric machine and fittings manufacturing industry
40	Communication apparatus ,computer and other electric installation manufacturing industry
41	Instrument and meter, stationery machine manufacturing industry
42	Handicraft and other manufacturing industry
43	Removal and processing of obsolete resource and material industry

Source: NBSC

Appendix 3 Provinces Classification Based on Geography

Eastern Region	Beijing	Tianjin	Hebei	Shanghai	Jiangsu
	Zhejiang	Shandong	Fujian	Guangdong	Hainan
	Liaoning				
Inland	Liaoning	Jilin	Heilongjiang	Shanxi	Anhui
	Jiangxi	Henan	Hubei	Hunan	Neimenggu
	Neimenggu	Guangxi	Chongqing	Sichuan	Guizhou
	Yunnan	Tibet	Shanxi	Gansu	Qinghai
	Ningxia	Xinjiang			

Appendix 4 Robustness Check: Whole Sample Results with Labor Share Growth Rate

	Whole sample (exp: 42.08%)			
	NP $\hat{m}'_0(z)$	$\hat{m}_1(z)$	NN($k, markup, TFP$) $\hat{m}_2(z)$ $\hat{m}_3(z)$	
<i>exint</i>				
0.1	-0.2099*** (0.024)	-0.0903*** (0.033)	-0.1839*** (0.063)	-0.7801*** (0.035)
0.25	0.0205 (0.024)	-0.0844*** (0.032)	-0.0703** (0.086)	-0.8004*** (0.061)
0.5	0.0713*** (0.025)	-0.0996** (0.039)	0.0131 (0.055)	-0.9341*** (0.119)
0.75	-0.0666** (0.026)	-0.0667*** (0.024)	-0.0345 (0.036)	-0.8716*** (0.061)
0.9	0.1051** (0.048)	-0.0634** (0.025)	-0.0509 (0.034)	-0.8468*** (0.039)
Mean of NP/NN	-0.094	-0.0545	-0.0628	-0.8571
<i>Total partial effect</i>				
VCM-FE	0.0220			
OLS-FE	-1.9945			
<i>N</i>	4,767			
Total obs:	42903			
R^2 :	0.1782			
h_{cvls}	0.1424			
J_{gcv} :	6			
Knots space	0.2533			

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

Figure 3.8: Whole Sample: Labor share growth rate

Bibliography

- Acemoglu, D., 2003. Labor-and capital-augmenting technical change. *Journal of the European Economic Association* 1 (1), 1–37.
- Allen, F., Qian, J., Qian, M., 2005. Law, finance, and economic growth in china. *Journal of financial economics* 77 (1), 57–116.
- Amiti, M., Freund, C., 2010. The anatomy of china’s export growth. In: *China’s growing role in world trade*. University of Chicago Press, pp. 35–56.
- Amsler, C., Prokhorov, A., Schmidt, P., 2016. Endogeneity in stochastic frontier models. *Journal of Econometrics* 190 (2), 280–288.
- Atkinson, A. B., 2009. Factor shares: the principal problem of political economy? *Oxford Review of Economic Policy* 25 (1), 3–16.
- Aw, B. Y., Roberts, M. J., Xu, D. Y., 2011. R&d investment, exporting, and productivity dynamics. *American Economic Review* 101 (4), 1312–44.
- Azmat, G., Manning, A., Reenen, J. V., 2012. Privatization and the decline of labour’s share: international evidence from network industries. *Economica* 79 (315), 470–492.
- Bai, C.-E., Qian, Z., 2010. The factor income distribution in china: 1978–2007. *China Economic Review* 21 (4), 650–670.
- Bentolila, S., Saint-Paul, G., 2003. Explaining movements in the labor share. *Contributions in Macroeconomics* 3 (1).
- Berkowitz, D., Ma, H., Nishioka, S., 2016. Capital-labor substitution, institutions and labor shares. Tech. rep.
- Berkowitz, D., Ma, H., Nishioka, S., 2017. Does capital-labor substitution or do institutions explain declining labor shares? University of Pittsburgh.

