Economic Integration and Firm Performance in Vietnam

Inaugural dissertation submitted by Huong Quynh Nguyen in fulfillment of the requirements for the degree of
Doctor rerum oeconomicarum
at the Faculty of Business, Economics and Social Sciences of the University of Bern.

Submitted by
Huong Quynh, Nguyen
from Hanoi, Vietnam

[2018]

The faculty accepted this work as dissertation on November 7th 2018 at the request of the two advisors Prof. Dr. Joseph Francois and Dr. Miriam Manchin, without wishing to take a position on the view presented therein.

Economic Integration and Firm Performance in Vietnam

Huong Quynh Nguyen

The World Trade Institute

Department of Economics, University of Bern

Inauguraldissertation zur Erlangung der Würde eines

DOCTOR RERUM OECONOMICARUM

der Wirtschafts und Sozialwissenschaftlichen Fakultät

der Universität Bern

Bern, October 2018

To my parents and my daughter

Preface

This thesis is an important milestone in my life as it records a long academic road in which I have met many great people who have accompanied me, whom I can learn from, and have helped me to grasp knowledge to pursue my dream to become a researcher in economics.

First and foremost, I would like to express my grateful thank to my principal supervisor Prof. Dr. Joseph Francois who has consistently encouraged and guided me from the very first step that I wrote this thesis. I have been asked interesting and unique research questions as well as introduced updated literature with novel research approaches by him. Especially, he gives me many opportunities to involve in communities of economic experts for my research presentation and discussion. Prof. Dr. Joseph Francois is not only my supervisor, but also my career mentor and my co-author in the fourth chapter of this thesis. He has spent many hours discussing with me whenever I have difficulties with my work, and gave me a constructive criticism of the manuscript. Hence, his guidance is extremely helpful for me to complete this thesis. Secondly, I am particularly indebted to my second supervisor Dr. Miriam Manchin who is very interested in my work and spent her valuable time to discuss and to give detailed comments on my thesis. I have learnt from both of my supervisors their endless passion of contribution to economic research, teaching and coaching for young generation, and cooperation in research for international economic development. I am also very much grateful to Dr. Robert Teh who was my mentor when I was doing my PhD intern at the Statistics and Economics Research Department at the World Trade Organization (WTO). I was impressed by and have learnt from his challenging questions and valuable discussion on literature and research methodologies. Dr.Roberta Piemartini at the WTO also spent her valuable hours to discuss with me about the thesis, and interesting topic(s) for the further studies. This extremely deepens my expertise in firm-data analysis and trade liberalization, especially localized trade impacts. I would like to very much thank my co-author Dr. Stephan Kyburz whom I have been working with for the fifth chapter of this thesis, and becomes my very best friend. We find our ways to work well together even often during lunch time at the Mensa of the University of Bern. Studying in Switzerland gives me an extraordinary opportunity to meet

and discuss with a large number of outstanding researchers. In particular, I would like to specially thank Prof.Dr.Aymo Brunetti, Prof.Dr.Harris Dellas, Prof.Dr.Christian Hilber, Dr.Thanh Le, Prof.Dr.Brian McCaig, Prof.Dr.Douglas Nelson, Prof.Dr.Thomas Cottier, Prof.Dr.Manfred Elsig, Prof.Dr.Ron Davies, Prof.Dr.Esteban Rossi-Hansberg and professors who taught me courses in Gerzensee Study Center. They gave helpful comments on my drafted thesis, my presentations, and discussed questions relevant to my research. Remarkably, I am grateful to Prof.Dr. Maximilian von Ehrlich, Prof.Dr.Blaise Melly, Dr.Eddy Bekkers, and Dr.Octavio Fernandez who gave valuable comments and suggestions on econometric techniques. Dr. Andreas Bachmann always answers my questions supportively either they are relevant to the PhD course works or to my PhD research. I also would like to thank Dr.Andreas Beerli, Mr.Pierre Sauve, Dr.Doris Oberdabernig, Dr.Konstantin Buchel, and Mr.Christophe Degain as well as other researchers for their very helpful and interesting discussions. Professional editing of the Editor in Chief, and comments from two anonymous referees of the Journal of Asian Economics are very helpful to improve chapter 2 which was published in 2017. Especially, the team of Academic English at the WTI and at the University of Bern were supportive in English editing and improving my presentation skills. Dr.Ioannis Kokkinis and Dr.Stephan Kyburz helped me many times to solve issues when I compiled LaTeX file for the thesis. My former colleagues in Foreign Trade University and friends from Vietnam also support me by searching for data and providing relevant reference materials. I would like to thank the invaluable support from Dr.Ha Doan, Dr.Hanh Pham, Dr. Khanh Pham, Mr.Hung Doan, Ms.Trang Nguyen, and officers from the General Statistics Office. I further thank all my colleagues at the World Trade Institute, the Economic Department at the University of Bern, and the World Trade Organization that have created my working hours at the office more enjoyable and unforgettable. Great classmates from the PhD courses as Andy, Andre, Micheal, Adela, Zhiyong, Marcus, Maria, Philip and Sarah also provided me ample inspiration when we shared nice memories in our PhD life at the Gerzensee study center. They all become my good friends.

Importantly, I would like to acknowledge the funding and the scholarship for my PhD study from the SECO/FTU cooperation project, the World Trade Institute research funding, and the World Trade Organization PhD support program.

Last but not least, my beloved family in Vietnam, especially my parents and my daughter, have lent me unconditional support so that I could study the PhD program, and also travel the world. Their support is because of their love, patience, understanding, sacrifice and generosity. Hence, I would like to express my deepest gratitude to them.

Huong Quynh Nguyen (Daisy Nguyen), Bern, October 2018

Contents

1 Introduction

1.1 Introduction

Since the Renovation (Doimoi) in 1986, Vietnam has transformed from the poorest country (per capital income was only USD200-USD300 per year in 1980s) to a middle-income country in 2011, and becomes a dynamic economy (GDP growth gains on average 6-7 % annually from 2000-current).[1] During the period from 2000-2010, Vietnamese manufacturing was restructured under dramatic business reforms and important supportive industry mixed policies, such as: zoning policy and foreign direct investment facilitation. In particular, complicated regulations in business ("red tapes") have been reduced, and domestic market has been open for highly competitive foreign investors and trade partners. Apart from that, cluster-oriented policies were implemented to attract investment into manufacturing, and to develop the industries in specific locations. The output value of Vietnamese manufacturing jumped nearly four times from 2000 to 2010.[2] Besides, foreign direct investment inflow as percent of GDP rose from 3.85% in 2000 to 6.90% in 2010 compared to the number of 2.15% and 2.96% respectively in countries from the East Asia and Pacific (World Bank's Statistics).

To make contribution to the large existing literature on firm performance analysis in developing economies, this thesis presents four novel studies answering research questions relevant to economic integration and firm performance in Vietnam. The term "economic integration" in this thesis has several implications. First, it refers to the efforts of manufacturing firms to catch up with leading performers in terms of total factor productivity within their industry and region (chapter 2). Second, it implies the inter-regional economic integration of provinces after the economic decentralization policies to provinces in the country (chapter 3). Third, it calls the attention to the performance of individual firm through their cooperation and competition within industrial district clusters, and with firms in the nearest upward and backward linked clusters (chapter 4). Forth, the

[1] The story of Vietnam's economic miracle (www.weforum.org).

[2] Output value of the Vietnamese manufacturing was 198,326.1 (Billion VND at 1994 price) in 2000, and rose to 722,222.2 (Billion VND at 1994 price) in 2010 according to the Statistics Book of Vietnam in 2005 and 2010 which can be downloaded from www.gso.org.vn

productivity growth upgrading of domestic firms through channels of labour mobility, technology transfer and learning effects from foreign invested partners within closer distance is also an important part of the inter-regional economic integration process (chapter 5). In brief, the economic integration is analysed in depth at regional level (key economic region), province level (chapter 2 and 3), district level (chapter 4), ward level (chapter 5) in conjunction with analysis for industry specific (chapter 2, and 4), provincial manufacturing characteristics (chapter 3) and firm characteristics (chapter 5) to draw a novel picture of the firm performance in the Vietnamese manufacturing in eleven years (from 2000 to 2010).

The thesis gives the first review and novel analysis of the economic integration of the manufacturing firms in the country, and draws policy implications by: (i) Investigating the impressive industrial development in the presence of firms in manufacturing in sixty-three provinces, and analysing the closing gap of productivity between the least and the best firms in light of the government's efforts in economic reforms (chapter 2); (iii) showing different impacts of key driving factors on provincial TFP for sixty-three provinces in Vietnam (chapter 3); (iv) Indicating industry specific evidence of negative congestion effects on firms performance to show that the Vietnamese manufacturing is in need of a better industrial policies to take advantages of industrial clustering, and clarifying the importance of zoning policy on the formation of industrial clusters in the country (chapter 4); (v) Mapping the evolution of foreign direct invested (FDI) firms at district level, and showing the significant results of distance decaying effects at ward-level that might be helpful for placed-base policies in FDI facilitation of the country (chapter 5). These studies might suggest initiatives for manufacturing development not only in Vietnam, but also in other developing countries that have similar economic conditions.

In addition, the thesis makes a significant contribution to data analysis and applied empirical studies on firm performance in terms of total factor productivity (TFP) estimation for manufacturing firms in Vietnam as a particular case study. Unique panel datasets were built by merging the micro-data of Vietnam with "traditional data" such as industrial-level trade data, provincial-level data, domestic input-output table, inter-country input-output table, and "non-traditional data", such as geographical data and map layers. With the modification and application of different econometric methods, the firm performance in terms of TFP is analysed in connection with business reforms, industrial clusters and FDI spillovers to draw crucial implications for business reforms, industry location-based development, foreign direct investment facilitation and small and medium enterprises (SMEs) enhancement in manufacturing industries of a transition economy such as Vietnam.

1.2 Summary of Firm-level Data

The micro-data sets are compiled using annual surveys of firm in Vietnam from 2000 to 2010. The survey has been conducted by the Vietnamese General Statistics Office (GSO) for legally registered firms in Vietnam that still run their business by the year when the firm is surveyed. The survey provides rich firm-level information, such as: revenue, profit, fixed assets, number of employees, firm ownership (foreign owned or domestic owned), location (63 provinces, 663 districts, 5198 wards), industry classification (VSIC 1993) at 2-digit (23 industries) and 4-digit (129 industries). Using each firm's unique ID key, a panel data was built for 11 years which includes about 287,650 observations. This data set is an unbalanced panel which records firm entering or exiting the industries, and some firms record missing for some years.

The following table lists selected studies using the firm-level datasets of the Vietnamese manufacturing, in different topics and time frame.

Table 1.1: List of relevant studies

Year	Author(s)	Topic	Data	Journal
2017	Huong Nguyen	Business reforms & TFP	2000-2010	Journal of Asian Economics
2017	Mai Vu., et al	TFP & wage premium	2000-2013	International Economics Journal
2016	Ha Doan., et al	TFP & misallocation	2000-2009	Asian Development Review
2016	Hanh Pham., et al	TFP spillovers	2000-2005	Journal of Development Studies
2015	Newman., et al	TFP spillovers	2009-2012	European Economics Review
2014	Ha and Kiyota	Trade Liberalization & TFP	2000-2009	The Japanese Economic Review

In this thesis, the firm-level data set was combined with inter-country input-output table (OECD) using the 2-digit classification ISIC.Rev3 which is equivalent to VISC 1993 in chapter 4. Chapter 5 investigates the TFP growth of firms at the 4-digit industry level which were converted to the 2-digit industry level in the domestic input-output table.

To develop different research questions relevant to proximity and car route travel distance, the thesis exploits firm's administrative location code at province, district and ward which can be merged with geographical information system data (GIS from gadm.org), as well as data of map layers (from www.openstreetmap.com).

1.3 Summary of Chapters

Chapter two: Business reforms and TFP in Vietnamese Manufacturing

Vietnam has implemented impressive business reforms, especially during two phases: Phase 1 from 2000 to 2005 when red tapes were cut to simplify business registration procedures; and phase 2 from 2006 to 2010 when one-stop-shops were established at each province to support for better business regulations. This paper provides the first evidence that technology gaps, in terms of total factor productivity difference between the least productive firms and the frontier firms, were narrowed across industries (and economic regions) in light of the reforms in Vietnam. Patterns of industry-level TFP growth rate were visualized for different key economic regions in each phase of the reforms. The results were drawn by exploring a novel micro-data set of manufacturing industries in Vietnam for the period between 2000 and 2010, and applying an advanced semi-parametric method proposed by Wooldridge (2009a) and Petrin and Levinsohn (2012) to measure firm-level total factor productivity (TFP).

Chapter three: Post-reform period in Vietnam: Mind the different impacts of TFP driving factors among provinces?

This chapter studies economic performance (TFP) of manufacturing at provincial level in light of economic decentralization in Vietnam after the adjusted law on Enterprises (since 2005). It is worth studying why some provinces perform better in the post-reform period because other provinces might learn from their success. It is also crucial to extract appropriate lessons for particular provinces due to their differences in geography, history and economic conditions.

The results of this chapter give a novel empirical evidence for key factors fostering provincial total factor productivity (TFP) in the context of decentralization in business regulations in the country from 2006 to 2010. The chapter highlights the different influences of these factors on TFP among groups of provinces by geography (North and South), by labour intensity, and by the presence of private firms. Particularly, impacts from provincial competitiveness on local TFP generate larger coefficients and more significant for provinces in the Southern region. Provincial competitiveness only shows positive impacts on provinces with a larger number of private firms (higher than median). Net effects of nearby leading productive provinces on provincial TFP are positive in the Northern provinces, but are negative in the Southern counterparts. These results may give some suggestions for regional development policies in Vietnam.

Chapter four: Industrial Clusters and Firm performance (joint with Joseph Francois)

Travel route plays an important role in inter-linking manufacturing firms (labour mobility, input trading, knowledge transfer), hence this paper presents novel findings for effects on firm performance induced from agglomeration effects of industrial clusters, and evidence that travel distance matters for cluster effects in Vietnam. Travel distance from a manufacturing firm to its intra-industry and cross-industry clusters as a proxy for travel cost in analysis of agglomeration effects is in focus of this paper. The travel distance is measured by the shortest car route using map data layers. Cluster orientation of zoning policies in Vietnam is also examined in this study.

Our results are two folds:(i) Evidence of agglomeration economies and congestion effects was found in manufacturing industries in Vietnam during 2005-2010.(ii) Results of the Logit estimation indicate that zoning policies in Vietnam is cluster-oriented since district with special economic zone(s) shows higher possibility of forming the clusters in most industries. The results suggest that policy makers in Vietnam may consider easing the congestion effects in the high-density clusters by establishing new industrial hubs, or supporting for emerging industrial districts with zoning policies.

Chapter five: Does proximity to foreign direct invested firm stimulate productivity growth of domestic firms? Firm-level evidence from Vietnam (joint with Stephan Kyburz).

This paper presents the first evidence for spatial components of foreign direct investment (FDI) spillover effects on total factor productivity (TFP) growth in Vietnam from 2005 to 2010. A unique micro-data set was compiled by geo-referencing domestic firms and their foreign counterparts using ward - the smallest administrative unit in Vietnam. In addition, to improve empirical studies on FDI spillovers, we apply a dynamic model of TFP growth using the first-difference two stage least square estimator. For the first time, spatial components of spillovers from foreign services suppliers in vertical linkages are investigated to reflect the importance of trade in services in manufacturing sectors.

This chapters give two key results. First, positive horizontal spillover effects are strongest, highly significant within radii of 2 to 10 km, and in a distinct decaying pattern within radii of 10 to 50 km. The significant effects are found for small and medium enterprises (SMEs) indicating SMEs also gain benefits from the spillovers within closer distance to foreign firms in the same industry. Second, vertical spillovers from the manufacturing sector are localized while vertical spillovers from the service sector are less geographically restricted. This study not only provides an advanced approach to investigate the spatial component of FDI spillovers, but may also suggest important implications for FDI facilitation policies in Vietnam and other developing countries.

2 Business reforms and total factor productivity in Vietnamese manufacturing

2.1 Introduction

Total Factor Productivity (TFP) has been regarded as a key driver of economic growth (Comin (2010)). TFP explains the growth of output that does not derive from the inputs used, but from the levels of technology efficiency (innovation), management and quality of inputs exploited in production (Syverson, 2011, and Van Beveren, 2012). A large number of scholars was interested in investigating TFP in light of reforms such as international trade liberalization,[1] and economic governance reforms [2] Much interest is paid to the catch-up of least productive firms to their frontiers by investigating TFP dispersion.[3] This is particularly important to developing economies where firms perform at low starting point of productivity. As noted by Brunetti et al. (1999), business regulations were among the biggest obstacles that hindered business activities of firms in South and Southeast Asia. Ease of doing business would encourage start-up and investment for advanced technologies, hence might enhance productivity catch-up of the least productive. Nevertheless, the TFP catch-up fashion in the context of business regulation reforms was not yet investigated for developing countries in the regions. This study, therefore, aims to give the first evidence of TFP gaps between least productivity and frontier firms across industries and key economic regions, and draw patterns of average industrial TFP growth for each region in two phases of business reforms in Vietnam from 2000 to 2010.

The period between 2000 and 2010 witnessed two impressive phases of business reforms in the country. Phase 1 - Red tapes cutting (2000-2005): the government simplified the administration procedures for start-up businesses. Phase 2 - One-stop shops setting (2006-2010): decentralized the control on business registration to local authorities of 63 provinces to support for further simplification in doing business. For example, UNIDO (2011a)

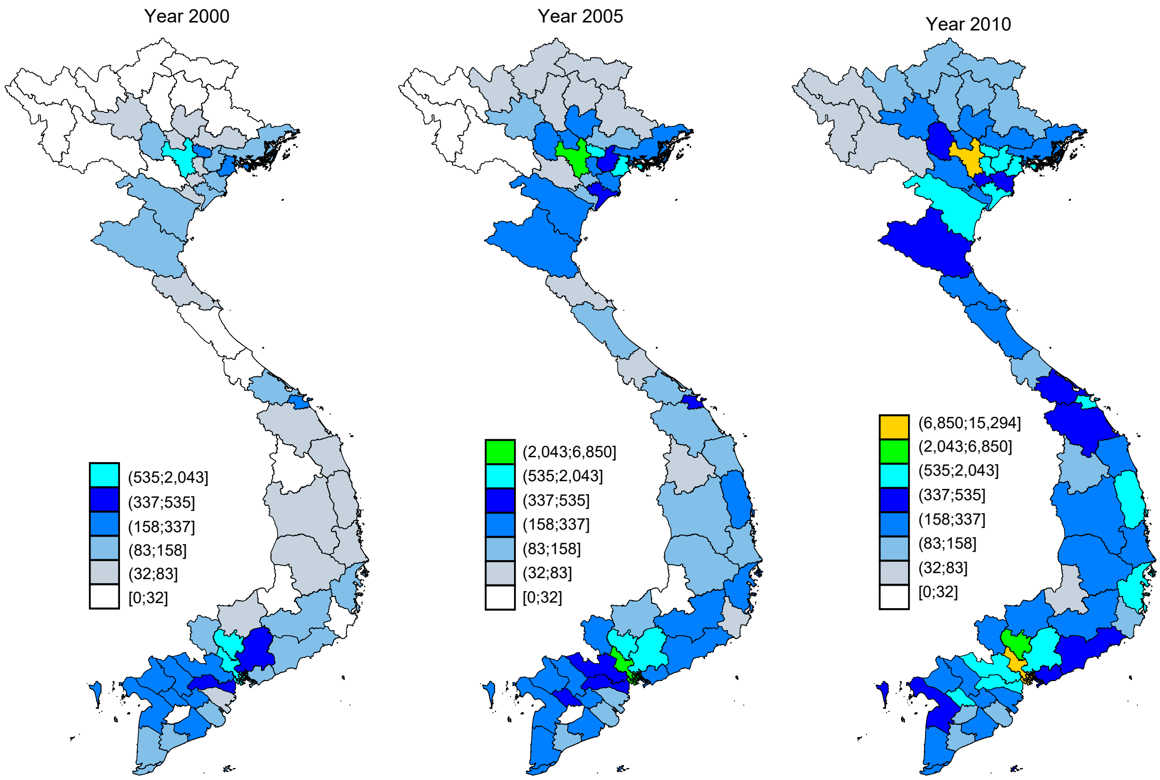
[1]See: Ha and Kiyota (2014), Topalova and Khandelwal (2010), Francois and Hoekman (2010)

[2]See: Acemoglu et al. (2005), Acemoglu and Robinson (2008), McCulloch and Malesky (2011), Alder et al. (2012), and Ghosh (2013)).

[3]See: Syverson (2011), Fox and Smeets (2011), Collard-Wexler (2011), and Hsieh and Klenow (2009).

reported that average time for formal business establishment in the periods between 1991-1999, 2000-2005, 2009-2011 respectively was 6-12 months, 50 days, and 5 days in Vietnam. Using our data, we could observe a boom in start-up business in Vietnam since 2000. Figure 2.1 shows an increase in number of firms and pattern of firms across provinces in Vietnam in 2010 compared to 2005 and 2000.

Figure 2.1: Number of firms by province, 2000, 2005 & 2010



Source: Firm-level data was drawn from the Vietnam Enterprise Survey (Vietnam General Statistic Office, 2000, 2005, and 2010). Administrative boundaries are based on Global Administrative Areas data (www.gadm.org). Several Vietnamese islands (e.g.: Hoang Sa and Truong Sa) are not displayed due to the limitation of the GADM administrative boundaries data.

This paper makes several contributions to the literature. First, a unique firm-level panel data set in 2-digit manufacturing industries is compiled from the Vietnamese enterprise survey for the period between 2000 and 2010. This study not only updates longer research period over recent literature but also considers two important phases of the reforms that were ignored in previous studies.[4] Second, the semi-parametric estimation introduced by Wooldridge (2009a) and Petrin and Levinsohn (2012) is modified for shaping a better production function by controlling for business cycles, and adding cluster standard errors for robust inference. Third, to improve existing literature, this research calculates the gaps in TFP between least productive and frontier firms not only across industries but also across key economic regions in Vietnam. Our key results are twofold.

[4]See the research time frame between 2000 and 2009 in Ha and Kiyota (2014), or from 2009 to 2012 in Newman et al. (2015), or between 2000 and 2012 but only for 6 manufacturing industries in Nguyen et al. (2016).

First, catching up is shown by much smaller TFP gaps between the least productive firms and the frontier firms in the second phase compared to the first phase of the reforms. Higher technology intensity industries observed larger TFP gaps in both phases. Second, interesting patterns of average TFP growth rate for each industry in different economic regions are shown for two phases of the reforms.

The rest of this paper is arranged as follows: Section 4.2 contains a literature review. Section 4.3 presents methodologies applied for firm-level TFP measurement, industrial TFP growth and TFP gaps. Section 4.4 describes the dataset. Section 2.5 discusses the empirical results. Section 4.7 draws conclusions and suggests further studies.

2.2 Literature Review

2.2.1 Production Function Estimation

Previous studies, which surveyed the TFP measurement methods, showed that semi-parametric method (Olley and Pakes, 1996) and its extensions (Levinsohn and Petrin, 2003, and Wooldridge, 2009a) could solve the endogeneity between inputs and unobserved productivity when estimating production function.[5] The robustness to measurement errors is also an advantage of the semi-parametric method (Van Biesebroeck, 2004).

To solve the bias issue of ordinary least square estimation for production function, Olley and Pakes (1996) first decomposed production function residuals into the firm's productivity and the random and zero-mean measurement errors, then used the inverse function of investment as the proxy for the unobserved productivity. Levinsohn and Petrin (2003) proposed to exploit intermediate inputs as the alternative proxy in case of lumpy investment. As claimed by Ackerberg et al. (2006), multicollinearity could happen when labour is correlated with the proxy, then the labour coefficient cannot be identified. To overcome this issue, Wooldridge (2009a) and later Petrin and Levinsohn (2012) suggested applying instrumental variables estimator using own lags of labour for its instruments. Newman et al. (2015) applied this method and added assumption of perfect competition market for each industry. However, empirical studies in TFP measurement by Petrin and Levinsohn (2012) and Newman et al. (2015) did not control for year effects, and ignored robust standard errors within clusters.

Studying imperfect competition markets, De Loecker and Warzynski (2012) exploited export information of Chilean micro-data to estimate Cobb-Douglas production function

[5] See: Van Biesebroeck (2004), Ackerberg et al. (2007), and Van Beveren (2012).

and translog production function. However, this method can only be applied for the micro-data that includes information on firm-level export activities. Working on panel of 3-digit Indian industry-level data, Maiti (2013) modified the estimation method proposed by Olley and Pakes (1996) and Levinsohn and Petrin (2003) by adding market imperfections in the context of international trade, and rearranging the production function. Given an ideal model for industry level TFP measurement, it is unclear how to get predicted labour coefficients from this method to measure firm-level TFP following Equation 5.23.

In short, the instrumental variables estimator proposed by Wooldridge (2009a) and Petrin and Levinsohn (2012) is most suitable to apply for the micro-data drawn from the Vietnamese enterprise survey. For the best application of the method to the context of the country, several modifications of the method are discussed in Section 2.3.2.

2.2.2 Measurement of Firm-level TFP and Dispersion of TFP

Thanks to an increasing availability of micro-data and a wide range of methods for production function estimation, various studies are applied to measure TFP at firm-level in the light of interesting economic topics. First, the common steps in TFP measurement are: (i) to apply the most appropriate method to estimate the production function parameters in a specific industry; and (ii) to use parameters estimated to measure TFP. Second, the TFP measured will be considered to be either directly the object for dispersion and tendency analysis, [6] or the dependent variable in estimating the impacts of spillovers effects,[7] or trade policy reforms [8]. Selected empirical studies are listed in Table 2.1.

Table 2.1: Summary of Literature on TFP

Year	Author(s)	Topic	Methodology	Country
2015	Newman et al	FDI spillovers	Wooldridge (2009a)	Vietnam
2014	Ha & Kyota	Trade liberalization	TFP index	Vietnam
2015	Arnold et al	Services liberalization	Olley and Pakes (1996)	India
2013	Deloecker	Learning by exporting	De Loecker (2013)	Slovenia
2012	Combes et al	Agglomeration effects	Levinsohn and Petrin (2003)	France
2012	Levinsohn & Petrin	Aggregate TFP growth	Wooldridge (2009a)	Chile
2009	Hsieh & Klenow	Misallocation	TFP index	China & India
2004	Syveson	Productivity dispersion	TFP index	United States

[6]See: Syverson (2011) for the survey of literature in determinants of TFP; Hsieh and Klenow (2009) for the case studies of India and China, Fox and Smeets (2011) for a case study of Denmark, and Syverson (2004) for the case study of the US.

[7]See Newman et al. (2015) for FDI spillovers effects in Vietnam.

[8]See Francois and Hoekman (2010) for the complete review of literature on trade in services, see Amiti and Konings (2007), and Topalova and Khandelwal (2010) for case studies of India and Indonesia respectively)

Regarding the studies on TFP dispersion, Syverson (2011) summarized selected research on patterns of TFP in light of competition (Syverson, 2004), sunk cost (Collard-Wexler, 2011), and input quality (Fox and Smeets, 2011), etc. TFP dispersion is explained differently in the literature. Fox and Smeets (2011) showed that for eight Danish industries, both human capital measures (schooling and gender) and wage bill equally influence the ratio of the 90th and 10th productivity percentiles. However, Syverson (2004) considered competition the key driver of the TFP gap. Industry de-licensing and size restriction policies were taken into account for the TFP difference analysis in India (Hsieh and Klenow, 2009). Hsieh and Klenow (2009) found no significant evidence of a link between TFP dispersion and de-licensing, labour market regulation, geographic measures and industry concentration for either China or India. Ha and Kiyota (2014) found an increase in the aggregate TFP growth from 2006 to 2007 thanks to trade liberalization when Vietnam joined the World Trade Organization. Nevertheless, there is little relevant evidence shown for TFP dispersion and industrial TFP growth in Vietnamese manufacturing in two phases of business reforms during 2000 to 2010.

2.3 Methodology

In this paper, measurement of firm-level TFP gaps between upper TFP firms and lower TFP firms includes three steps. (i) In the first step, production function is estimated for firms within each industry using instrumental variables as suggested by Wooldridge (2009a) and Petrin and Levinsohn (2012). (ii) In the second step, TFP are measured using the parameters estimated from the industry specific production function. (iii) From the results of measured TFP, TFP gaps are calculated following Hsieh and Klenow (2009).

2.3.1 Theoretical Framework of Firm's Behaviour

This paper assumes perfect competition in studied industries, which is similar to Newman et al. (2015) who estimate manufacturing firm-level TFP in Vietnam using method of Wooldridge (2009a). The firm's behavior basically followed the theoretical framework introduced by Olley and Pakes (1996): A profit-maximizing incumbent observes the productivity itself, and decides to continue to produce if its productivity exceeds a threshold of productivity $\bar{\omega}_t$; Otherwise, it leaves the market and is compensated with "sell-off values of its plant" (Olley and Pakes, 1996).

The Bellman equation is used to describe the behaviour of a profit maximizing firm (Olley and Pakes, 1996). The simplified Bellman equation is as follows:

$$V_t(\omega, k) = Max(\phi, Sup_{i_t \geq 0}\Pi_t(\omega, k) - c_t(i) + \beta E[V_{t+1}(\omega, k)|J_t] \tag{2.1}$$

Where:
$V_t(\omega, k)$: Current value of the firm which is the function of current productivity (ω), and capital stocks (k).[9] β is the discount factor. $E[V_{t+1}(\omega, k)|J_t]$ is the expected future value of the firm. J_t is the current information set. $\Pi_t(\omega, k)$ and $c_t(i)$ are respectively the profit function, and cost as the function of investment i.

Second, self-selection process of firms is:

$$\chi_t = \begin{cases} 1, & \text{if } \omega_t \geq \bar{\omega}_t. \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

Firm exits when $\chi_t = 0$, or stays when $\chi_t = 1$.[10]

2.3.2 Production Function Estimation

The assumptions in production function proposed by Wooldridge (2009a) is similar to assumptions stated in Levinsohn and Petrin (2003) and Olley and Pakes (1996): (i)The intermediate inputs have increasingly monotonicity in unobserved productivity; (ii) The unobserved productivity follows Markov rule; (iii) The unknown function of productivity is approximated by the third order polynomials function of capital stocks and intermediate inputs (Levinsohn and Petrin, 2003), or investment flows (Olley and Pakes, 1996);

The Algorithm of Production Function Estimation by Wooldridge (2009a) is summarized in Appendix 2.A.2.

This papery applies instrumental variables estimator (Wooldridge, 2009a). The production function is estimated for firms within each industry which follows the common practice in literature.[11] To improve the measurement of TFP in existing studies using method of Wooldridge (2009a) such as Newman et al. (2015) and Petrin and Levinsohn (2012), this paper adds some modifications: (i) year fixed effects are included in the estimation to control for business cycle, and annual shocks such as: the introduction of the Law on Enterprise in 2000, the amendment of the Law on Enterprise in 2005, the acces-

[9] The age of firm is excluded in the Bellman equation and in the production function. This simplification followed existing literature such as Levinsohn and Petrin (2003), Petrin and Levinsohn (2012), and Newman et al. (2015).

[10] $\bar{\omega}_t$ and i_t are decided by the firm under the Markov perfect Nash Equilibrium. These functions with t subscript means they depends on the market factor structure and factor prices at time t.

[11] See: De Loecker (2013), Newman et al. (2015), and Petrin and Levinsohn (2012).

sion of Vietnam to the World Trade Organization in 2007, and the global financial crisis in 2009, (ii) cluster-robust standard errors at firm-level are measured for robust inference as indicated in (Cameron and Miller, 2015).

Moreover, the business regulation reforms in Vietnam, which created better business investment environment, reduced entry barriers, and stimulated more competition, possibly encouraged the firm's investment in capital, and stimulated firm's expenditure on materials for more production. As noted in Section 3.1, the year 2006 is regarded as the stepping stone in ease of doing business in Vietnam. Hence production function is estimated in two sub-periods 2001-2005 and 2006-2010.[12] Significant results of Chow tests (Chow, 1960), which were conducted for each 2-digit industry, confirm year 2006 is the known structural break in the period between 2000 and 2010.

2.3.3 Measurement of TFP, TFP growth and TFP gap

Measurement of TFP

Equation 5.23 presents how ln(TFP) can be measured from parameters estimated from production function estimation in Section 2.3.2. This approach rules out the measurement errors of value-added.

$$\hat{\omega_{ijt}} = v\hat{a}_{ijt} - \hat{\beta}_l l_{ijt} - \hat{\beta}_k k_{ijt} \tag{2.3}$$

Where: $\hat{\omega_{ijt}}$: the log of estimated TFP of firm i in industry j at year t. $v\hat{a}_{ijt}$ is the value-added estimated from the production function. l_{ijt}, k_{ijt}: log values of labour and real accumulated capital stocks respectively. $\hat{\beta}_l$: labour parameter, $\hat{\beta}_k$: capital parameter estimated from the production function of industry j.

Average growth rate of TFP

The average growth of TFP by industry for each phase of the reforms is measured using weighted average TFP of each industry.

- Weighted average TFP of industry j in economic region k at year t

[12] As production function estimator proposed by Wooldridge (2009a) requires lagged variable at time $t-1$, we include year 2000 in the first sub-period, and include year 2005 in the second sub-period.

$$LnTFP_{jkt} = log(\sum \gamma_{ijkt} * TFP_{ijkt}) \tag{2.4}$$

Where share of real value-added of firm i in industry j in economic region k is:

$$\gamma_{ijkt} = \frac{VA_{ijkt}}{\sum VA_{ijkt}} \tag{2.5}$$

- Average TFP growth rate of industry j in economic region k for each phase (5 years):

$$g_{jk,2001:2005} = \left[\frac{lnTFP_{jk,2005}}{lnTFP_{jk,2001}}\right]^{1/4} - 1 \tag{2.6}$$

One can calculate average growth for the second phase using Equation 2.6.

TFP gap

Following Hsieh and Klenow (2009), the gap in TFP is measured as the difference in TFP between frontier firms (at 90th, or 75th percentile of TFP for each industry) and least productive firm (respectively at 10th, or 25th percentile of TFP):[13]

$TFPgap_{90/10} = ln(TFP)_{p90}$ - $ln(TFP)_{p10}$

$TFPgap_{75/25} = ln(TFP)_{p75}$ - $ln(TFP)_{p25}$

Where:

$ln(TFP)_{p90}$: ln(TFP) of the frontier (firm at 90th percentile)

$ln(TFP)_{p10}$: ln(TFP) of the least productive (firm at 10th percentile)

and alternatively:

$ln(TFP)_{p75}$: ln(TFP) of the frontier (firm at 75th percentile)

$ln(TFP)_{p25}$: ln(TFP) of the least productive (firm at 25th percentile)

[13] Hsieh and Klenow (2009) exclude firms with 1% highest and 1% lowest values of ln(TFP) to prevent the outliers. However, I still keep all the observations for the randomness.

As TFP is expressed in log values, the TFP gap indicates interesting implication. As noted by Syverson (2011), the gap can be inverted to the ratio between the output created by the TFP frontier firm and the output generated by the least productive firm, assumed that both using the same inputs: $Ratio = e^{TFPgap}$. In this study output is replaced by value-added as we use value-added as independent variable to estimate production function.

To fit the calculation of TFP gaps for the case study of Vietnam, the TFP gaps were calculated not only for each 2-digit industries but also across key economic regions and other regions in the country. Key economic regions were assigned by the government since 1997 to take advantages of the local region's natural resources and comparative advantages as well as to support for satellite provinces. Four key economic regions in Vietnam are: (i) The Northern key economic region includes Ha Noi (capital), Hai Phong, Vinh Phuc, Bac Ninh, Hung Yen, Quang Ninh, and Hai Duong. (ii) The Central key economic region consists of Da Nang, Thua Thien Hue, Quang Nam, Quang Ngai, and Binh Dinh. (iii) The Mekong River Delta economic region covers the area of Can Tho, An Giang, Kien Giang, and Ca Mau. (iv) Provinces in the Southern economic region are Ho Chi Minh, Dong Nai, Ba Ria-Vung Tau, Binh Duong, Binh Phuoc, Tay Ninh, Long An, and Tien Giang.

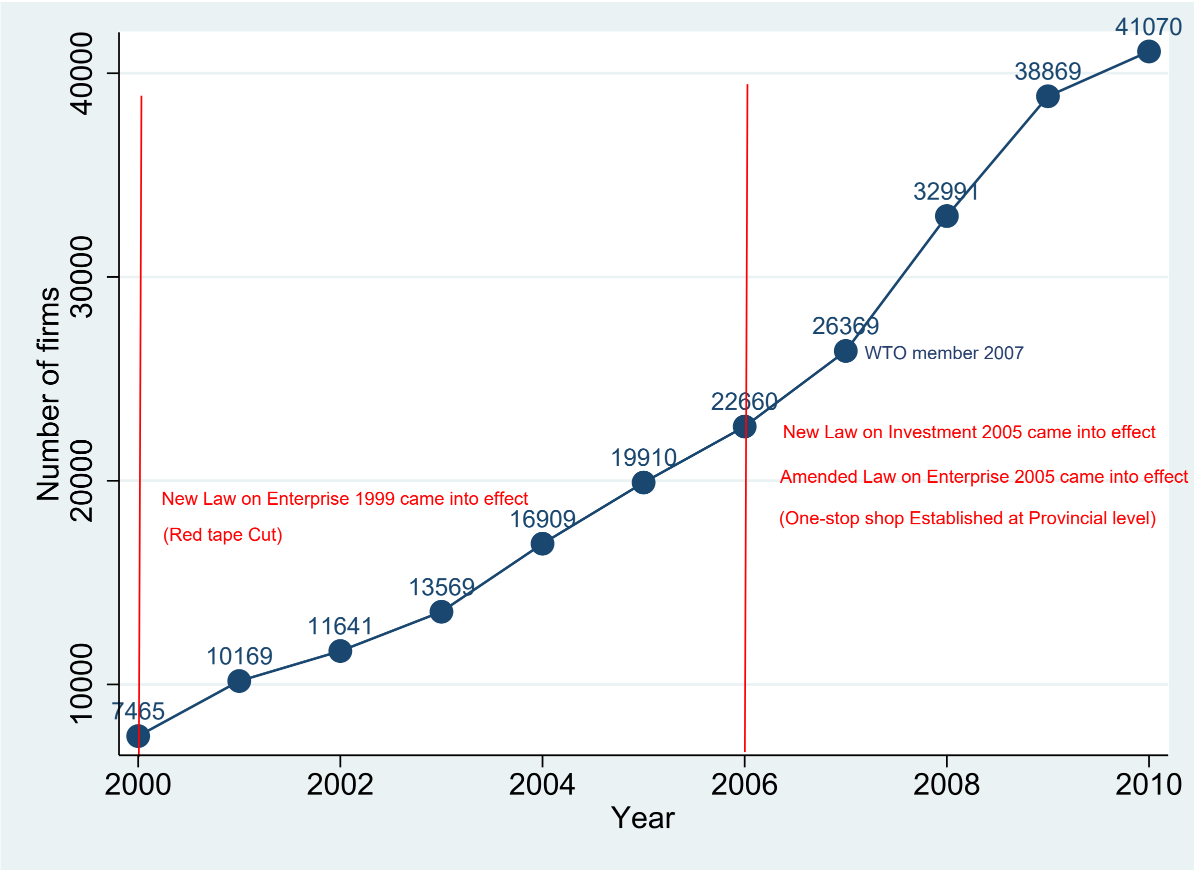
2.4 Data Description

This paper uses the rich manufacturing firm-level data drawn from the Vietnam Enterprise Survey for an eleven-year period (2000–2010). The survey has been conducted annually by the Vietnam General Statistics Office (GSO).[14] The data set contains annual information of legally registered enterprises (including business establishments) which were still doing business until December 31^{st} of the year surveyed. The firm-level information includes: establishment year, revenue, profit, expenditure, wage bills, number of employees, firm types, net fixed assets, etc. Each firm is identified by a unique id key which are then compiled into an unbalanced panel. The identification for firms in the dataset using the code assigned by the GSO for each firm through the years. Number of firms by years is shown in the Figure 2.2.

Table 2.2 depicts how key variables are constructed and presents the statistics summary. The measurement of value-added uses firm-level information of total profit, total wage, and depreciation. The depreciation ratio is assumed to be 10%. Different deflators

[14] The data has been aggregated and published annually in the Vietnam Statistical Yearbook. See more details in: www.gso.gov.vn.

Figure 2.2: Number of firms, 2000-2010



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2000-2010, 2-digit VSIC 1993, 15-37.

are used to convert the nominal values in the current price to the base year price which is year 2000. Specifically, the producer price index (PPI) of each industry is used to deflate output and value-added. [15] Capital stocks are converted to the price of base year 2000 by the gross fixed capital formation deflator. [16] Annual GDP deflators are used for nominal values of materials and services.[17]

The industries in the dataset are classified using the 2-digit Vietnam Standard Industrial Classification 1993 (hereafter named VSIC 1993) provided by the Vietnamese General Statistics Office.[18] To overcome the constraint in the number of observations in some industries, I merge related industries with each other (See Appendix 3.A.1 for

[15]Source: I calculated the Index with base year 2000 by using the annual Producer Price Index (PPI) by Industry provided by the General Statistic Offices of Vietnam at www.gso.gov.vn

[16]The calculation of deflators use the annual nominal gross fixed capital formation values of Vietnam from the World Bank, www.worldbank.org

[17]Source: The World Economic Outlook, www.imf.org

[18]VSIC 1993 is in line with the International Standard Industrial Classification (ISIC Rev.3) which was introduced by the United Nations. http://unstats.un.org/unsd/cr/ctryreg/ctrydetail.asp?id=1448

Table 2.2: Firm-level Data Description (Mean of log values and number of observations), 2000-2010

Variable	Measurement	2000-2005		2006-2010		2000-2010	
Labour	Number of labour	3.63	78225	3.13	161659	3.29	239884
Value-added	Profit + Wage + Depreciation	20.02	71896	19.83	149885	21.01	228271
Capital stock	Net book value of fixed asset	20.59	71830	20.35	152797	20.42	224627
Material	Total cost - Wage - Investment	21.52	64446	21.42	132902	21.45	197348

*Source:*Data drawn from the Vietnam Enterprise Survey (2000-2010). Industries listed in Appendix 3.A.1. *Notes:* The depreciation ratio is 10%. Value-added, capital stocks, materials are deflated values. Variables are in log values.

more details). Industries with the high concentration (i.e.: consider the Herfindahl Index using labour share), such as: 16 (Tobacco), 23 (Oil and Refinery Oil products), and 37 (Recycling), are excluded (a similar practice can be found in Newman et al., 2015).