- Bernard, A. B., Eaton, J., Jensen, J. B., Kortum, S., 2003. Plants and productivity in international trade. *American economic review* 93 (4), 1268–1290.
- Bernard, A. B., Jensen, J. B., 1997. Exporters, skill upgrading, and the wage gap. *Journal of international Economics* 42 (1-2), 3–31.
- Bernard, A. B., Jensen, J. B., 1999. Exceptional exporter performance: cause, effect, or both? *Journal of international economics* 47 (1), 1–25.
- Blanchard, O. J., Nordhaus, W. D., Phelps, E. S., 1997. The medium run. *Brookings Papers on Economic Activity* 1997 (2), 89–158.
- Blancher, M. N. R., Rumbaugh, M. T., 2004. China: international trade and WTO accession. No. 4-36. International Monetary Fund.
- Böckerman, P., Maliranta, M., 2011. Globalization, creative destruction, and labour share change: evidence on the determinants and mechanisms from longitudinal plant-level data. *Oxford Economic Papers* 64 (2), 259–280.
- Boggio, L., Dall’Aglia, V., Magnani, M., 2010. On labour shares in recent decades: A survey. *Rivista Internazionale di Scienze Sociali*, 283–333.
- Craven, P., Wahba, G., 1978. Smoothing noisy data with spline functions. *Numerische mathematik* 31 (4), 377–403.
- Daudey, E., García-Peñalosa, C., 2007. The personal and the factor distributions of income in a cross-section of countries. *Journal of Development Studies* 43 (5), 812–829.
- De Loecker, J., Warzynski, F., 2012. Markups and firm-level export status. *American Economic Review* 102 (6), 2437–71.
- Dinopoulos, E., Segerstrom, P., 1999. A schumpeterian model of protection and relative wages. *American Economic Review* 89 (3), 450–472.
- Dorn, D., Katz, L. F., Patterson, C., Van Reenen, J., et al., 2017. Concentrating on the fall of the labor share. *American Economic Review* 107 (5), 180–85.
- Drysdale, P., Song, L., et al., 2000. China’s entry to the WTO: An overview. Routledge, Taylor & Francis Group.
- Elsby, M. W., Hobijn, B., Şahin, A., 2013. The decline of the us labor share. *Brookings Papers on Economic Activity* 2013 (2), 1–63.

- Eubank, R. L., 1999. Nonparametric regression and spline smoothing. CRC press.
- Girma, S., Gong, Y., Görg, H., Yu, Z., 2009. Can production subsidies explain china's export performance? evidence from firm-level data. *The Scandinavian Journal of Economics* 111 (4), 863–891.
- Guerriero, M., Sen, K., 2012. What determines the share of labour in national income? a cross-country analysis. working paper.
- Guo, K., N'Diaye, M. P., 2009. Is China's export-oriented growth sustainable? No. 9-172. International Monetary Fund.
- Guscina, A., 2006. Effects of globalization on labor's share in national income. International Monetary Fund.
- Harrison, A., 2005. Has globalization eroded labors share? some cross-country evidence. working paper.
- Henderson, D. J., Parmeter, C. F., Su, L., 2015. Nonparametric threshold regression: Estimation and inference. working paper.
- Hung, J. H., Hammett, P., 2016. Globalization and the labor share in the united states. *Eastern Economic Journal* 42 (2), 193–214.
- Jaumotte, M. F., Tytell, I., 2007. How has the globalization of labor affected the labor income share in advanced countries? No. 7-298. International Monetary Fund.
- Jiang, X., 2008. The determinants and trends of china's exportable structure. *China Economic Journal* 1 (2), 191–202.
- Kalecki, M., 1938. The determinants of distribution of the national income. *Econometrica: Journal of the Econometric Society*, 97–112.
- Karabarbounis, L., Neiman, B., 2013. The global decline of the labor share. *The Quarterly Journal of Economics* 129 (1), 61–103.
- Keller, W., 1998. Are international r&d spillovers trade-related?: Analyzing spillovers among randomly matched trade partners. *European Economic Review* 42 (8), 1469–1481.
- Keller, W., 2002. Trade and the transmission of technology. *Journal of Economic growth* 7 (1), 5–24.
- Lawrence, R. Z., 2015. Recent declines in labor's share in us income: a preliminary neoclassical account. Tech. rep., National Bureau of Economic Research.