2.5 Empirical Results

2.5.1 TFP dispersion

Figure 2.3 & 2.4 provide details about the average TFP gaps between the 75th and 25th percentiles, and 90th and 10th percentiles across industries in four key economic regions (Northern, Central, Mekong River Delta, and Southern) and other regions in two phases of the reforms. 2-digit industries were classified into 2 technology intensity groups: low, and medium and high technology intensity. The classification of technology intensity is heavily based on definition of the Organization for Economic Cooperation and Development (OECD).[19]

In general, Figure 2.3 & 2.4 indicate that TFP disparity in phase 1 was much larger than in phase 2. Apart from that, on average higher TFP gaps existed in higher technology intensity industries in comparison to lower technology intensity industries.

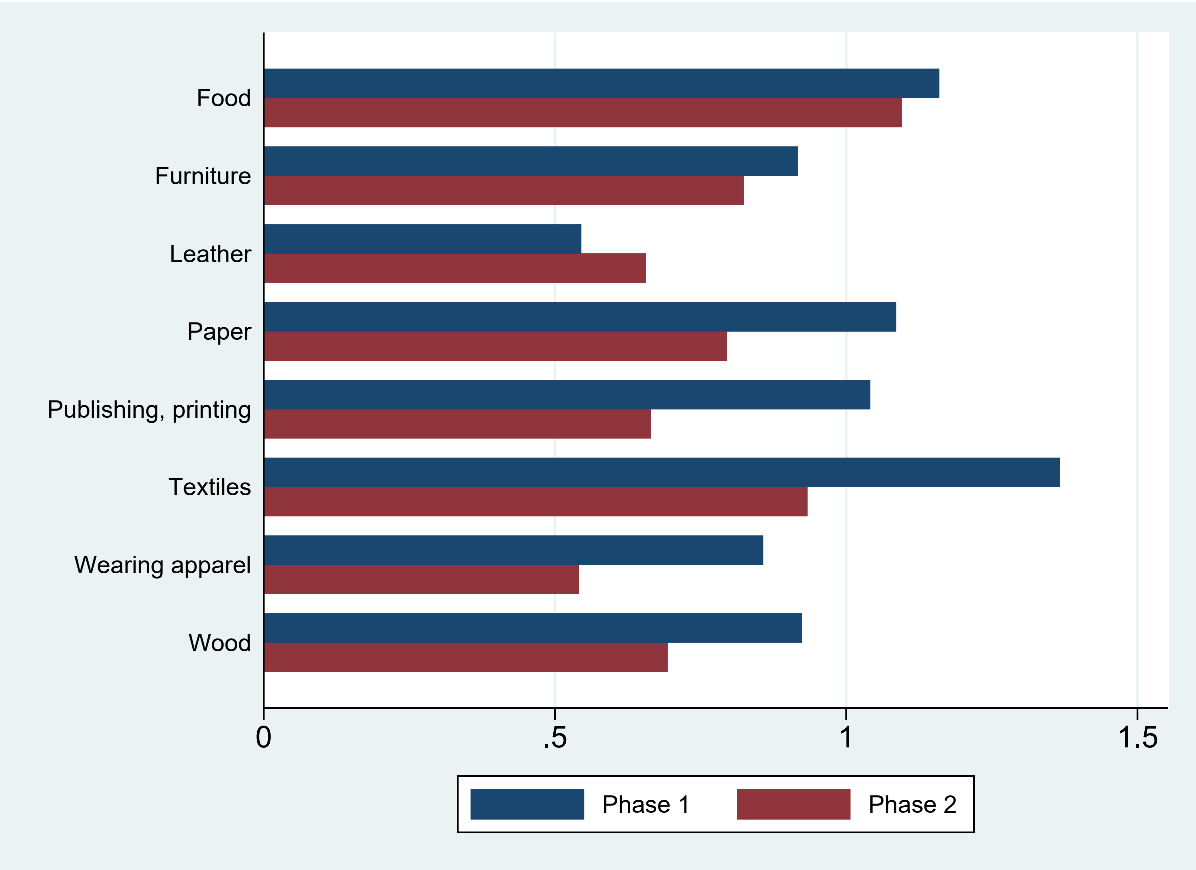
To be more specific, taking the exponential of 0.74, which is the average TFP$gap_{75/25}$ of the manufacture of wood products (industry coded 20) in the Northern key economic region: in the first phase, one could obtain the value-added ratio which is 2.09:1 for 75th:25th percentile firms; In the second phase, the ratio drops to 1.33:1 for the same industry in the same region. In addition, the manufacture of chemical products (industry code 24), which is a high technology intensity industry in the Northern region, reported the value-added ratio at extremely high value 13.5:1, and 2.41:1 respectively in the first and the second phase.

[19]See more details in: https://www.oecd.org/sti/ind/48350231.pdf

In short, the least productive firms catch up closer to frontier TFP firms in the second phase. These findings of TFP catch-up are shown in Figure 2.3 and Figure 2.4 for all regions. More competition and less entry barriers thanks to the reforms might support the catch-up of least productive. As technology imitation and knowledge transfer are more difficult to happen in higher technology intensity industries, given that the ease of doing business was implemented, larger TFP gap is still observed in the group.

Hsieh and Klenow (2009) report that the output ratio of 75th:25th percentiles is equal to 5.0:1 in India (1995), 3.6:1 in China (2005), and 3.2:1 in the US (1997) for manufacturing in general. However, it is hard to compare the results from literature with the results of Vietnam in this paper due to differences in research time frame and methods.

Figure 2.3: TFP Gap in Low Technology Industries, Phase 1 & Phase 2



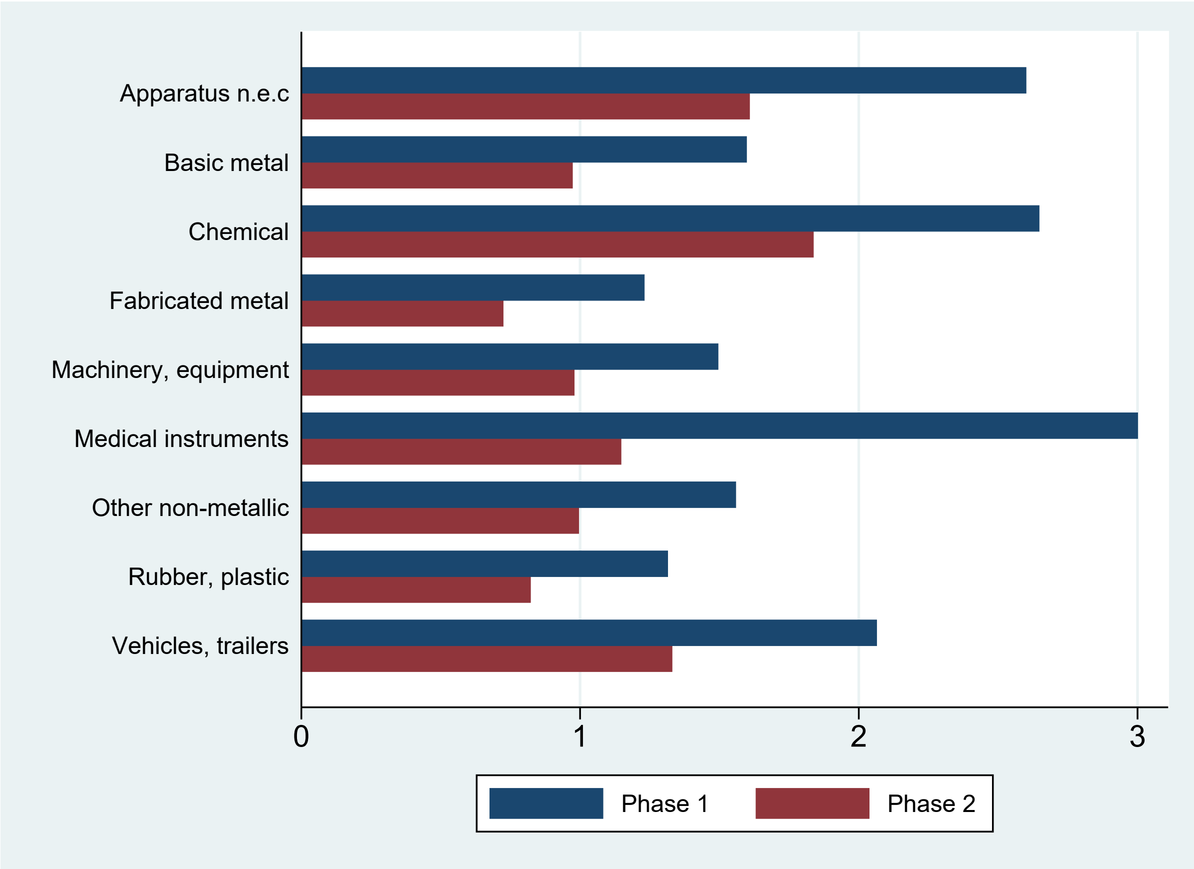
Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2000-2010, 2-digit VSIC 1993. List of industries is in Appendix 3.A.1.
Notes: Key Economic Region 1: Northern, region 2: Central, region 3: Southern, region 4: Mekong Delta, region 5: others.

2.5.2 TFP growth

Interesting patterns of TFP growth by industries across economic regions in phase 1 and phase 2 of the reforms are visualized in Figure 2.5 and Figure 2.6 accordingly. Interestingly, more industries with positive and faster growth rate are observed in the first phase in other areas rather than four economic regions (See Figure 2.5). Slowdown of TFP growth is shown for several industries in negative TFP growth rates in the second phase, especially in other regions (See Figure 2.6). The Southern key economic region, which is the biggest economic hub of Vietnam, performed at more stable TFP growth rate during two phases of the reforms. The youngest key economic region - Mekong Delta - and other areas were in deeper slowdown of TFP in the second phase compared to the Northern, the Southern and the Central regions.

Some examples are sketched out as follows. In phase 1, TFP grew fastest in industry 33 (medical, precision and optical instruments) in other areas which is not economic

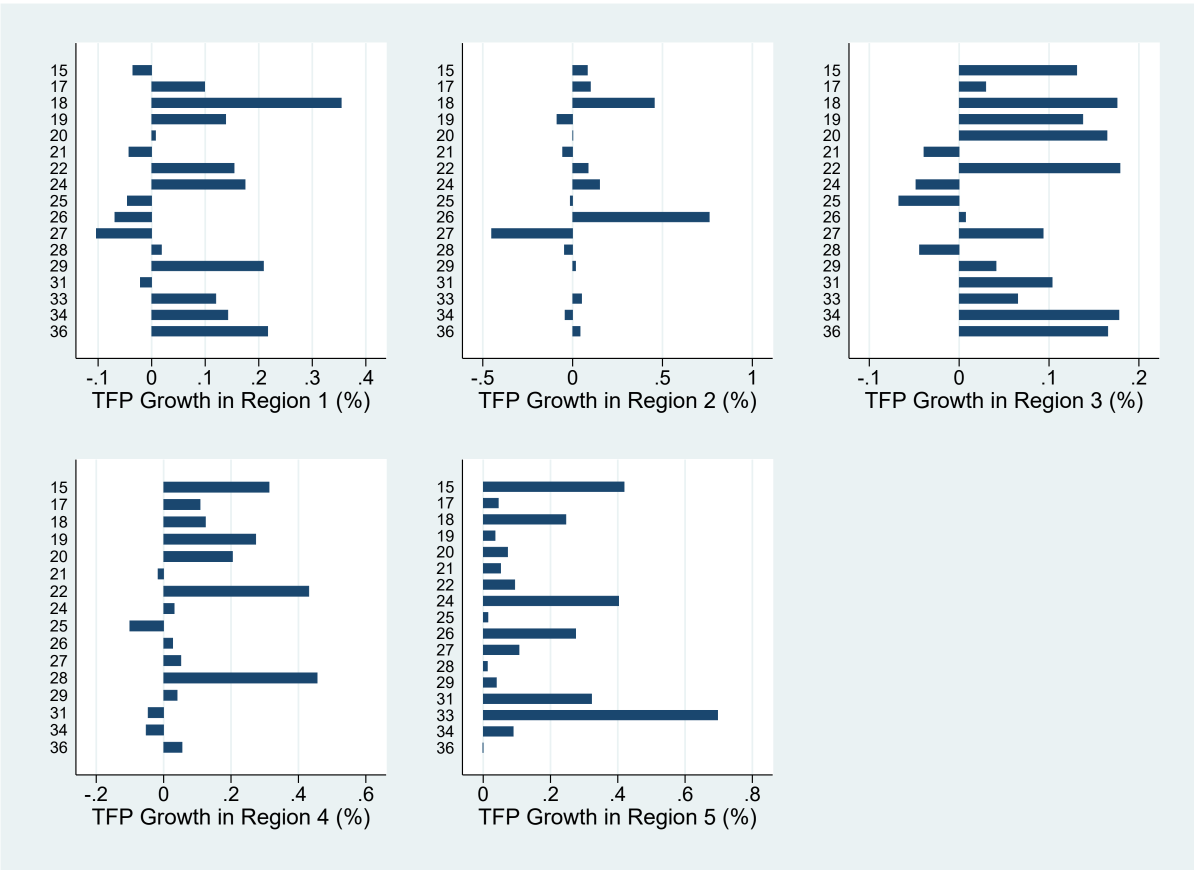
Figure 2.4: TFP Gap in Medium & High Technology Industries, Phase 1 & Phase 2



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2000-2010, 2-digit VSIC 1993. List of industries is in Appendix 3.A.1.
Notes: Key Economic Region 1: Northern, region 2: Central, region 3: Southern, region 4: Mekong Delta, region 5: others.

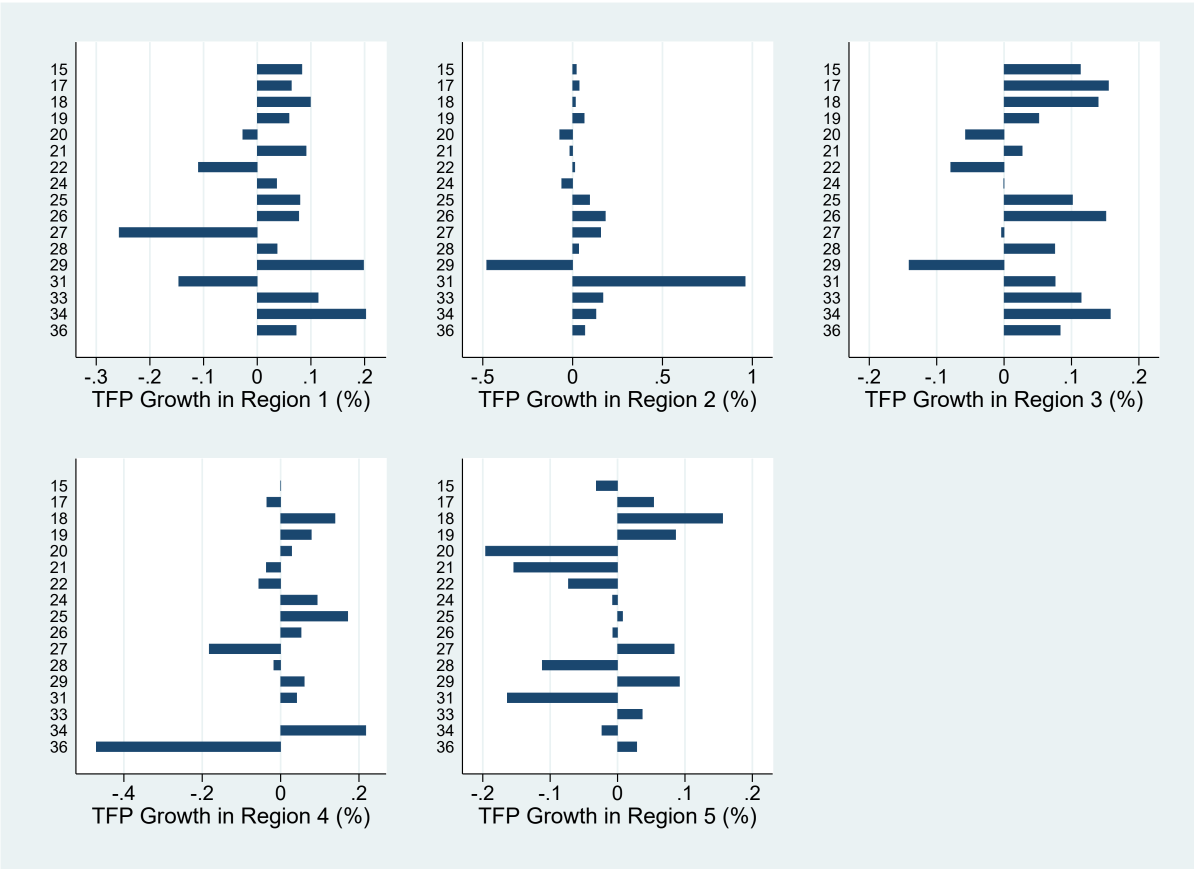
regions. The quick growth might be due to the lower starting point of TFP level of this industry in the region. Positive and high growth rate was shown in industry 18 (wearing apparel) for all regions. However, negative growth rate was observed in industry such as 27 (basic metal) in Northern and Central economic regions, or in industry 25 (rubber and plastic products) for four key economic regions. In phase 2, negative growth was still existing in industry 27 (basic metal) in the Northern and Mekong River Delta regions. The Central region gained highest growth rate in the country for TFP in industry 31 (electrical machinery and apparatus n.e.c, Television and Communication Equipment).

Figure 2.5: Average TFP Growth Rate, Phase 1



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2000-2010, 2- digit VSIC 1993. List of industries is in Appendix 3.A.1.
Notes: Key Economic Region 1: Northern, region 2: Central, region 3: Southern, region 4: Mekong Delta, region 5: others.

Figure 2.6: Average TFP Growth Rate, Phase 2



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2000-2010, 2-digit VSIC 1993. List of industries is in Appendix 3.A.1.
Notes: Key Economic Region 1: Northern, region 2: Central, region 3: Southern, region 4: Mekong Delta, region 5: others.

2.6 Conclusion

The early phase (2000-2005) of the economic reforms in Vietnam saw the impressive implementation of "cutting the red tape". This policy has reduced the administrative barriers for enterprises to enter manufacturing industries. In the second phase (2006-2010), "cutting the red tape" continues to reduce documents required in business regulations. Furthermore, "one-stop shops" were established to decentralize the central authorities control on business regulations to provincial authorities. The reforms have provided more convenient and efficient public administrative services to enterprises as well as encourages better interaction between local authorities and entrepreneurs. Interesting findings of this study reveal that the TFP gap was narrowed much more in the second phase of the reforms across industries and economic regions. That indicates an example of the TFP catch-up of the least productive to the frontier after business reform in a developing country such as Vietnam. Slower catch-up in TFP was still seen more in high-technology intensity industries than in low-technology intensity industries in phase 2. In addition, interesting patterns of average TFP growth rate for each industry in different key economic regions were visualized in two phases of the reforms. Given that faster TFP catch-up is found, TFP slowdown is shown for several industries in different key economic regions in Vietnam after the first phase of the reforms. With more availability of micro-data, a comparative study on TFP catch-up of the least productive firms to the frontier firms across industries in developing countries in Asia is also a promising research.

2.A Appendix

2.A.1 List of selected industries

Industry	Classification
Food Products	15
Textiles	17
Wearing Apparel; Dressing & Dying of Fur	18
Tanning & Dressing of Leather	19
Products of Wood	20
Paper products	21
Publishing, Printing & Reproduction of Recorded Media	22
Chemical Products	24
Rubber & Plastic Products	25
Other Non-metallic Products	26
Basic Metals	27
Fabricated Metal Products	28
Machinery & Equipment, Office, Accounting & Computing Machinery	29 (29+30)
Electrical Machinery & Apparatus n.e.c; Television & Communication Equipment	31 (31+32)
Medical, Precision & Optical Instruments	33
Motor Vehicles, Trailers & Semi-trailers & Other Transport Equipment	34 (34+35)
Furniture; Manufacturing n.e.c	36

Several industries are merged: industry classified as 29 includes VSIC-1993 industry 29 and 30; industry classified as 31 includes VSIC-1993 industry 31 and 32; industry classified as 33 includes VSIC-1993 industry 33 and 34. Industry classified as 15 (the Manufacture of Food) includes VSIC-1993 4-digit 1511 1512 1513 1514 1520.

2.A.2 Algorithm of Production Function Estimation

Equation 2.7 presents a Cobb-Douglas production function for an individual firm i:[20]

$$va_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it} \tag{2.7}$$

Where: va_{it} is the log of value-added, l_{it} is the log of number of labours, k_{it} is the log of capital stocks. Similar to Olley and Pakes (1996) and Levinsohn and Petrin (2003), Wooldridge (2009a) separates the error terms into ω_{it} which is the unobserved productivity of firm, and ϵ_{it} which is the measurement errors of the value-added and the unpredictable shocks. ϵ_{it} is orthogonal on current and past inputs:

[20] I use the notation va_{it} which is the log of value-added while Wooldridge (2009a) states the dependent variable is the log of output in general. This is in line with the value-added approach of Levinsohn and Petrin (2003).

$$E[\epsilon_{it}|k_{it}, l_{it}, m_{it}, k_{i,t-1}, l_{i,t-1}, m_{i,t-1}, k_{i1}, l_{i1}, m_{i1}] = 0; t = 1, 2...T \quad (2.8)$$

The assumption in equation 2.8 can be strengthen by adding the serial independence of ϵ_{it}, i.e: ϵ_{it} is random, is uncorrelated with its past values, and has zero mean.

Demand of intermediate inputs is: $m_{it} = m(\omega_{it}, k_{it})$. When m_{it} strictly increases in ω_{it}, it allows for the inversion of $m(.)$, yields: $\omega_{it} = f(m_{it}, k_{it})$. [21]

Because unobserved productivity ω_{it} follows the Markov rule:

$$\omega_{it} = E[\omega_{it}|\omega_{i,t-1}] + a_{it} = g[f(m_{i,t-1}, k_{i,t-1})] + a_{it} \quad (2.9)$$

Orthogonal condition of a_{it} is given by: [22]

$$E[a_{it}|k_{it}, k_{i,t-1}, l_{i,t-1}, m_{i,t-1}] = 0 \quad (2.10)$$

The shocks a_{it} is uncorrelated with k_{it} because capital stock is accumulated from k_{t-1}. The in-correlation of a_{it} with other variables in lagged values are obviously owing to the timing of choices were made in the past.

In this light, Wooldridge (2009a) sets up two simultaneous equations which jointly estimate the coefficients of capital stocks and labour inputs:

$$va_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + f(k_{it}, m_{it}) + \epsilon_{it} \quad (2.11)$$

$$va_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + g[f(k_{i,t-1}, m_{i,t-1})] + a_{it} + \epsilon_{it} \quad (2.12)$$

Orthogonal condition for equation 2.11 is given in equation 2.8. Orthogonal condition for equation 2.12 is:

---

[21] This is the assumption documented by Levinsohn and Petrin (2003). When the method of Olley and Pakes (1996) is applied, the investment is used instead of the intermediate inputs.

[22] a_{it} is assumed to be orthogonal to k_{it} and all past values of (k_{it}, l_{it}, m_{it}).

$$E[(\epsilon_{it} + a_{it})|k_{it}, l_{it}, m_{it}, k_{i,t-1}, l_{i,t-1}, m_{i,t-1}, k_{i1}, l_{i1}, m_{i1}] = 0 \qquad (2.13)$$

Wooldridge (2009a) notes that equation 2.12 could be estimated using instrument variable (IV) when $f[g(.)]$ is completely unspecified. The author also states that lags would be added to test the overidentification restrictions but at the cost of losing observations. Following Wooldridge (2009a), Petrin and Levinsohn (2012) instrumented l_{it} with its own lag; and other exogenous variables, which are shown in the orthogonal condition (Equation 2.13), act as their own instruments.

2.B Acknowledgement

I would like to thank Joseph Francois, Andreas Bachmann, Eddy Bekkers, Ha Doan, Hanh Pham, Hung Doan, Thanh Le, Stephan Kyburz, Octavio Fernandez, Miriam Manchin, Robert Teh, participants in the World Trade Institute PhD Colloquium (April 2016), participants in the Rimini Conference in the Waterloo University (September 2016), officers from the Vietnam General Statistics Office (GSO), and other researchers for their interesting discussions and suggestions. I am also grateful to helpful comments of the Editor in chief and two anonymous referees from the Journal of Asian Economics. This work was supported by a funding provided by the joint project of the Swiss State Secretariat for Economic Affairs (SECO) and the World Trade Institute (WTI) and the Foreign Trade University (FTU); the World Trade Organization (WTO) PhD support program; the WTI research funding; and the FTU research grant. However, the paper does not necessarily reflect views and ideas from these acknowledged organizations and people.

A peer reviewed version of this chapter has been published in the Journal of Asian Economics (51) 2017 under the title “Business reforms and total factor productivity in Vietnamese Manufacturing industry".

A previous version of this chapter is the first part in SECO/WTI working paper (No.9/2016) titled “Ease of doing business reforms in Vietnam: implication for total factor productivity in manufacturing industries".

3 Post-reform period in Vietnam: Mind the different impacts of TFP driving factors among provinces?

Huong Quynh Nguyen

3.1 Introduction

The pattern of regional economic development is crucial to a country's overall growth and development strategy, and for this reason it attracts a great deal of attention in the different strands of literature focused on explaining factors influencing growth and innovation, including both development economics [1] and urban economics [2]. Particularly, studies of "regional growth and regional decline" between the North and the South regions, coastal and non-coastal areas of a specific country are well reviewed by Breinlich et al. (2014).[3] Importantly, by considering iceberg transport cost, Desmet and Rossi-Hansberg (2010) introduce a novel theoretical model to analyse the spillovers across cities with multiple industries instead of "within city" analysis. While causal effects of agglomeration, local institutional variation, and interdependence among cities on regional economic performance are widely discussed theoretically and empirically in the literature, little evidence of these effects on local innovation were indicated for Vietnam in the post period of business regulation decentralization since 2005.

Because of differences in geography, history, economic conditions, and governance settings, regional disparities of economic performance in one country can last for decades and the analysis of it may give important initiatives for policy makers. This paper contributes to the literature by adding novel empirical evidence of key TFP driving factors for 63 provinces in Vietnam in five years of the post business reform period (from 2006-2010).

[1] See: Acemoglu et al. (2005), Acemoglu and Robinson (2008), and Dixit (2015).

[2] See: Duranton and Puga (2004) discuss theoretical models for agglomeration effects, Combes et al. (2012) use firm level data to explain why big cities in France perform better.

[3] Developed countries that were in the review are Italy, England, France and Japan. Developing countries that were referred to in the study are China and India.

According to Cung (2008), there was higher level of openness to business reforms and less "state-linked" of the local authorities in the South than in the North in Vietnam during 2000-2003. Since 2006 central government has reduced their control in sub-regional level after the amendment of Law on enterprises. However, analysis of provincial productivity in the post-reforms period in Vietnam from 2006 to 2010 is still neglected. This research thus highlights empirical evidence on how the key factors influence local TFP differently in the North and the South of Vietnam in the post-reforms period. More particularly, it explores further if more presence of private firms in the province could enhance the performance of local TFP. It also presents results for provinces more specialized in high labour intensive industries because these industries are important in the development path of Vietnam (during 2000-2010) when labour intensity is one of comparative advantages of manufacturing in Vietnam (see: Dinh (2013) and Anh et al. (2014)). We define provinces specializing in high labour intensive industry for provinces that show labour intensity higher than the median value of the country. Additionally, apart from receiving positive spillover effects from more productive performers (Desmet and Rossi-Hansberg, 2010), in practice provinces may face with competition with these provinces. [4]Thus, this study takes a further step to investigate the transfer of net effects between spillovers effects and competition effects from leading innovative provinces to others in Vietnam.

In short, research questions of this study are: (i) What are factors driving provincial TFP in Vietnam? (ii) Could a better local governance for private business could stimulate the provincial manufacturing productivity in Vietnam? (iii) whether leading productive provinces transfer stronger productivity spillovers (positive net effects) to, or impose tougher competition (negative net effects) on other provinces?; (iv) Are there different impacts of these driving factors between the Northern region and the Southern region of Vietnam along with dimensions such as: labour intensity (because labour intensive manufacturing is the comparative advantage of Vietnam, see chapter 2 in Dinh (2013)), or on provinces with denser presence of private firms?

The research time frame of this study is within five years from 2006 to 2010 after Vietnam implemented a big mile stone in business regulations. Since 2005, the country has officially decentralized the business management to 63 provincial governments (according to the amended Law on enterprises and the new Law on investment of Vietnam in 2005). During 1999-2003, the Northern provinces were claimed to be likely more under the "state-

[4]In 2005, nearly the half of a number of provinces (31 provinces) was indicated as "fence-breakers" who authorized extralegal investment incentives for FDI investment competition. This happened due to the fiscal decentralization in Vietnam in 1996, since then provinces have been under pressure of stimulating domestic and foreign business to increase tax revenue after suffering from budget deficits, especially in small-scaled provinces. Fence-breaking does not imply a negative meaning but rather refers to reformers in taxation and land policies to attract foreign investment among provinces (Vũ et al., 2007).

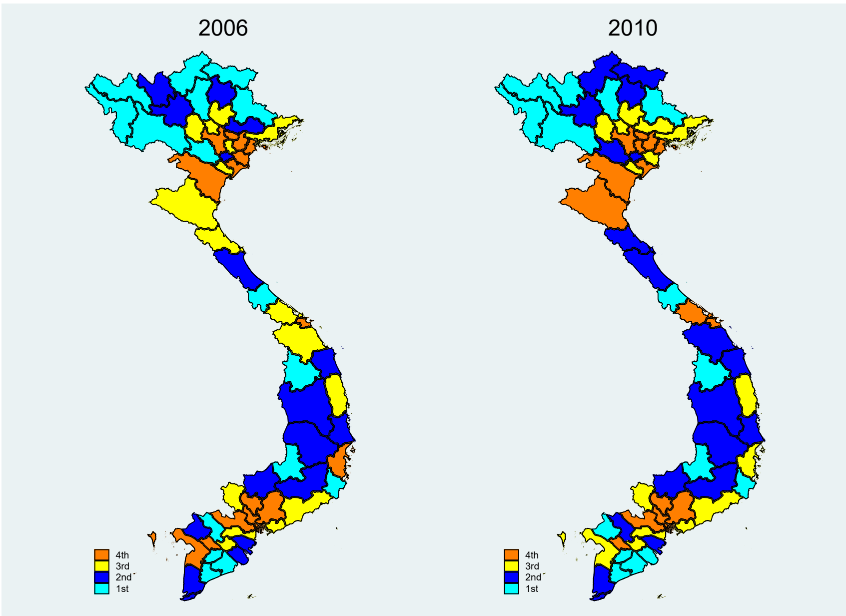
linked" subsidy from the central government (Cung, 2008), hence they were more sluggish in the ease of doing business for the private sector.[5] For the post period of business reforms (2006-2010), the exit ratio and entry ratio in the South are 0.70% and 0.83% respectively higher than in the North (Table 3.1). These statistics reveal that the Southern industries were under higher competition and more open than in the Northern market. Private firms by province are illustrated in Figure 3.1 for further investigation.

Table 3.1: Average Ratio of Firm's Exit and Entry by Region

	Exit Ratio			Entry Ratio		
	North	South	All	North	South	All
Average (%)	1.17	1.97	1.57	1.16	1.99	1.58
Observation	126	128	254	157	160	317

Source: Provincial data was aggregated from firm-level data which was drawn from the Vietnam Enterprise Survey (Vietnam General Statistic Office, 2005-2010).
Note: The ratio for each province is calculated by number of firm entering (or exiting) of the province divided by total number of entry (or exit, respectively) of the country by year. Exit ratio ratio for year 2010 cannot be calculated. Entry ratio for year 2005 cannot be calculated.

Figure 3.1: Private firms by province



Source: Author's calculation using firm-level data in 2006 and 2010.
Note: The log number of private firms is arranged by province in 1^{st}, 2^{nd}, 3^{rd} and 4^{th} percentile for each year. Private firms include private-owned firm and FDI firms. Because of limitation in boundary data, some islands belong to Vietnam (such as Hoangsa and Truongsa) were not mapped.

Figure 3.2 visualizes the provincial total factor productivity (TFP) in positive linear correlation with the provincial competitiveness index (PCI) (see details of the calcula-

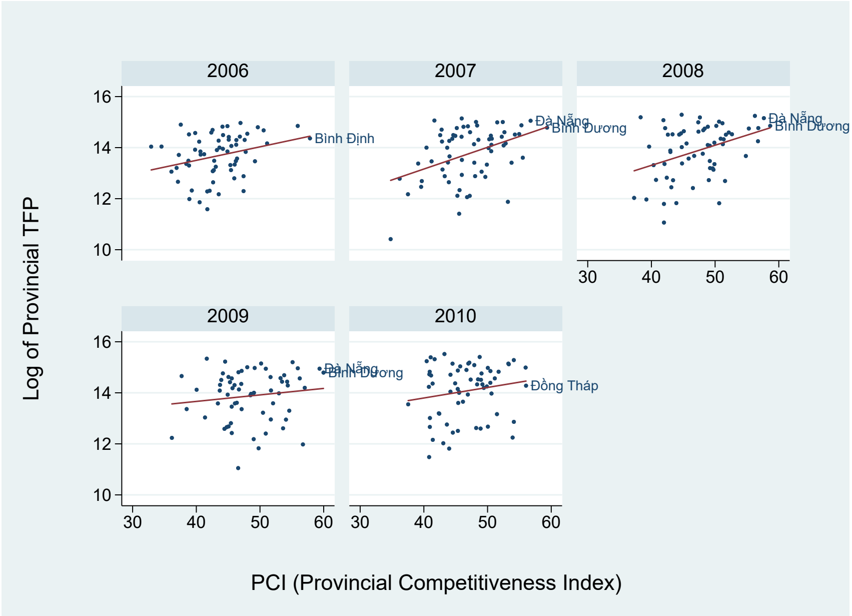
[5]For example, contribution of state-owned business in growth of Bacgiang and Quangninh (Northern Vietnam) was respectively 76% and 74% (1999-2002), while Binhduong (Southern Vietnam) gained only 2% of growth from the state-owned business in the same period (Cung, 2008).

tion in Section 3.5.2).[6] Local authorities ranked the highest provincial competitiveness index in Vietnam were Binhdinh (in the year 2006), Danang and Binhduong (in the year 2007, year 2008, and year 2009), and Dongthap (in the year 2010) (Figure 3.2). Interestingly, all these leading reformers in doing business are provinces in the South of Vietnam.[7]Competitiveness index by province, which is recalculated by the author from sub-indices, is presented in Figure 3.3.

[6]PCI is recalculated by the author using information of consistent sub-indices for the research time frame from http://orgeng.pcivietnam.vn/gioi-thieu-pci-c2.html

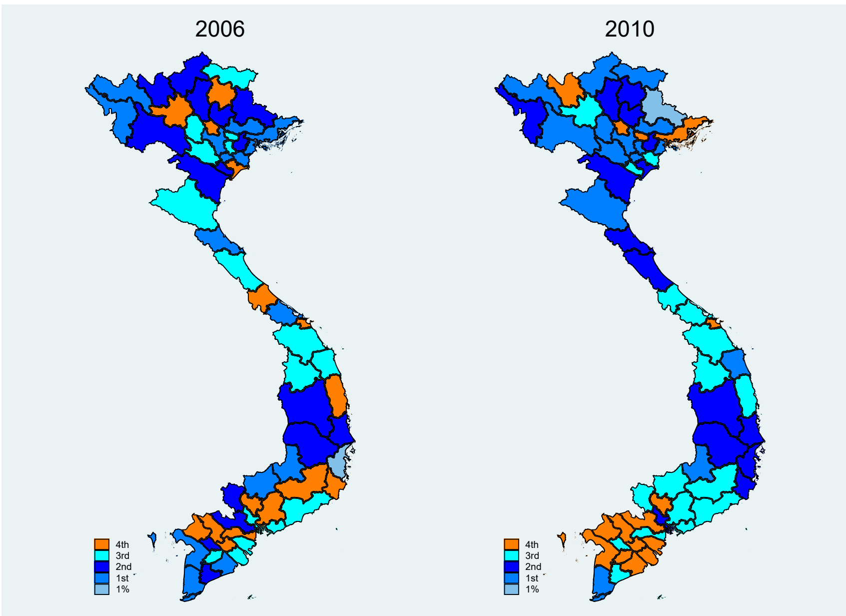
[7]Quangninh, Phutho, Hatinh, Nghean, Namdinh (Northern) were noted as "State-captured"; Northwest Vietnam, Hoabinh, Sonla, Langson, Baclieu were namely "Poor and struggling"; Southern provinces such as Hochiminh city, Danang, Binhduong, and Dongnai were nominated as "reformers" with "low-state sector dependence" (Edmund Malesky cited by Cung (2008))

Figure 3.2: Provincial TFP and Competitiveness Index



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2006-2010, and PCI data downloaded from http://eng.pcivietnam.org.
Note: Only names of provinces with PCI > 97 percentile were shown in the graphs. Provincial TFP is average weighed TFP from firm-level TFP which is measured following Wooldridge (2009a) and Petrin and Levinsohn (2012) (See Nguyen et al. (2016) for more details of firm-level TFP estimated using this measurement). PCI is an unweighted index.

Figure 3.3: Provincial Competitiveness Index by Province



Source: Author's calculation using data of sub-indices downloaded from http://eng.pcivietnam.org.
Note: The indices are arranged by province in 1^{st}, 2^{nd}, 3^{rd} and 4^{th} percentile for each year. Because of limitation in boundary data, some islands belong to Vietnam (such as Hoangsa and Truongsa) were not mapped.

The provincial competitiveness index (PCI) is used as a proxy for governance quality in Vietnam.[8] This index is calculated from data of the annual survey on provincial economic governance of Vietnam (which has been published by the Vietnam Chamber of Commerce and Industry since 2005). To solve the endogeneigty of governance quality raised in the literature as the inverse of the influence of entrepreneurs on governance (for example: Carbonara et al. (2016)), two instruments are proposed for the IVs estimation. First, after the fiscal decentralization in Vietnam since 1996, local authorities of bigger provinces are under more pressure of reforms in extending local business scale to collect more tax revenues to afford for their expenditure Vũ et al. (2007). Second, poorer provinces are more depending on the subsidy from central government (e.g.: from state owned enterprises), and their manufacturing were in limited number of industries (Vũ et al., 2007). Thus, quality of local governance is instrumented by state-linked capital stocks per head (similar to the indicator in Cung (2008)) and the province's dynamic structure in manufacturing industries (the calculation follows (Combes et al., 2010)). This argument is consistent with the development path of manufacturing in Vietnam towards the industrialization and modernization during 2006-2010 (Anh et al., 2014).

The empirical results generated from IVs estimation confirm that on average better provincial governance for private business sector stimulates the local productivity in manufacturing.[9]

Interestingly, the impacts of provincial governance are more significant and larger: (i) in the Southern region than in the Northern region of Vietnam, (ii) in provinces specializing more in higher labour intensive industries, and (iii) in provinces having more private firms (than median value) during the post-business reform from 2006 to 2010.

We exploit the firm-level data to aggregate provincial information which is then merged with the provincial data on human capital (number of university students), provincial competitiveness index, geographical data on shortest distance among provinces. Instrumental variables estimation is applied to solve the endogeneity issue (Baum et al., 2007) between local productivity and provincial government quality.

The following sections are organized as follows: The next section summarizes relevant literature in agglomeration economics and development economics. The third section proposes our specification. The forth section describes data set and measurement of variables. The fifth section presents empirical results and discusses the policy implications.

[8]See similar proxy in Malesky (2010) and McCulloch et al. (2013).

[9]Quality of provincial government was ranked by private enterprises in annual surveys for seven sub-indices: 1) low entry costs for business start-up; 2) easy access to land and security of business premises; 3) a transparent business environment and equitable business information; 4) minimal informal charges; 5) has limited time requirements for bureaucratic procedures and inspections; 6) sound labor training policies; and 7) fair and effective legal procedures for dispute resolution. See more details of the data source in http://orgeng.pcivietnam.vn/gioi-thieu-pci-c2.html

The last section draws conclusions.

3.2 Literature review

Two strands of relevant literature are reviewed in this section. The first strand is development economics which is relevant to driving factors of regional growth relating to institution, trade or human capital. The second strand is the literature in which density of cities and specialization of industries are analysed as factors influencing regional economic performance (agglomeration economics). There are also studies including both strands of the literature. For example: "institution", "human capital", and "specialization" are summarized as three driving forces explaining why cities grow differently (Storper, 2010).

North (1994) reviews the role of institutions in productivity enhancement and discusses the lowering of firm's transaction cost due to improvement in performance of government. Similarly, Acemoglu et al. (2006) suggest that good quality of government fosters better economic performance. Djankov et al. (2006) also confirm that growth is brought about by the ease of doing business which reduces business cost for firms. Breinlich et al. (2014) give a fundamental review for ample theoretical and empirical studies for regional growth in both developed and developing countries, however, the case of Vietnam was not reported.