- Levinsohn, J., Petrin, A., 2003. Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies* 70 (2), 317–341.
- Lu, D., 2010. Exceptional exporter performance? evidence from chinese manufacturing firms. manuscript, University of Chicago.
- Luo, C., Zhang, J., 2010. Declining labor share: is china's case different? *China & World Economy* 18 (6), 1–18.
- Ma, X., 2018. Labor market segmentation by industry sectors and wage gaps between migrants and local urban residents in urban china. *China Economic Review* 47, 96–115.
- Ma, Y., Tang, H., Zhang, Y., 2014. Factor intensity, product switching, and productivity: Evidence from chinese exporters. *Journal of International Economics* 92 (2), 349–362.
- Ohlin, B., 1952. *Interregional And International Trade*. Vol. 39. Harvard University Press.; Cambridge.
- Park, B. U., Mammen, E., Lee, Y. K., Lee, E. R., 2015. Varying coefficient regression models: a review and new developments. *International Statistical Review* 83 (1), 36–64.
- Perugini, C., Vecchi, M., Venturini, F., 2017. Globalisation and the decline of the labour share: a microeconomic perspective. *Economic Systems* 41 (4), 524–536.
- Petrin, A., Poi, B. P., Levinsohn, J., et al., 2004. Production function estimation in stata using inputs to control for unobservables. *Stata journal* 4, 113–123.
- Piketty, T., Yang, L., Zucman, G., 2017. Capital accumulation, private property and rising inequality in china, 1978-2015. Tech. rep., National Bureau of Economic Research.
- Piketty, T., Zucman, G., 2014. Capital is back: Wealth-income ratios in rich countries 1700–2010. *The Quarterly Journal of Economics* 129 (3), 1255–1310.
- Rodrik, D., 1998. Has globalization gone too far? *Challenge* 41 (2), 81–94.
- Stockhammer, E., 2013. Why have wage shares fallen? an analysis of the determinants of functional income distribution. In: *Wage-led growth*. Springer, pp. 40–70.
- Stolper, W. F., Samuelson, P. A., 1941. Protection and real wages. *The Review of Economic Studies* 9 (1), 58–73.
- Sun, Y., Carroll, R. J., Li, D., 2009. Semiparametric estimation of fixed-effects panel data varying coefficient models. In: *Nonparametric econometric methods*. Emerald Group Publishing Limited, pp. 101–129.

- Sun, Y., Malikov, E., 2018. Estimation and inference in functional-coefficient spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*.
- Tian, J., 2018. Debt and total factor productivity growth in chinese industrial firms: Evidence from nonparametric threshold regression. manuscript, West Virginia University.
- Tian, J., Wang, T., 2018. Recasting the trade impact on labor share: A fixed effect semiparametric estimation study. manuscript, West Virginia University.
- Wang, J., Yang, L., 2009. Efficient and fast spline-backfitted kernel smoothing of additive models. *Annals of the Institute of Statistical Mathematics* 61 (3), 663–690.
- Young, A. T., Zuleta, H., 2015. Do unions increase labor shares? evidence from us industry-level data. *Eastern Economic Journal*, 1–18.
- Zhang, X., Lu, Y., 2014. Effects of commodity trade structure variations on labor’s share of income in chinaan empirical study based on data of production industry sectors. *China Economist* 4, 009.
- Zhou, M., 2015. *Labors Share of Income: Another Key to Understand Chinas Income Inequality*. Springer.
- Zhu, L., Jeon, B. N., 2007. International r&d spillovers: Trade, fdi, and information technology as spillover channels. *Review of International Economics* 15 (5), 955–976.