The business reforms in Vietnam since 2000 and especially the business decentralization since 2005 have spurred a large number of studies on local institutional variation of the country. For example, quality of local governance in Vietnam has been paid much attention in connection with the level of foreign direct investment to each provinces (Malesky and Taussig, 2009), or to the formation of business (Malesky, 2010). Local authorities from the North were claimed to be more likely dependent on the state-owned business while leading business reformers and "fence-breakers" were provinces in the South of Vietnam (Cung, 2008). Although the literature focuses on economic growth in Vietnam, these studies did not investigate driving factors for TFP level at provinces, and how the impacts of the factor are different between the Northern and the Southern provinces. Moreover, it is also important to investigate which TFP driving factors are more important for which regions in Vietnam because the economy of each region was formed historically in different settings. This question is still not answered in the existing literature for the case study of Vietnam.

Previous studies in agglomeration economies commonly indicate that total factor productivity is higher in bigger cities (see Combes et al. (2012) for case studies of cities in France), and larger cities gains more labour productivity (Glaeser (2010)). In a study ofCombes et al. (2008), the combination between agglomeration and market potentials was computed by an inverse discounted index of distance to nearby areas and density of the areas. Theoretical model introduced by Desmet and Rossi-Hansberg (2010) investigates the factor mobility and diffusion across space among cities. In the model of Desmet and Rossi-Hansberg (2010), congestion and agglomeration are the two forces that form both partial employment pattern and aggregate growth; higher trade cost among cities implies more regional concentration of production which creates innovation, and high innovative areas induce diffusion to other areas in its neighbourhood.[10] While Desmet and Rossi-Hansberg (2010) only refer to the innovation spillovers from excessively innovative provinces to others, this study considers the peers effects which is the interaction between spillovers effects and competition effects from the elite provinces - "the peer net effects".

This study follows the theoretical framework and empirical studies that combine these two strands of literature in development economics and agglomeration economics. It estimates the influences of the local government quality on provincial total factor productivity (Dixit (2009) and Dixit (2015)), considers other productivity driving factors, such as labour density (net effects of congestion and agglomeration effect, see: Combes et al. (2012) and Desmet and Rossi-Hansberg (2010)), human capital (number of undergraduates and college students, Romer (1990)), and the presence of FDI firms (as the proxy for FDI spillovers, Javorcik (2004)).

3.3 Theoretical model

This section summarizes the backbone theoretical model, which is applied in this study, is introduced by Desmet and Rossi-Hansberg (2010). This model shows the factor mobility and diffusion across space.

In this model, (i) congestion and agglomeration are the two forces that form both partial employment pattern and aggregate growth, (ii) higher trade cost implies more regional concentration of production which creates innovation, (iii) more innovation in a certain area diffuses to other areas in its neighbourhood.[11] The model is summarized as follows:

* **Assumptions of the model**: Production of one good (no specialization or inter-industry innovation effects). Factor mobility is frictionless, trade is the results of agents with land portfolio across locations. Land $\in$ [0,1], total population is $\overline{L}$. Space is divided

[10]See example of supporting evidence of human capital spillovers in Rosenthal and Strange (2008).

[11]See for example of supporting evidence of human capital spillovers in Rosenthal and Strange (2008).

into "provinces" in which there is one local government for each province.[12] Provinces are assumed to be small and innovation spillovers across provinces. Free mobility: utility is equalized across regions.

* **Agent**

$$\max \text{ E} \sum_{t=0}^{\infty} \beta^t U(c(\ell,t))$$

subject to: $\omega(\ell,t) + \dfrac{\overline{R(t)}}{=} p(\ell,t)c(\ell,t)$ for all ℓ and t

$p(\ell,t)$: is good price $\omega(\ell,t)$: is wage at location l and time t. $\overline{R(t)}$: is total land rent. $\dfrac{\overline{R(t)}}{\bar{L}}$: dividend from land ownership.

* **Producer**

Producer in location ℓ at time t

$$\max(1-\tau(\ell,t))(p(\ell,t)Z(\ell,t)L(\ell,t)^{\mu} - \omega(\ell,t)L(\ell,t))$$

Where:

Production per unit of land: $y(L(\ell,t)) = Z(\ell,t)L(\ell,t)^{\mu}$

$\mu < 1$

$Z(\ell,t)$: total factor productivity

$L(\ell,t)$: labour per unit of land

$\tau(\ell,t)$: tax imposed on profit and paid to the local government.

* **Government(risk-neutral)**

$$max_{\Phi(\ell,t)} \int_G \frac{\phi(\ell,t)}{a-1} p(\ell,t)Z(\ell,t)L(\ell,t)^{\mu} d\ell - I\Psi(\Phi)$$

Provincial government influences on local innovation by buying innovation at a price $\Psi(\Phi)$ per unit of land from the tax revenue imposed on the firms. Firms in each province

[12] Desmet and Rossi-Hansberg (2010) divide space into counties.

can equally access to the local innovation. Consequently, this improves the level of local TFP $z_\ell Z_i(\ell,t)$ in which: $Pr[z < z_\ell] = (\frac{1}{z})^a$.

* **Innovation diffusion**

Each province gains the best spatially discounted technology of the previous year (Trade surplus is in a high productivity location, trade deficit is in low productivity location):

$$Z_i(\ell, t+1) = \max\ e^{-\delta|\ell - r|} Z(r,t)$$

* **Iceberg transport cost**

A unit of good is transported from province ℓ to province r, then price good in r is $p(r,t) = e^{\kappa|\ell - r|} p(\ell, t)$.

* **Goods market clearing condition (sequentially)** Stock of excess $H(O,t)$ between location O and location ℓ:

$$H(O,t) = \frac{\partial H(\ell,t)}{\partial \ell} = \theta(\ell,t)x(\ell,t) - c(\ell,t)\left(\Sigma\theta(\ell,t)L(\ell,t)\right) - \kappa|H(\ell,t)|$$

Change in the excess surplus is the difference between the quantity produced and the quantity consumed. The good market clears if $H(1,t) = 0$

* **Labour market clearing condition** $\int_0^1 L(\ell,t)d\ell = \overline{L}, \forall$

3.4 Baseline Specification

The baseline estimation is proposed in Equation 3.1 (Section 3.5.2 gives more details about how to calculate these variables in equation 3.1 using firm-level dataset and provincial-level dataset):

$$y_{p,t} = X_{p,t}\beta + YEAR_t\alpha + MUNI_p\theta + YEAR_tMUNI_p\sigma + u_{p,t} \tag{3.1}$$

Where :

$y_{p,t}$ is $logTFP_{p,t}$: the total factor productivity of province p at time t.

$X_{p,t}$ includes $PCI_{p,t}$ and other control variables for province time-variant characteristics.

$PCI_{p,t}$ is the provincial competitiveness index proxied for the quality of local government in business reforms for the private sector (including both domestic and foreign direct invested firms). The index covers categories for smart local authorities supporting firms to start-up, better access to land and facilities with transparent policies, less informal cost, training, and effective legal procedure, etc. (See Section 3.5.2 for more details). These categories are consistent to evaluate the quality of local economic governance (Dixit, 2009).

In Equation 3.1, PCI is observed at time t. The timing of $PCI_{p,t}$ follows Acemoglu et al. (2006, p.6) such that good quality of government at time t induces better economics performance at time t. This setting is reasonable as the improvement of local governance during the year reduces cost for enterprises (North, 1994).

Since there are concerns of endogeneity between $TFP_{p,t}$ and $PCI_{p,t}$, excluded instruments are chosen for $PCI_{p,t}$. After the fiscal decentralization in Vietnam since 1996 (less subsidy from the central government), local authorities of provinces with more state-linked business are under more pressure of reforms in extending local business scale to collect more tax revenues to afford for their expenditure. Hence, $PCI_{p,t}$ of a province p is instrumented by its local business scale of state-owned enterprises $STATELINK_{p,t-1}$ (i.e.: $STATELINK_{p,t-1}$ is proxied by its total real value of capital stocks per capita). Additionally, poorer provinces depend more on the subsidy from central government, e.g.: from state owned enterprises, and their manufacturing industries are in limited in number (Vũ et al., 2007). Thus $PCI_{p,t}$ is also instrumented by the province's dynamic structure in manufacturing industries $DIVERSITY_{p,t-1}$ (the calculation follows (Combes et al., 2010)). In addition, Lagged value $PCI_{p,t-1}$ is also included as an instrument. The robustness check is made by replacing $PCI_{p,t-1}$ with square of $PCI_{p,t-1}$.

Exogenous instruments Z_p including $STATELINK_{p,t-1}$, $DIVERSITY_{p,t-1}$, and $PCI_{p,t-1}$ are assumed:

$$E(Z_{p,t-1}PCI_{p,t}) \neq 0$$
$$E(Z_{p,t-1}u_{p,t}) = 0$$

To prove the validity of these instruments used in the model, in the presence of heteroskedasticity (non i.i.d. errors), Lagrange multiplier (LM) version of the Kleibergen-Paap (2006) rk statistic test is applied (Baum et al., 2007), and its p-value is reported in Table 3.2. F values for weak identification test (Kleibergen-Paap) are also presented in Table 3.2.

Other control variables are lagged one year to assume the orthogonal condition such that $E(X_{p,t-1}u_{p,t}) = 0$. $X_{p,t-1}$ includes explanatory variables that possibly have an influence on productivity following the existing literature, such as: labour density (Combes et al., 2012), human capital (Romer, 1990), and FDI spillovers from the presence of FDI firms Javorcik (2004). Net effects from leading performers (spillovers and competition) are also considered under the context of inverse distance to the leading provinces. This also counts for the ice-berg transport cost referred by Desmet and Rossi-Hansberg (2010).

Moreover, $YEAR_t$ is year fixed effect dummy controlling for unobservable time-variant provincial characteristics.

$MUNI_p$ is used to control for the municipality fixed effect dummy. The municipalities are the five biggest provinces directly under the control of the central government in Vietnam (Hanoi capital, Hochiminh city, Haiphong, Danang, and Cantho). The dummies and their interaction with year fixed effects could control for the changes (in economic conditions, policies, etc.) by years across municipality and non-municipal areas. The interaction between year and province is not added as the change in year-provincial-level policies are proxied by $PCI_{p,t}$. Moreover, the provincial level dataset is constrained by number of observations.

3.5 Data Analysis

3.5.1 *Data description*

A panel provincial data set for 63 provinces in Vietnam is built by aggregating the firm-level data from the annual enterprise survey of Vietnam.[13] The data is collected and managed by the General Statistics Office of Vietnam, GSO (See more details of the firm-level data in Nguyen (2017a)). The aggregated data set includes provincial-level information about total factor productivity (TFP), total capital stocks, number of workers, number of foreign invested firms, etc. in manufacturing industries.[14] Industry classification is 2-digit VSIC-1993 which is equivalent to the 2-digit international classification ISIC Rev.3 constructed by the United Nation.[15] Some adjustments of industrial code are explained in Appendix 3.A.1.

To obtain further control variables (such as: labour force, number of students), provincial data is downloaded from the GSO website: www.gso.gov.vn.

The performance of local government is calculated by the author using sub-indices that were downloaded from www.pcivietnam.org.

More details of variable constructions are explained in Section 3.5.2.

3.5.2 *Measurement of variables*

1. **Dependent variable** Provincial total factor productivity $LogTFP_{pt}$: is the weighted average provincial TFP of province p at time t. It implies the level of technology efficiency in manufacturing of the province. TFP_{ipt} at firm-level is measured for each 2-digit VSIC1993 classified industry using the instrumental variables estimation method of Wooldridge (2009b) and Petrin and Levinsohn (2012). The limitation of this paper is that we lack of data on land used by each firm. Hence, the inputs for firm-level production is real value of capital and labour following and Petrin et al. (2004).[16] Using weighted average counts for the structure of industries in each

[13]Before 2008, there were 64 provinces in the country. Since 2008, Hatay province has been merged with Hanoi capital.

[14]See more details of firm-level data in Nguyen (2017a).

[15]See more details in this link: https://unstats.un.org/unsd/cr/ctryreg/ctrydetail.asp?id=1448

[16]For the case study of Vietnam: the similar method of TFP estimation is applied in Nguyen (2017a) for firm-level TFP and analysis in industry-region TFP gap; Newman et al. (2014) also use similar method to estimate TFP for Vietnamese firm.

province.[17]

The Weighted average TFP of province p at year t is:

$$LogTFP_{pt} = log(\sum \gamma_{ipt} * TFP_{ipt}) \tag{3.2}$$

Where share of labour of firm i in province p at time t is :

$$\gamma_{ipt} = \frac{L_{ipt}}{\sum_i L_{ipt}} \tag{3.3}$$

2. **Independent variable** $X_{p,t-1}$ are controlling for:

- Labour density $LABOR_{p,t-1}$ is calculated by number of residents whose age is 15 years upwards per km^2 by province (see Combes et al. (2012) for the similar construction of labour density).

- FDI density by province $FDI_{p,t-1}$ is the presence of foreign firms in each province which is a ratio of number of foreign firms over total number of firms in manufacturing locating in the province.

- Human capital of the province $STUDENT_{p,t-1}$ is proxied by a number of university and college students by province. [18]

- Labour intensity $LABOUR - INTENSITY_{p,t-1}$ is the ratio of number of workers over total capital stocks of each province. This variable is used to indicate if the province has more potentials in higher labour intensive industries if the province has value of labour intensity higher than the medium value of labour intensity in the country.

- As the development of each province is not limited within its border but influenced by its neighbourhood (Storper, 2013), we investigate net influence of elite provinces on other provinces. Thus, we hypothesize that when spillovers effects from the elites are higher than their competition effects, the net effect is positive. This means one province could gain from other elites.

[17] Labour share being used as weight to calculate average TFP is also applied in Combes et al. (2012).

[18] In the case of Daknong province, as there are several missing values for some years, we calculate the values of the missing years by the average value of the two nearest years. Dienbien also has missing values, I replace missing values with number of students of this province in vocational school.

The peers net effects $PEERS_{p,t-1}$ is the sum of $TFP_{q,t-1}$ from elite province q discounted by its distance to province p. Inverse distance implies decaying effects of distance that closer elites could transfer more spillovers. [19]

$$PEERS_{pq,t} = \sum_{q,q \neq p} \frac{TFP_{q,t}}{Distance_{pq,t}} \tag{3.4}$$

Where:

The elites (peers) are province q that gains $TFP_{q,t}$ value ranging in the forth quartile (higher than 75th percentile TFP) of the provincial TFP distribution at year t. $Distance_{pq,t}$ is measured by the shortest geodetic distance between two centroids of province p and elite province q.[20]

- Provincial competitiveness index ($PCI_{p,t}$)

  In the literature, the interaction of the province fixed effects variable and the year effects variable is used to absorb the effects of decentralization to the province (for an example, see Ahrend et al. (2014)). However, the interaction could not indicate the impacts of changes in provincial policies. Hence, I include the provincial competitive index PCI_{pt} instead of using the interaction of province and time effects. The index could be used as the proxy for quality of local economic governance that could ensure the "three prerequisites of market economies" (Dixit (2009), p.2): "security of property rights", "enforcement of contracts", and "collective action".

  Data to calculate $PCI_{p,t}$ used in this study is from the annual surveys on "assessment and ranking of the economic governance quality of provincial authorities in creating a favourable business environment for development of the private sector" from 2006 to 2010 in 63 provinces in Vietnam. The survey has been run annually by the Vietnam Chamber of Commerce and Industry (VCCI) and the United States Agency for International Development (the USAIDS) since 2005 for selected provinces and since 2006 for all 63 provinces.

  The $PCI_{p,t}$ used in this study is calculated as an unweighted index by summing up seven sub-indices consistently surveyed for 5 years in the research time frame. The seven sub-indices are: "1) low entry costs for business start-up; 2) easy access to

[19] For examples of research in inverse weighted distance approach, see Combes et al. (2012), Keller (2002), and Bodman and Le (2013). Combes et al. (2008) construct an index for average discounted distance to large potential markets.

[20] Distance across provinces is assumed to be proportional with travel costs, and travel costs not only prevent the direct business interactions among firms from locations, but also hinder the face-to-face meetings among government officials from different provinces. Storper (2013) noted that in face-to-face contacts imitation and competition are stimulated.

land and security of business premises; 3) a transparent business environment and equitable business information; 4) minimal informal charges; 5) has limited time requirements for bureaucratic procedures and inspections; 6) sound labor training policies; and 7) fair and effective legal procedures for dispute resolution.[21]

- Two excluded instruments for $PCI_{p,t}$ (in log values) (included instruments are other exogenous control variables):

  $STATE-LINK_{p,t}$ is state-owned capital stocks per capita. It is calculated from the firm-level data by summing up real value of total capital of state-owned enterprises for province p (except for join companies between the state and foreign investors), then dividing by number of citizen older than 15 years' old in the province.

  $DIVERSITY_{p,t}$ is measured as the inverse-Herfindahl index: j is denoted for industry, p is denoted for province, t is denoted for year (Combes et al., 2008).

$$DIVERSITY_{p,t} = \frac{L^2_{p,t}}{\Sigma_j L^2_{j,p,t}} \tag{3.5}$$

Table A.2 in the Appendix summarizes the measurement of variables and tests applied for the validity of Instruments.

[21] Other sub-indices are not added in the recalculation because they are not included in the surveys for all the selected years, such as: 8) limit crowding out of private activity from policy biases toward state, foreign, or connected firms; 9) proactive and creative provincial leadership in solving problems for enterprises; 10) developed and high-quality business support services"(See more details in http://eng.pcivietnam.org).

3.6 Empirical results and Discussion

In Table 3.2, IVs regression results in column (1), (2), and (3) using all samples of 63 provinces. Column (4) and (5) show the results for Northern and Southern provinces separately. Because labour intensive industries are regarded as important strategies for Vietnam as an alternative for agriculture, in column (6) and (7), the results are presented for two groups of provinces: (i) more specialization in labour intensive industries (log of labour intensity higher than median) and (ii) less specialization in labour intensive industries (the rest of the samples). The last two columns: (8) and (9) respectively reveal the results for (i) provinces with denser presence of private firms (log number of private firms in the province is higher than median value of the country) and (ii) provinces with less presence of private firm (vice versa).

In general, agglomeration effects, labour intensity, human capital, FDI density and expansion of private firms in the lagged values positively and significantly enhance the local TFP of current time (See Table 3.2).

Particularly, Table 3.2 indicates that provincial competitiveness fostered the local TFP, however, the effects are stronger and more significant in the Southern region (column 4 & 5), in the group of provinces with higher labour intensity (column 6 & 7), and in the group of provinces experienced more presence of private firms (column 8 & 9).

The different impacts of other key factors on provincial productivity in the North and the South of Vietnam are shown in column (4) and (5) of Table 3.2.

In particular, labour agglomeration generates quite similar causal effects for provinces in the North and in the South. On average, the coefficients are 0.041 for the North, and 0.045 for the South (these results are consistent with the range from 0.02-0.10 in (Combes et al., 2012)).

Regarding to human capital, its causal effects of 1% rise in log number of students range from 0.116-0.201 on average (column (1)-(3) in Table 3.2). These results confirm for the importance of human capital in productivity stimulation. However, human capital induces less significant and lower impacts in the North (column (4) & (5) in Table 3.2).

Apart from that, 1% increase in FDI density induces stronger FDI spillovers in the Northern provinces (coefficient = 0.304) than in the Southern provinces (coefficient = 0.266).

Besides, increase in labour intensity only generates positive effects in the Southern region. This might be due to the fact that development of high labour-intensity manufacturing industries fit more with the economic conditions of the South (Dinh, 2013).

In addition, provinces with a 1% larger expansion of private firms in the North gain higher TFP than province with less private firms in the same region, i.e.: 0.927% higher (Table 3.2, column (4)). The same figure for the South is only 0.532% (Table 3.2, column (5)).

Interestingly, the net effects of leading provinces stimulate positive and more significant impacts on Northern provinces (Table 3.2, column 4) while they are negative for the local TFP of Southern provinces (Table 3.2, column 5). This might be due to the more competition effects in the Southern region.

As shown in column (6) and (7) of Table 3.2, provinces with higher labour intensities experience a greater influence on local TFP from the driving factors relatively to provinces with less labour intensity, except for the impacts from FDI density. This is in favour of the fact that high labour intensive industries are the comparative advantage of Vietnam and they are the heart of the manufacturing development path in Vietnam from 2005-2010. The less influence of FDI density in high labour intensive provinces still needs further investigation, nevertheless, it might be because FDI spillovers are generated more through capital intensity investment.

Provinces with a higher number of private firms gain more positive impacts, which are induced from local government competitiveness for private sectors, FDI spillovers, and high labour intensive industry specialization, on their TFP.

Results in Table 3.3 are generated for the robustness check with the similar specification, but the IV used is the square (lagged one year) of $PCI_{p,t}$ instead of the lagged value: $PCI_{p,t-1}$. The results are in similar pattern to the results in Table 3.2. Tests for IVs in both results support for the identification and reject the weak IVs hypothesis.

Table 3.2: Driving factors of provincial TFP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All samples	All samples	All samples	North	South	More light industries	Less light industries	More private firm	Less private firm
Provincial Competitiveness	0.033*	0.059**	0.063**	-0.013	0.068***	0.084*	0.000	0.051***	0.011
	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.01)	(0.04)
Labour density	0.032**	0.034*	0.037*	0.041**	0.045**	0.078***	0.023*	0.021*	0.259***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.05)
Human capital	0.116***	0.201***	0.196***	0.054+	0.115**	0.131**	0.080*	0.034	0.235***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)	(0.03)	(0.07)
FDI density	0.267***	0.321***	0.329***	0.304**	0.266***	0.268***	0.355***	0.353***	0.155+
	(0.04)	(0.04)	(0.04)	(0.09)	(0.07)	(0.07)	(0.05)	(0.04)	(0.08)
Labour intensity	0.289***	0.265**	0.265**	0.240	0.215+			0.524***	0.105
	(0.07)	(0.09)	(0.09)	(0.18)	(0.13)			(0.07)	(0.14)
Peer net effects	0.030	0.062	-0.104	0.306**	-0.210+	-0.089	0.071	0.033	-0.063
	(0.06)	(0.07)	(0.12)	(0.12)	(0.12)	(0.10)	(0.08)	(0.05)	(0.12)
More presence of private firms	0.753***			0.927***	0.532***	0.757***	0.599***		
(Dummy = 1 if log(private) > median)	(0.11)			(0.17)	(0.13)	(0.14)	(0.18)		
North (Dummy =1)		0.203	0.971*			0.420	0.029	0.382***	-0.086
		(0.16)	(0.43)			(0.28)	(0.17)	(0.10)	(0.26)
North#Peer net effects			0.229+						
			(0.12)						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE#Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	204	204	204	99	105	100	104	126	78
F value of weak identification KP test	92.736	79.966	77.536	16.718	45.370	15.780	49.905	55.895	28.250
P-value of underidentification LM statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: Standard errors (SE) robust to heteroskedasticity are in brackets. High intensity are provinces with labour intensity higher than median. More privatization are provinces with number of private firms higher than median. KP test is Kleibergen-Paap rk. LM test is Lagrange multiplier test.
+ $p < 0.10$;* $p < 0.05$;
** $p < 0.01$;
*** $p < 0.001$.

Table 3.3: Driving factors of provincial TFP (robustness check)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All samples	All samples	All samples	North	South	More light industries	Less light industries	More private firm	Less private firm
Provincial Competitiveness	0.034*	0.060**	0.065**	-0.012	0.069***	0.086*	0.000	0.051***	0.012
	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.01)	(0.04)
Labour density	0.033**	0.034*	0.037*	0.040**	0.046**	0.078***	0.023*	0.021*	0.258***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.05)
Human capital	0.116***	0.201***	0.196***	0.054+	0.115**	0.131**	0.080*	0.034	0.235***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)	(0.03)	(0.07)
FDI density	0.267***	0.321***	0.329***	0.302**	0.267***	0.269***	0.355***	0.353***	0.155+
	(0.04)	(0.04)	(0.04)	(0.10)	(0.07)	(0.07)	(0.05)	(0.04)	(0.08)
Labour intensity	0.289***	0.267**	0.268**	0.243	0.215+			0.525***	0.106
	(0.07)	(0.09)	(0.09)	(0.18)	(0.13)			(0.07)	(0.14)
Peer net effects	0.029	0.060	-0.108	0.306**	-0.213+	-0.092	0.071	0.032	-0.064
	(0.06)	(0.07)	(0.12)	(0.12)	(0.12)	(0.11)	(0.08)	(0.05)	(0.12)
More private firms	0.752***			0.926***	0.530***	0.753***	0.599***		
(Dummy = 1 if log(private) > median)	(0.11)			(0.17)	(0.13)	(0.14)	(0.18)		
North (Dummy =1)		0.210	0.987*			0.433	0.029	0.385***	-0.082
		(0.16)	(0.43)			(0.30)	(0.17)	(0.10)	(0.26)
North#Peer net effects			0.231+						
			(0.12)						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE#Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	0.998	0.997	0.997	0.998	0.998	0.998	0.998	0.999	0.997
F value of weak identification KP test	75.116	63.115	61.194	14.118	38.947	13.866	38.013	50.163	23.510
P-value of underidentification LM statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: Standard errors (SE) robust to heteroskedasticity are in brackets. High intensity are provinces with labour intensity higher than median. More privatization are provinces with number of private firms higher than median. KP test is Kleibergen-Paap rk. LM test is Lagrange multiplier test.
+ $p < 0.10$;* $p < 0.05$;
** $p < 0.01$;
*** $p < 0.001$.

3.7 Conclusion

Local government has been fully established to decentralize the central authorities control on business regulations to local authorities since the adjusted law on enterprises came into effects since 2006 in Vietnam. This is a following step of business reforms after the New law on enterprises in 2000 allows more private firms to register in the country. Additionally, the decentralization has transferred more power to local authorities to make decisions when they were in the cross-road for local economic development (e.g.: thirty-one fence-breaking provinces gave extralegal incentives for foreign investment in 2005). The decentralization has provided more convenient and efficient public administrative services to enterprises, and encourages better interaction between local authorities and entrepreneurs, especially private sector. Nevertheless, the policy and other key driving factors influences differently across Vietnam because of differences in geographical advantages, historical economic conditions, and economic structure.

As there is scepticism of high growth rate in output which might be incorrectly reported by provincial authorities (Cung, 2008), analysing provincial TFP aggregated from firm-level data prevents the concerns. On average, agglomeration effects, density of fdi, labor intensity, human capital, and the expansion of private firms positively influence provincial productivity. More interestingly, local government competitiveness for private sector in the Southern provinces shows more support to the provincial productivity than the Northern provinces. This might be because there were less likely of state-linked and more market orientation in the Southern region as Cung (2008) mentioned. Apart from that, net effects of nearby leading productive provinces are found positively in the North possibly due to stronger spillovers effects. Nevertheless, the result shows negative effects of peer net effects in the South possibly due to stronger competition among provinces.

The results are in favour of further local business reforms for private sector in the Northern region of the country. Moreover, local governments in Northern provinces may stimulate more business interaction among provinces by enhancing local infrastructure, make larger investment in education, and improve favourable economic conditions to attract advanced technologies from FDI. As high labour intensive industry are a comparative advantage of the South, it could be enhanced towards more advanced technology. Provincial government in the South may give more supports to develop local education, may take better advantage of the local governance quality to create larger number of economic hubs (to exploit the positive impacts of labour density), and to attract more FDI with higher technology. Further research on taxation incentives of each province and their policy for specific sectors would give a more complete picture for policy makers in the country.

3.A Appendix

3.A.1 *Selected industries*

Table A.1: List of selected industries

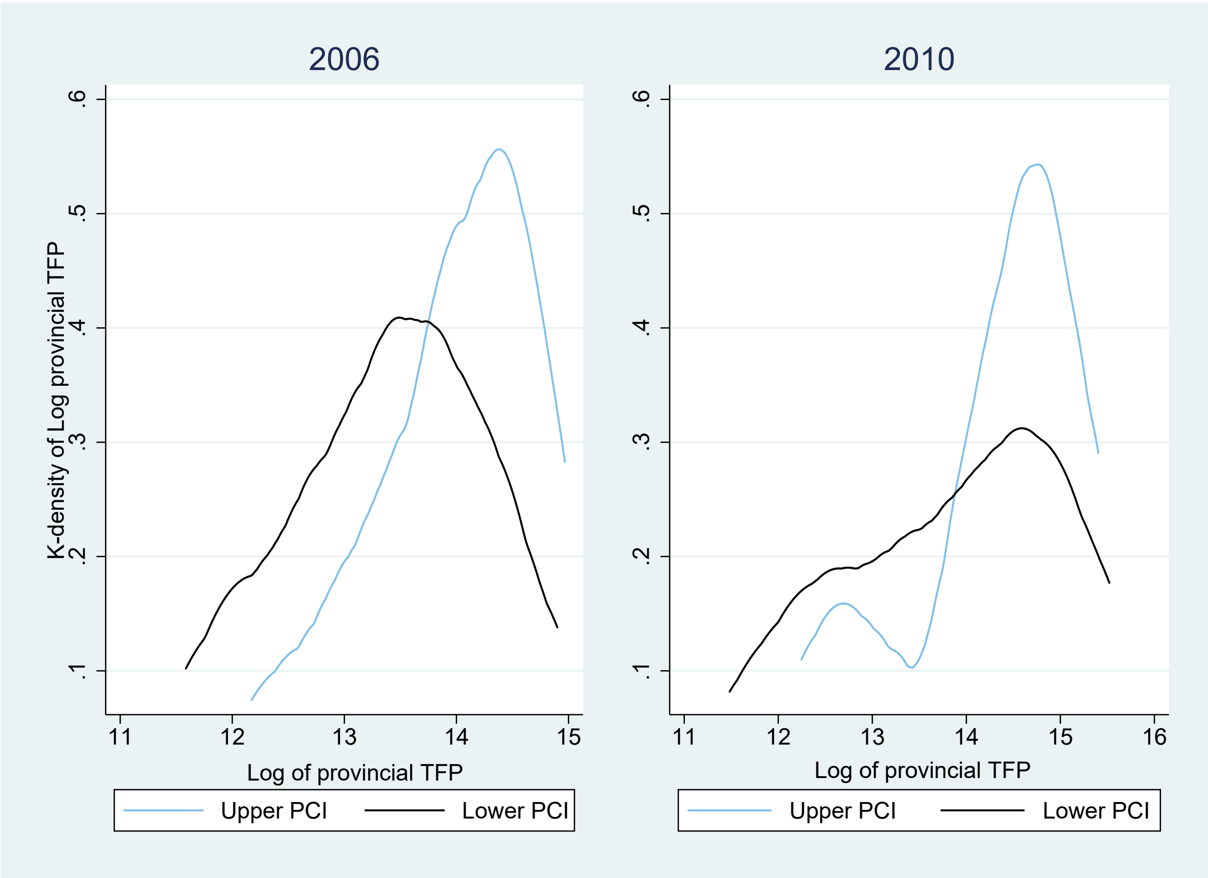
Industry	**Classification**
Food Products	15
Textiles	17
Wearing Apparel; Dressing & Dying of Fur	18
Tanning & Dressing of Leather	19
Products of Wood	20
Paper products	21
Publishing, Printing & Reproduction of Recorded Media	22
Chemical Products	24
Rubber & Plastic Products	25
Other Non-metallic Products	26
Basic Metals	27
Fabricated Metal Products	28
Machinery & Equipment, Office, Accounting & Computing Machinery	29
Electrical Machinery & Apparatus n.e.c; Television & Communication Equipment	31
Medical, Precision & Optical Instruments	33
Motor Vehicles, Trailers & Semi-trailers & Other Transport Equipment	34
Furniture; Manufacturing n.e.c	36

Several industries are merged together: industry classified as 29 includes VSIC-1993 industry 29 and 30; industry classified as 31 includes VSIC-1993 industry 31 and 32; industry classified as 33 includes VSIC-1993 industry 33 and 34. Industry classified as 15 (the Manufacture of Food) includes VSIC-1993 4-digit: 1511, 1512, 1513, 1514, 1520.

3.A.2 *Kernel distributions of provincial TFP*

Kernel distributions of log provincial TFP are plot in Figure A.1 for two groups: (i) upper Provincial Competitiveness Index (PCI), and (ii) lower PCI. Higher PCI is proxied for better performance of governance. The distribution indicates that for $logTFP$ is approximately greater than 14, provinces with better governance performance gain higher TFP than the below-median PCI group in both year 2006 and year 2010. The stronger evidence is given in the graph for the year 2010.

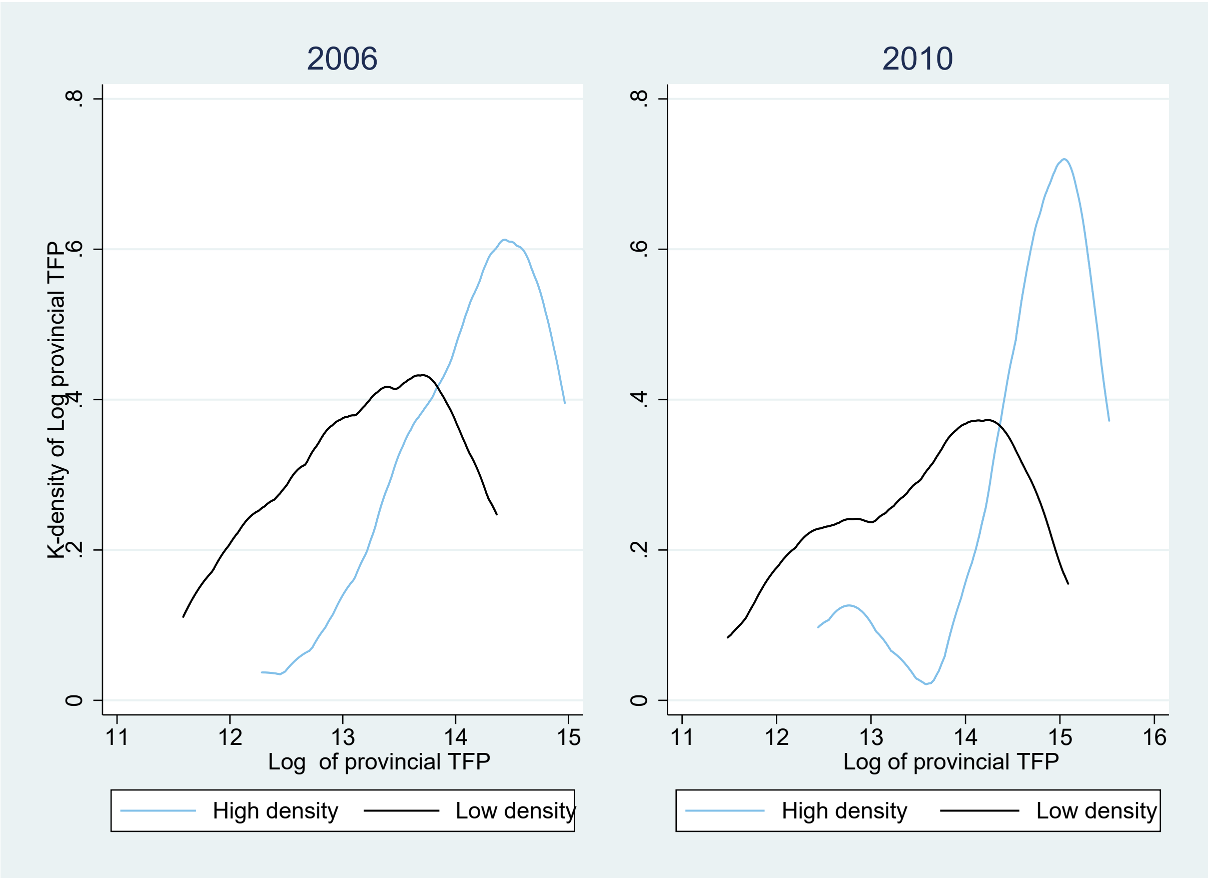
Figure A.1: Kernel distributions of Provincial TFP for upper PCI and lower PCI provinces



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2006-2010, and PCI data downloaded from http://eng.pcivietnam.org.
Note: Provincial TFP is average weighed TFP from firm-level TFP which is measured following Wooldridge (2009a) and Petrin and Levinsohn (2012). PCI is an unweighed index.

Kernel distributions of provincial log TFP by high (above-median) and low (below-median) labour density groups of provinces are plot in Figure A.2 for two years 2006 and 2010. Similarly to Combes et al. (2012) for the case study of provincial TFP in France, Figure A.2 indicates that denser provinces in Vietnam obtained higher log TFP on average for both selected years.

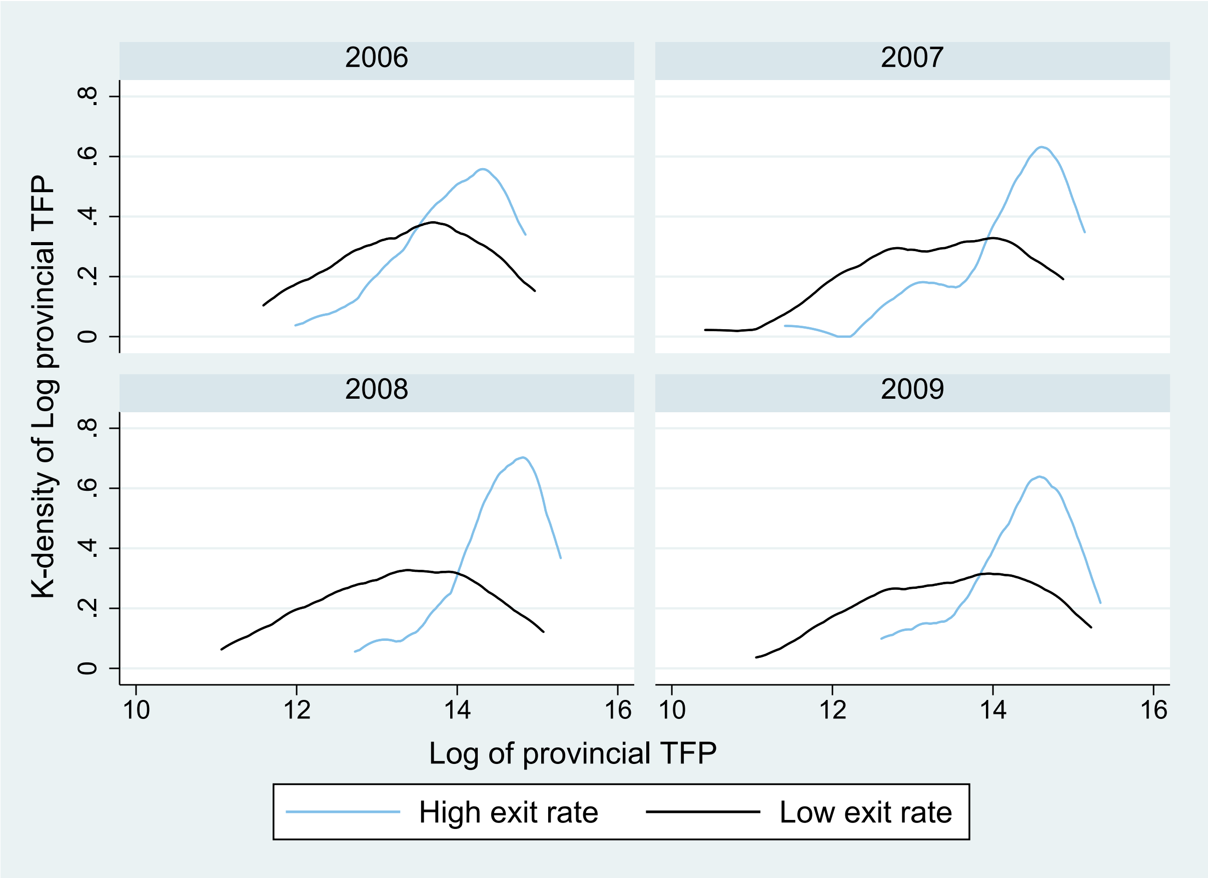
Figure A.2: Kernel distribution of Provincial TFP for High and Low labour density provinces



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2006-2010. Information about labour per km square is calculated by the author using data downloaded from www.gso.org.vn
Note: Provincial TFP is average weighed TFP from firm-level TFP which is measured following Wooldridge (2009a) and Petrin and Levinsohn (2012). Labour density is calculated as number of citizens per km square whose age is at least 15 years old (Combes et al., 2012).

Apart from that, Kernel distributions of provincial TFP in Vietnam by groups of high and low firm's exit ratio (Figure A.3) show that higher productivity provinces (above-median) forces more number of less productive firms to exit the market on average (in Combes et al. (2012), the argument about "Darwinian selection of firms" is in favour of a larger city where there is tougher competition).

Figure A.3: Kernel distribution of Provincial TFP for High and Low exit ratio provinces



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2006-2010.
Note: Provincial TFP is average weighed TFP from firm-level TFP which is measured following Wooldridge (2009a) and Petrin and Levinsohn (2012). Exit ratio of manufacturers in each province is calculated as share of number of manufacturers exiting by province in total number of manufacturers exiting in the country for each year. Year 2010 is excluded as it is the final year of the time series.

Table A.2: Measurement of Variables & Tests applied for Instruments

Province-level Variable	Measurement	Source
Provincial TFP	Firm-level TFP weighted by labour share	Enterprise survey of Vietnam Firm-level TFP calculated in Nguyen (2017b)
Provincial Competitiveness	Unweighted Recalculated from sub-indices	http://orgeng.pcivietnam.vn/
Labour density	People older than 15 years' old per km^2	www.gso.org.vn
Human capital (log value)	Number of students in college or University	www.gso.org.vn
FDI density (log value)	Ratio of FDI firms relative to total firms	Firm-level data Vietnam enterprise survey
Light industry specialization Labour intensity (log value)	Ratio of labour relative to real capital stocks	Firm-level data Vietnam enterprise survey
Peer net effects (log value)	Weighted TFP of peers by their inverse distance ratio	Firm-level data Vietnam enterprise survey
More presence of private firms (dummy =1)	Province with number of private firms higher than median value	Firm-level data Vietnam enterprise survey
North (dummy=1)	Provinces in the Northern regions	Firm-level data Vietnam enterprise survey
North#Peer net effects	Interaction term	
State-link (log value)	State-owned real value capital stocks per citizen older than 15 years' old	Firm-level data Vietnam enterprise survey
Diversity (log value)	Inverse Herfindahl index of Labour	Firm-level data Vietnam enterprise survey
Test for IVs		
F value of weak indentification KP test	Kleibergen-Paap rk Wald F statistic	Baum et al. (2007)
P-value of underidentification LM statistics	LM is Lagrange multiplier test statistic	Baum et al. (2007)

3.B Acknowledgement

I would like to specially thank Joseph Francois, Miriam Manchin, Manfred Elsig, the participants from PhD colloquium in 2016 at the WTI, and the participants in international Rimini conference in Waterlooo in 2016 for their invaluable comments on this paper. I also thank Stephan Kyburz for sharing the boundary data of provinces in Vietnam. Nevertheless, the paper does not necessarily reflect the views and ideas from these acknowledged people and organizations.

A previous version of this chapter is the second part in the SECO/WTI working paper (No.9/2016) titled “Ease of doing business reforms in Vietnam: implication for total factor productivity in manufacturing industries" by Huong Quynh Nguyen.

4 Industrial clusters and firm performance

Joint with Joseph Francois

4.1 Introduction

Early in the book "Principle of Economics", Marshall (1920) coined the term industrial district in the phrase "thickly peopled industrial district". Greenstone et al. (2010) noted that a geographical concentration of industries exists in most countries. Importantly, a cluster approach to industrial policy may better support a firm's "broad-based competitiveness" because it allows for broad-based policies that would not usually an individual firm, though collectively individual firms benefit (Ketels et al., 2010). With an industrial cluster (e.g. Silicon Valley), firms may take advantage of technology diffusion when locating closer to other firms in the same industry (Ellison and Glaeser, 1997).

Vietnam offers an interesting case for the study of industrial clusters and firm performance, as it is a transition economy in Asia which has recently made great efforts to industrialize its economy.[1] Indeed Newman et al. (2014) looked at the improvement in firm-level productivity by industrial clusters according to geographical administrative units of Vietnam during the period 2002-2007,[2] while Howard et al. (2015) have examined industrial co-agglomeration in Vietnam in cross-section with data for 2007.[3] Theoretically, Duranton and Puga (2004) have noted that agglomeration economies and congestion costs influence firms when firms cluster and shape cities. However, less attention has been focused on the two-sided effects of industry-specific district clusters in a developing country such as Vietnam.

In accordance with other supporting manufacturing industry policies, the Vietnamese government has implemented zoning policies to enhance infrastructure in specific geo-

[1] Historically, village-based industrial clusters were existing in crafts and traditional products in the country, such as well-known guild streets "Hang" in Hanoi Old Quater, or Hadong silk village and Battrang pottery village. Modern clusters of light industries concentrate more in Hochiminh city and its satellite towns (Southern area of the country) while capital intensive hubs locate in higher density in Hanoi and its neighbour provinces (Ketels et al., 2010).

[2] by district and province.

[3] In this study, the authors excluded small firms less than 30 workers.

graphical industrial clusters (economic zones, EZs). According to Arnold (2015), there were 292 Industrial Zones, and 3 Export Processing Zones in Vietnam by 2014. These EZs created jobs for 2.28 million employees, and attracted 85.993 billion USD from foreign direct investment in 2014.[4] Several studies have also examined the economic implications of Zoning policies for developing countries.[5] Nevertheless, there is little empirical evidence given for a connection between zoning policies and industrial cluster development.

Our focus in this chapter is on a dynamic spatial analysis with firm level data data for 2005 to 2010 of whether and to what extent industry-district clusters influence the performance of firms clustering in the same industrial district in Vietnam. Our spatial approach includes unique data on the distance of individual firms from intra-industry and cross-industry clusters when identifying agglomeration effects.[6] Additionally, as Ketels et al. (2010) have claimed that zoning policies in Vietnam until 2010 had not been "oriented towards clusters", we test whether policies have nonetheless supported the formation of industrial clusters in the country during the 2005-2010 period.

Our contributions are: (i) new measurement of industrial district cluster of labour in Vietnam (inter-industry and cross-industry clusters); (ii) novel findings for not only agglomeration effects but also congestion effects on performance of firms in Vietnam; (iii) unique utilization of road map data for measuring travel distance between firms and clusters. The utilization of transport route data is particularly relevant for emerging markets where road networks are less developed, with driving distances more binding as geographic barriers; (iv) we develop a Logit framework for evaluating cluster orientation of zoning policies in Vietnam.

We develop a unique and novel firm-level data set of manufacturing industries in Vietnam for the period between 2005 and 2010. It features firm-level data drawn from the yearly enterprise survey of the General statistic office of Vietnam, merged with other relevant datasets of Vietnam, such as: administrative data at district level (from GIS data), industry-level data of imported intermediate to Vietnam (from UNComtrade), and economic zones data (from the Vietnam Ministry of Investment and Planning). These data are further merged with geographic data (travel distance by car from map data layers) and processed to generate road network distances between firms and clusters.

To investigate the agglomeration and congestion effects, we apply the GMM framework

[4]See Figure 4.A.1 for the map of Industrial Zones in Vietnam.

[5]See for example:The World Bank (2011), Wang (2013), Arnold (2015), and Francois and Davies (2015). For the case of Chinese Special Economic Zones, Wang (2013) indicated the economic impact of these zones on local economy.

[6]For example, see Fujita and Thisse (2013) for the theoretical model explaining for agglomeration equilibrium in which shipping expense is costly.

proposed by Roodman (2009). The results show that on average agglomeration effects (positive) are significant in low-density clusters, while the congestion effects (negative) are shown in high-density clusters. However, mixed results are indicated for industries at different levels of technology intensity. Second, because system GMM is not suitable for short time series with a small number of firms within each industries, we apply fractional polynomial regression (Royston and Altman, 1994), and find a bell-shaped curve for TFP trends when density of cluster increases in twelve of twenty industries.[7] The bell-shaped also indicates the dominance of agglomeration effects for TFP in low-density clusters, and the congestion effects are stronger in high-density clusters. Last but not least, results of the Logit regression reveal that zoning policies stimulate higher possibility of forming industrial district clusters in most of the manufacturing industries in Vietnam.

The content of this paper is arranged as follows. Section 4.2 summarizes relevant studies on agglomeration effects. Section 4.3 proposes the methodology used to define an industry-cluster and our specification. Section 4.4 describes our data. Section 4.5 and 4.6 present and discuss our empirical results. Finally, Section 4.7 draws a conclusion and gives implications for our studies.

4.2 Literature review

In this section we summarize two strands of literature related to the questions addressed here.

4.2.1 Agglomeration effects from industrial cluster

Recent studies have shown results explaining geographic pattern of industries, how industries co-agglomerated, and agglomeration effects.[8] In terms of firm heterogeneity advantages, firms may benefit from their location (Clark et al., 2003)[Porter.M, chapter 3]. The natural advantages of a location and spatial distribution were also analysed by Glaeser and Kerr (2009).

Redding (2010) has emphasized "location of economic agents relative to one another in space" which is the "second nature geography" in new economic geography literature (beginning with the study of (Krugman, 1990)), and referred to the focus of neoclassical economic geography literature on the "first nature geography" (climate, topology and re-

[7] Other industries observed fast increasing then flat trend of productivity.

[8] See examples respectively in: Duranton and Overman (2005), Ellison and Glaeser (1999),Ellison et al. (2010), Combes et al. (2012) and Combes and Gobillon (2015).

soure endownments). Redding (2010) mentioned that firms locate closer to large markets because of love of variety, increasing return to scale and transport cost, and indicated evidence from rich literature showed that wages and production are in connection to market access (such as: proximity to sources of market demand and to sources of supply of inputs, etc.). However, according to Redding (2010), causality, confounding factors of the link between wages and market access, and the "discrimination of agglomeration forces" have not yet been paid due attention; Apart from that, Redding (2010) also reviewed studies that measured the concentration of economic activities (in agglomeration of industry or geographical concentration), or the effects generated by reallocation of firms and workers. The author concluded that there are still lack of empirical evidence in multiple level geographic distribution of economic activities and spatial scales of agglomeration forces.

Micro-data sets with geographic information have been widely used in literature to analyse the pattern of local industries, and draw crucial implications for regional industry policies in economic development. Several important studies have paid attention to industry-specific clusters. For example, Ellison and Glaeser (1997) analysed spillover effects from the localization of individual U.S manufacturing industries. They found that the movement of labour force could bring accumulated experience and skills across enterprises. Specifically, face-to-face communication and knowledge sharing in a high-density cluster of labour force with skills in a common working field would probably stimulate productivity. Duranton and Puga (2004) referred to the term "urban specialization" or "localisation economies" as clusters of firms in the same industry within close proximity (for more than one industry). The authors noted that in the study of Glaeser (1999), proximity plays important roles in stimulating knowledge transfer from elite workers to less skilful workers. In a study of localisation economies, Callois (2008) showed a bell-shaped relation between proximity among small firms in similar sectors and their performance.

In general, agglomeration effects have been decomposed into localization and urbanization externalities in the literature. Localization externalities from specialization (within industry) are termed Marshallian externalities (Marshall, 1890), while urbanization externalities induced from production diversity (across industries) are defined as Jacobian externalities (Jacobs, 1969).[9] Mixed empirical results are provided for Marshallian and

[9]For a complete survey of literature in Mashallian and Jacobian externalities, see Crafts and Klein (2015) and Fazio and Maltese (2015) for examples. In particular, Crafts and Klein (2015) calculated Marshallian externalities equivalent to L_{cs}/L_c (where L_{cs} is the employment of sector s in city c, and L_c total employment in city c), and measure the Jacobian externalities as a Hirschman-Herfindahl index for the case study of the United States. Similar calculations of these externalities were proposed by Fazio and Maltese (2015) for the case study of Italian firms. The authors modified the index of Marshallian externalities such that is it faced by each firm, and also referred to Porterian externalities which were defined as local competition. Fazio and Maltese (2015) also provided literature which proposed different index for Marshallian and Jacobian externalities.

Jacobian externalities in the recent literature. For example,Crafts and Klein (2015) found the positively strong Marshallian externalities but weak Jacobian externaltieis on labour productivity in the United States from 1880 to 1930. They indicated the positive effects of Jacobian externalities only for large cities while the negative effects of diversification were found in small cities. In the study of Fazio and Maltese (2015), Marshallian externalities positively influence TFP level of small and medium Italian firms but not the TFP growth, however, Jacobian externalities are the driving factor of growth in firm-level TFP. However, empirical studies in Jacobian externalities across industries showed little evidence of direct connections among industries. Current literature of foreign direct investment (FDI) spillovers and international trade has exploited linkages among industries from input-output table to demonstrate backward and forward linked industries when calculating spillovers Javorcik (2004) and weighted tariff (see (Topalova and Khandelwal, 2011), (Amiti and Davis, 2012), and (Brandt et al., 2017)). In our paper, we develop further this approach from the literature of trade and FDI, and exploit inter-country input-output table (OECD, version 2015) to indicate the importance of transport road by car between firms and forward and backward linked industry clusters for the case study of Vietnam. Hence, we investigate the agglomeration effects generated through these linkages in inter-industry clusters, apart from intra-industry clusters of the country.

Analysing impact of agglomeration and dispersion forces is important because they influence the internal structure of the city, and the analysis may give useful information for implementing policies in transportation investment, urban development and taxation (Ahlfeldt et al., 2015). Duranton and Puga (2004) have summarized and developed theoretical frameworks for how agglomeration effects influence on firms that shape urban areas. Three crucial channels of the agglomeration effects are: "sharing", "matching", and "learning". "Sharing" means indivisible facilities, wide variety of inputs, gains from specialisation, and risks. "Matching" is relevant to qualities of matches and matching probabilities. "Learning" includes generating, diffusing and accumulating knowledge. Importantly, the authors indicated that urban cities can generate two-sided effects: agglomeration economies and congestion costs. Desmet and Rossi-Hansberg (2009) introduced two forces of geographic dispersion which are technology diffusion and congestion costs. Their paper also discussed knowledge transfer, pecuniary externalities, labor market pooling as main forces of geographic concentration. [10] Ahlfeldt et al. (2015) estimated agglomeration and dispersion forces in a quantitative theoretical framework with a gravity equation for

[10] Applying theoretical framework of Desmet and Rossi-Hansberg (2009),Desmet et al. (2015) analysed the geographical pattern of manufacturing and services in India compared to China and the United States. Their conclusion is that the manufacturing in India reached the highest labour productivity in "medium-density" places, which was different from patterns of manufacturing in China and the United States.

bilateral commuting flows. This frame work can be applied to investigate the influence of intervention, such as change in transport technology, on the economic activities within city. Interestingly, minimum travel time measured by public transport (in 1936, because travel by car was not popular in that year) and by car in Berlin (1986 and 2006) were calculated by ArcGIS for bilateral connection between city blocks (15,937 x 15,937) in Berlin (Ahlfeldt et al., 2015). Heblich et al. (2018) proposed a novel structural estimation in a gravity equation for commuting flows, and aimed to evaluate the impact of railway construction. Data from commuting flows in London in 1921 by steam railway is analysed by Heblich et al. (2018) in conjunction with historical data of population, values of land and transport network in London (back to early nineteenth century). The findings from Heblich et al. (2018) indicated the "first large-scale separation of workplace and residence" by steam railway travellers in London because the steam railway could reduce travel time in comparison to slower travel times by human or horse power. Evaluating place-based policies by accounting for the influence of these policies on location choice of firms, Gaubert (2018) showed that subsidizing policies for smaller cities in France induce negative aggregate effects and do not decrease the spatial disparity. Gaubert (2018) decomposed the firm's productivity advantage in dense areas into advantage of the agglomeration externalities and the endogenous of sorting into the areas of higher productive firms, and found that firm sorting generates nearly half of "elasticity of productivity to density".

In existing literature, driving distance or travel time can be also calculated from database of Google Application Programming Interfaces (APIs) (Ozimek and Miles, 2011), or MapQuest APIs (Voorheis, 2015), however, the use of Google APIs and MapQuest is now restricted by the suppliers (Huber and Rust, 2015). Instead of using these data, Huber and Rust (2015) propose a new method using non-commercial data from OpenStreetMap. The advantage of this method is that it can calculate the minimum travel distance (or travel time) by car from the database of OpenStreetMap without cost. We apply the method by Huber and Rust (2015) in our paper. However, the caveat of using this method is that geographical data from OpenStreetMap is updated everyday (year 2016) while our micro-data covers the years from 2005 to 2010. Hence, it can hardly control for the improvements in road infrastructure in Vietnam throughout the years, and the difference in travel time due to peak or off-peak hours using either geodetic distance or real route distance (other databases of map layers may also have this disadvantage).

Table 4.1 summarizes selected studies relevant to agglomeration effects of industrial clusters (horizontal and vertical). Few evidence from the literature of industrial clusters have investigated the roles of geographical proximity between firms and industrial clusters (horizontal as well as vertical clusters) in agglomeration effects.

Table 4.1: Summary of Selected Literature in Agglomeration Economies

Author (Year)	Agglomeration	Collocation	Horizontal	Vertical	Proximity	Key results	Country
Glaeser (1999)			X		X	Distance stimulates knowledge transfer	
Callois (2008)			X		X	Bell-shaped relation: proximity & productivity	
Combes et al. (2012)	X					Firms more productive in larger cities	France
Newman et al. (2014)			X	X		Clusters enhance productivity	Vietnam
Desmet et al. (2015)	X					Productivity slowdown in medium-density areas	India
Howard et al. (2015)	X	X				Firm location based collocation index	Vietnam
Howard (2017)					X	Proximity less important than social network	Vietnam

4.2.2 Zoning policies evaluation

Most recent studies of Enterprize zones (EZs) in Europe and the US focused on evaluating the effects of EZs on employment, however, they gave mixed results (Neumark and Simpson, 2015). For instance, Neumark and Kolko (2010) mapped EZs in California (the US) and its "ring" boundaries, and found that EZs did not reach the primary target in increasing employment. In opposition, Mayer et al. (2017) applied method of Difference in Difference and Logit estimation to indicate positiveness effects of EZs policies in municipalities in France, and concluded that this program can increase the probability that firms locate in EZs. Studies of EZs in developing countries such as Asian focused on FDI attraction. As noted by Arnold (2015), FDI attraction in EZs is relevant to the investment-driven stage of one country's development that was first introduced by Porter (1990). Most recently Wong and Buba (2017) evaluate the growth in nightlights in the zones of one country in comparison to the growth in nightlights emitted in the whole country to evaluate the performance of special economic zones across countries, and find mixed results. The paper also indicates the positive externalities of special economic zones in neighbour areas (with distance less than 50 km to the zones) for the case study of Vietnam.

Table 4.2 summarizes main results of selected studies that evaluate the effects of zoning policies. Existing studies did not pay due attention to empirical evidence whether place-based policies in EZs encourage the establishment of industrial clusters.

Table 4.2: Summary of Literature in Industrial Cluster

Author (Year)	Key results	Country
Neumark and Kolko (2010)	SEZs do not increase employment	California (the US)
Wang (2013)	SEZs increase FDI	China
Neumark and Simpson (2015)	Summarize mixed results of SEZs effects on employment	EU & the US
Mayer et al. (2017)	Positive effects of urban enterprise zones on employment	France
Wong and Buba (2017)	Positive effects of special economic zones on neighbour areas	Vietnam

4.3 Methodology

In this section we first summarize the theoretical framework of Duranton and Puga (2004) that we apply in this study. Theoretically, this paper follows micro foundation theory of urban agglomeration in the link with sharing, matching and learning channels (Duranton and Puga, 2004). Duranton and Puga (2004) developed existing literature on agglomeration and suggested theoretical model for agglomeration in the context of imperfect competition with the consideration of proximity. Instead of examining the agglomeration at city level as sketched out in Duranton and Puga (2004), we extend our research at district level at which industrial cluster is investigated in Vietnam.

Second, we evaluate the economic impact of an industry-district cluster on the performance of manufacturing firms in 2 steps.[11] (i) We estimate firm-level total factor productivity (TFP) using instrumental variables estimator initiated by Wooldridge (2009a) and Petrin and Levinsohn (2012). (ii) We investigate the economic impact of the industrial cluster on firm performance by regressing firm-level TFP on dummy variable indicating whether a firm locating in an industrial district cluster or not. In addition, firm's distance to the nearest cluster (horizontal and vertical industries), the presence of foreign firms, and difference in firm ownership (FDI firms or domestic firms) are also controlled for in the empirical specification. Spatial analysis with the investigation of travel roads for industrial clusters makes sense especially for developing countries like Vietnam where 25% of surveyed firms ranked transport "severe or major constraint" in 2005 (The World Bank, 2006).

Finally, we propose a logit model to investigate the role of zoning policies in the formation of industry-district clusters for each industry.

[11] See Combes and Gobillon (2015) for the two-steps approach.

4.3.1 Concept of Industrial Clusters

Existing literature contains a large number of studies examining the density of employment. For example, Giuliano and Small (1991) investigated the employment sub-center in Los Angeles by exploiting information on number of employees, and setting boundary for distance. Giuliano et al. (2005) followed the concept of employment clusters which was first introduced by Giuliano and Small (1991). The employment center was defined as a cluster if it reached a "minimum employment density of D, and together containing total employment of at least E" (Giuliano et al. (2005), p.21). duranton and Puga (2000) calculated the index for specialization and diversification of cities using the ratio of employment shares in a city and in an industry. Most recently, Delgado et al. (2015) documented an algorithm of industry cluster as a "group of closely related industries" with information of industry co-location, input-output linkages, and labour similarity.

In this study, to define a district is an industrial district cluster, we construct an index of a specific industry-district cluster of labour.[12]

We specify whether a district is a high-density labour cluster of a specific 2-digit industry by constructing a district-industry index. This index indicates the labour density of industry j in district d at year t (See Equation 4.1).

$$I_{djt} = \frac{L_{djt}}{\sum_j L_{jt}} \tag{4.1}$$

Where:

L_{djt} is the total number of workers in district d that worked in industry j in year t.

$\sum jL_{jt}$ is the total number of workers in industry j in year t.

This index focuses on the density of workers in a specific manufacturing industry in a district. Therefore, these clusters are hubs of firms that specialize in a particular manufacturing industry. From that measurement, we define an industrial cluster, which has a high labour density district-industry, as follows: From the distribution of the index within one industry, we can refer to a district as a cluster of a specific industry if its I_{djt} is higher than the mean value of the index in the industry j.

[12]This index is similar to the Marshallian index which is calculated at city level in the study of Crafts and Klein (2015).

$$Cluster(Mean) : Cluster_{djt} = \begin{cases} 1, & \text{if } I_{djt} > \mu_{jt} \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

Where μ_{jt} is the mean value of index I_{djt} in the industry j at year t.

Alternatively, increasing the threshold of the index to a higher density of labour, we define a district as a cluster in Equation 4.3:

$$Cluster(SD) : Cluster_{djt} = \begin{cases} 1, & \text{if } I_{djt} > \mu_{jt} + \sigma_{jt} \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

Where:

μ_{jt} is the mean value of index I_{djt} in the industry j at year t. σ_{jt} is the standard deviation of index I_{djt} from μ_{jt} in the industry j at year t. This concept defines an industrial district with higher density.

4.3.2 Baseline Specification

In this section, we first present our approach to evaluate the impact of industry clusters on firm performance using system generalized method of moments (GMM) introduced by Blundell and Bond (1998). We then propose a logit model to investigate the probability that zoning policies possibly play an a key role in forming industrial clusters.

Industrial Clusters and Firm Performance

To estimate the effects of the industry-district clusters on firm's TFP, our specification is presented as follows :

$$\begin{aligned} Y_{idj,t} &= \alpha_0 + Y_{idj,t-1} + \alpha_{clus} CLUSTER_{dj,t} + \alpha_{near} NEAREST_{idj,t} \\ &\quad + \alpha_{fdi} FDISHARE_{dj,t} + \alpha_{imp} INTERMEDIATE_{j,t} \\ &\quad + \alpha_{fdidum} FDIDUM_{idj,t} + \alpha_t T_i + \epsilon_{idj,t} \end{aligned} \tag{4.4}$$

Where:

$Y_{idj,t}$: Log of TFP of firm i in district d and industry j at time t. We estimate TFP by using instrumental method introduced by Wooldridge (2009a) and Petrin and Levinsohn (2012).[13]

$CLUSTER_{dj}$: dummy variable = 1 if firm i is in the cluster of industry j within district d.

$NEAREST_i$: the nearest route distance by car from centroid of district d to centroid of the nearest industrial district cluster c in industry j is calculated using travel route data for Vietnam downloaded from www.openstreetmap.org (on October 11th 2016) by applying the method of (Huber and Rust, 2015).

We assume that inputs and outputs are transported by car as freight by car transport counted for dominant parts of total freight in Vietnam.[14] Real travel time may be investigated as an alternative proxy for distance for further research. Nevertheless, due to the limitation in measurement of travel time, we do not have variables such as travel time by different transportation, or at different time, hence the variable distance of car route is currently chosen at our best.

$FDISHARE_{dj}$: % share of workers in foreign firms in industry j within district d. This variable is used as a proxy for presence of foreign firms in industry j within district d.[15]

$INTERMEDIATE_j$: value of intermediate inputs in industry j imported to Vietnam (in log). This variable is used to isolate the impact of import on the total factor productivity of firm i. It may explain the integration level of each industry to the global economies when the local market has been opened up, imported intermediate inputs might bring high technology but also can create competition.

$FDIDUM_{idj}$: dummy variable = 1 if the firm is a foreign invested enterprise, and the dummy = 0, otherwise. The firm can be either a 100% foreign invested firm, or a joint venture owned by foreign investors and domestic investors.

T_i : Time fixed effects of firm i.

---

[13]In Newman et al. (2015), the author use similar approach for the Vietnamese micro-data from 2009 to 2012, however, industry classifications were different from our study. See also Nguyen (2017a) for the similar application of Wooldridge (2009a) and Petrin and Levinsohn (2012) in calculation of TFP in manufacturing of Vietnam from 2000-2010.

[14]According to (Union, 2011), freight by railways was merely 2% of total freight in Vietnam.

[15]We follow Aitken and Harrison (1999) to consider the presence of foreign firms in a specific industry. However, as we do not observe the foreign equity for each firm, we only use the output share of foreign invested firms in general similar to Newman et al. (2015). We take a further step to localize the presence of foreign firms in each industry to district boundary.

In Equation 4.4, we estimate the impact of agglomeration effects generated from the industry-district cluster by controlling for whether a firm i located inside the cluster ($CLUS_{dj}$), and the distance of its location's district d to the nearest cluster k in the same industry j ($NEAR_{dkj}$).

Apart from that, the role of firm's foreign ownership is controlled by the dummy variable $FDIDUM_{idj}$. Importantly, TFP_i in the previous year is also considered in the baseline specification as in our assumption TFP follows the Markov process. Time fixed effects and industry fixed effects are included.

Firm fixed effect is an issue when unobservable time-invariant factors potentially correlate with control variables in Equation 4.4. The solution for fixed effects is to first difference both sides of the Equation, however this method is addressed to face endogeneity problem: when first differencing Equation 4.4 in which lag of dependent variable $Y_{idj,t-1}$ was included as one of control variables, endogeneity can exist. This happens because the first difference of the lagged dependent variable includes the $\epsilon_{idj,t-1}$ (see Equation 4.5).

$$E[Y_{idj,t-1} - Y_{idj,t-2} | \epsilon_{idj,t} - \epsilon_{idj,t-1}] \neq 0 \tag{4.5}$$

To overcome the endogeneity issue for the dynamic model in Equation 4.4 using a "small-T, large-N" panel data, we apply system generalized method of moment (GMM) following Blundell and Bond (1998), Blundell and Bond (2000), and using xtabond2 command in Stata (Roodman, 2009).

Zoning policies and Formation of Clusters

In this sub-section, we examine the possibility that a district could become an industry-specific cluster district of labour when Zoning policies were implemented in the district. Obviously, zoning areas give firms more favourable incentives that attracts firms to locate in the zones for taking advantages of labour pooling, sharing facilitations and reducing transport cost of inputs, technologies and ideas. Wong and Buba (2017) summarize backward linkages (suppliers and firms collocate in close proximity, Ottaviano and Puga 1998) and forward linkages (demand channel, Marshall 1920, Ottaviano and Puga 1998, and Farole (2011)) as two crucial channels that EZs could stimulate the clustering of firms for efficiency and interdependency in a "self-enforcing agglomeration mechanism".

We exploit the Zone dummy variable as the proxy for the treatments of Zoning policies, i.e: the available amenities of one district such as advanced infrastructure, closer distance to transport system, favourable tax policies for workers, etc.

In addition, density of a district is controlled by the log value of the number of manufacturing firms within the area.

We proposed that the logit model is used to predict the probability that the district is an industrial cluster (as being defined for two cases in Equation 4.2 and 4.3 respectively) under the Zoning policies treatment.

The estimation is conducted for each industry. From this specification, we could evaluate the sign of margin effects brought by the Zoning policies. Positive sign of the margin effects can be interpreted that Zoning policies led to the higher probability of a district to be an industry specific cluster of labour. Because the labour density varies differently among industries, baseline regressions are applied for each industry.

Logit model (1)

$$Pr[Cluster_{djt} = 1|X_{t-1}] = Pr[I_{djt} > \mu_{jt}] \tag{4.6}$$

Logit model (2)

$$Pr[Cluster_{djt} = 1|X_{t-1}] = Pr[I_{djt} > \mu_{jt} + \sigma_{jt}] \tag{4.7}$$

Where:

X_{t-1} are control variables including:

$ZONE_{dj,t-1}$: dummy variable = 1 if district d has SEZ(s) locating within its boundary in year $t-1$.

$NUM_{d,t-1}$: log number of firms in district d in year $t-1$. This variable is proxied for the size of the district d.

$YEAR_t$: Year effects. REG_p: Region fixed effects is added to absorb the fixed effects of each region such as natural and cultural conditions due to its geographical location. It is important to note that the master plan of zoning policies is in accordance with the regional economic development plan (See the summary of target industries in EZs by regions in Appendix 4.A.3). We also apply an interaction term between year effects and regions fixed effects to control for the changes in policies across region through years.

I_{djt}: Industry specific labour density index as being defined in Equation 4.1.

μ_{jt}: the mean value of index I_{djt} in the industry j at year t.

σ_{jt}: the standard deviation of index I_{djt} from μ_{jt} in the industry j at year t.

Theoretically, endogeneity may happen due to simultaneity when: $EX_{jt}\sigma_{jt} \neq 0$. This may happen if the government decides to implement zoning policies in a specific location because the location has some advantages such as labour pooling and capacity of technologies (can be proxied through number of firms in the area), land for manufacturing, infrastructure development or industrial development strategies of one region (yearly changes over regions). To prevent the endogeneity issue, independent variables (in the Logit model) are lagged one year , then $EX_{j,t-1}\sigma_{jt} = 0$.

4.4 Data Description

This paper uses panel data of manufacturing firms that is compiled from the Vietnamese Enterprise Survey (2005-2010). This survey has been conducted annually by the General Statistics Office of Vietnam and includes legally registered firms or firms that still run their business by December 31st of each year. Annual aggregate description of this survey is presented in the Statistical Yearbook of Vietnam. Firm-level datasets (compiled from this survey) were used in several studies but in different time frame.[16] We exploit yearly firm-level information on values of capital stock, profit, cost, number of employees, output, location (district-level), and firm types. Manufacturing industry code is defined by the 2-digit Vietnamese Standard Industry Classification 1993 (VSIC 1993). This classification is provided by the Vietnamese General Statistics Office in the survey. Importantly, VISC 1993 is in line with ISIC Rev.3 provided by the United Nations. This allows us to merge the micro-data with other sectoral data of Vietnam organized in ISIC Rev3 such as data of trade volume from World Intergrated Trade Solution (WITS).[17] Manufacturing of tobacco, oil and petrol, and recycling were excluded due to too few observations. A list of the 20 industries used in this study is presented in the Appendix 4.A.2.

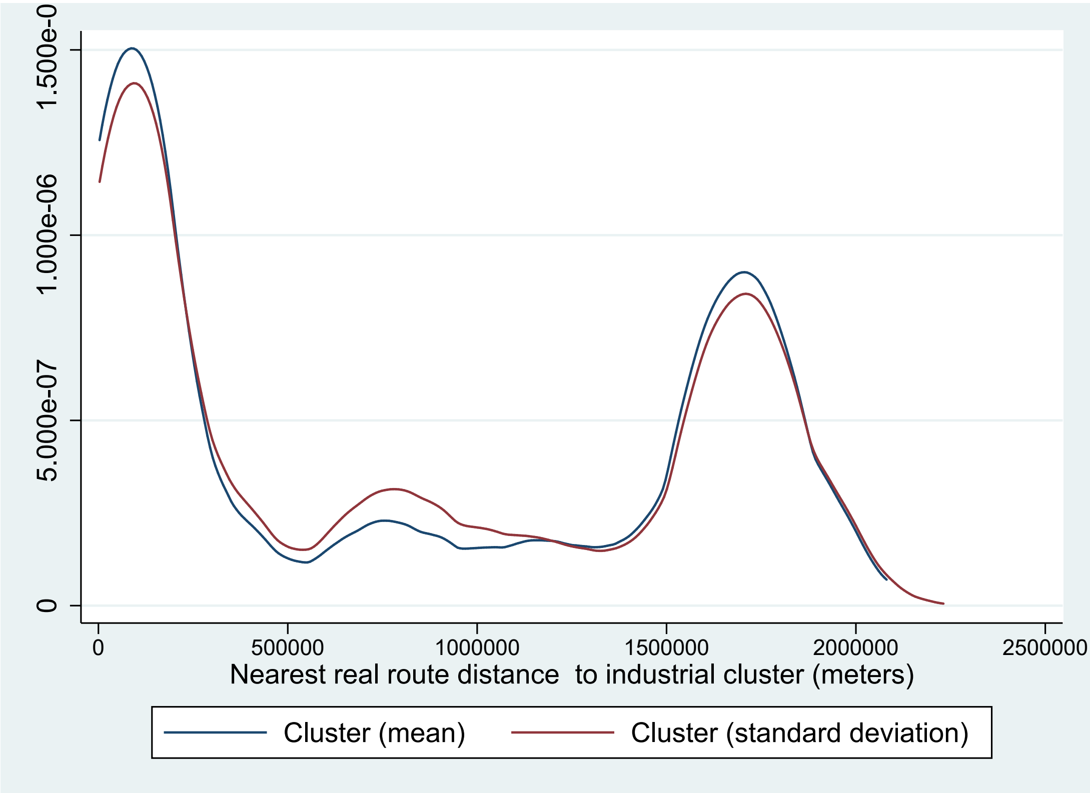
The data include 664 districts and 63 provinces.[18] Location at district-level in the

[16]For instance, Howard et al. (2015) studied the year 2007 with sample of manufacutures that hired more than 30 workers. Newman et al. (2015) investigated the survey with period from 2009-2012, also with firms hired more than 30 workers. In a study by Ha and Kiyota (2014), the research time frame was 2000-2009 with firms employed more than 20 workers.

[17]See Nguyen (2017a) and Nguyen (2016) for the description of the Vietnamese manufacturing micro-data.

[18]According to the General Statistics Office of Vietnam, there were 697 districts in the country until 2009 (for more details see http://www.gso.gov.vn/default.aspx?tabid=386&idmid=3&ItemID=8589). However, as there may be some districts where manufacturing firms are not established, the micro-data only records 674 districts in total. In 2008, under the regulations of the government, Hatay province was

Figure 4.1: Kernel Distribution of Real Route Distance From Each District to Its Nearest Cluster



Source: Authors' calculation using the real route data downloaded from website of Openstreetmap www.openstreetmap.org in October 2016.

micro-data is merged with Administrative boundary data of Vietnam from the Global Administrative Area data (www.gadm.org).[19] Due to issues of confidentiality, a firm's address was not provided. As our research aims at the localization of a 2-digit industry at district level, and the area of one district is quite small, we assume all firms in the same district have the same location which is specified by the longitude and latitude of the district's centroid. By doing this, we could position the specific location of an industrial district, and calculate the real route distance between the district and its nearest cluster as explained in section 4.3.2. The Kernel distributions of the real route distance from each district to its nearest clusters by definition of equation 4.2 and 4.3 are presented in Figure 4.1.

merged into Hanoi. In this research, we merged Hatay to Hanoi for the whole 2005-2010 time frame. 515 observations with missing information on districts were excluded.

[19]We specially thank Stephan Kyburz for providing us the administrative boundary data of Vietnam from global administrative area data (www.gadm.org). See Nguyen (2016) for more details of the combination of Vietnamese boundary data from the global administrative area data (www.gadm.org) and Vietnamese firm-level data.

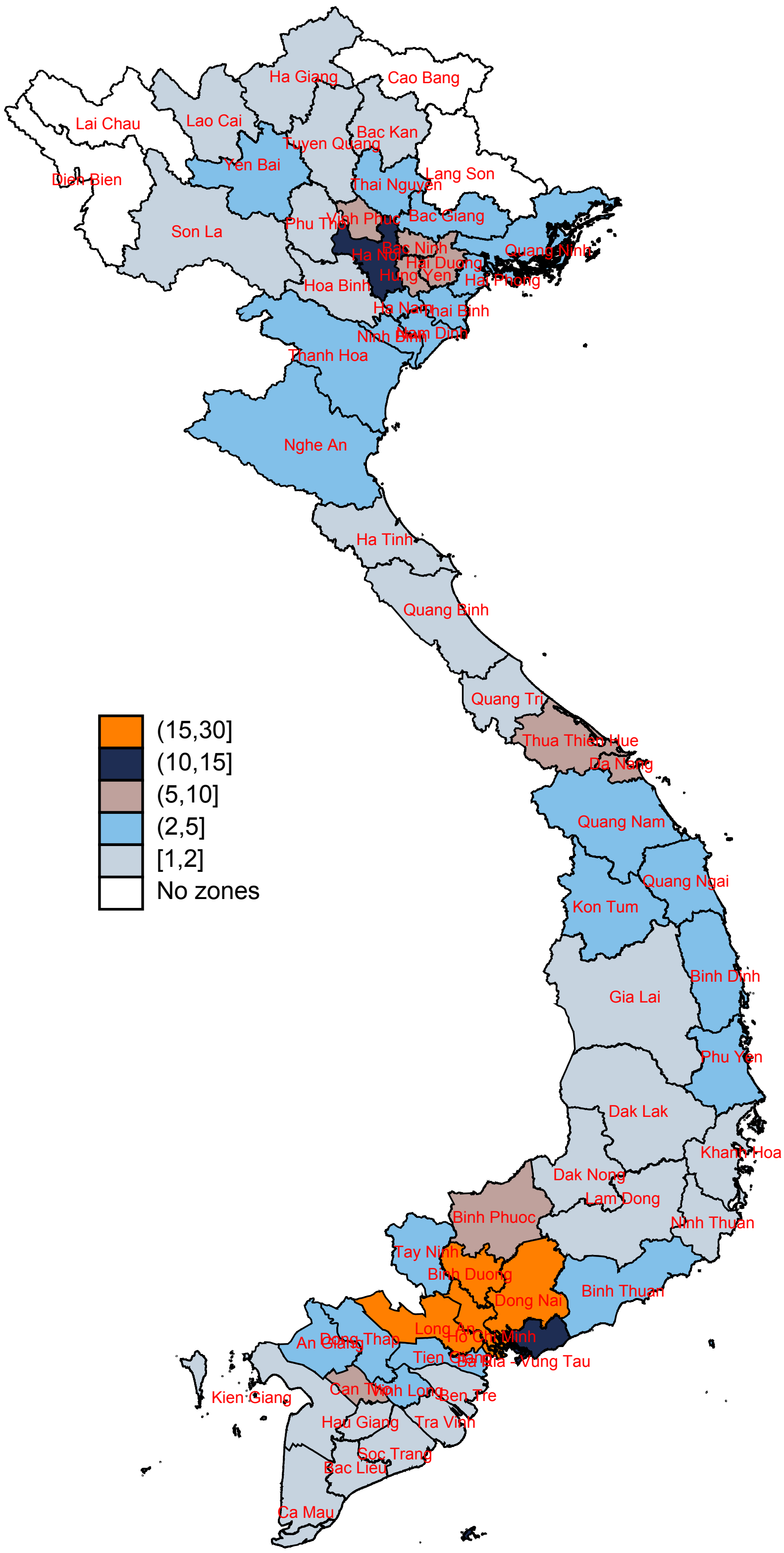
To georeference EZs in each district, we use an internal report on EZs in 2010 conducted by the Vietnamese Ministry of Investment and Planning in 2011. For zones whose location information is missing from the report, we look for the zone's district address from its website or from the website of its provincial EZs management board. Figure 4.2 visualizes the number of EZs by province in Vietnam.

The micro dataset is also merged with the sectoral-level trade dataset of Vietnam from the World Integrated Solutions (WITS) using 2-digit industrial classification ISIC Rev.3.[20] This information helps to evaluate the level of international integration of each manufacturing industry. In addition, data of year 2008 from the inter-country input-output table is also combined with micro-data.[21] The data provides the coefficients for domestics industrial linkages. We assumed the coefficients did not change for the research time frame (2005-2010).

[20]See more details of the dataset in: http://wits.worldbank.org/

[21]The data was downloaded from OECD's website: https://www.oecd.org/sti/ind/input-outputtablesedition2015accesstodata.htm

Figure 4.2: Number of EZs by province in Vietnam 2010



Source: Authors' visualization using number of EZs in Vietnam in 2010 (the Vietnamese Ministry of Investment and Planning, Annual report, 2011). Several Vietnamese islands, such as Hoangsa and Truongsa, were not mapped due to the limitation of the boundary data.

4.5 Empirical Result of Firm Performance and Industrial Clusters

4.5.1 Intra-industry

This section presents the empirical results for agglomeration effects of cluster on firm's performance by estimating impacts of the location of firm in a cluster (in the same industry) and distance to other clusters (in the same industry) on firm-level productivity. The results were drawn from the estimation using system GMM (Roodman, 2009). Robust results are shown for both types of clusters which are Cluster (Mean) - as defined in Equation 4.2, and Cluster(SD) - as defined in Equation 4.3. Additional information from fractional polynomial regression is provided for each industry.

Clusters of Intra-Industry Firms

For the case of lower density cluster (Mean), results of Model 1 and Model 3 in Table 4.3 indicates that on average locating in a cluster of firms producing the same goods stimulates a firm's productivity. In addition, an increase in the real route distance to the nearest cluster shows a negative impact. This implies that staying closer to the cluster, a firm performs better. These results confirm the agglomeration effects generated within and near a cluster of a specific industry which is the first evidence of "localisation economies" (Duranton and Puga, 2004) in Vietnam. The agglomeration effects within the same industry ("localisation economies") can be explained to be generated through "sharing", "matching", and "learning" mechanism as noted by Duranton and Puga (2004).

Regarding to higher density cluster (SD), the results in Model 2 and Model 4 also indicate the decaying effects of agglomeration from the nearby industry-district cluster on the firm-level TFP. However, locating inside a higher density labour cluster - Cluster (SD) show a negative impact on firm performance. This can be explained due to the "urban crowding" or "congestion" effects as discussed by Duranton and Puga (2004). We provide a robustness check for the results of agglomeration effects from the intra-industry clusters in Section 4.5.1 in which the fractional regression method is applied for each industry.

These findings for lower density cluster and higher density cluster in our research are novel and give more profound analysis in comparison to previous research in clustering in developing countries. For example: Newman et al. (2014) generally found positive productivity spillovers from clustering in Vietnamese manufacturing during 2002-2007. In another study for the case study of manufacturing clusters in Cambodia, Chhair and

Newman (2014) showed negative influence of clusters on the performance of formal firms which were explained due to the competition of firm within a cluster. Similarly to Newman et al. (2014), Chhair and Newman (2014) did not look at the level of density of the cluster, and the role of upstream and downstream clusters as well as distance between the firm and other clusters in the same industry.

Table 4.3: Intra-Industry Clusters and Firm Performance, 2005-2010

Firm-level Log(TFP)	Model 1	Model 2	Model 3	Model 4
Firm in industrial district cluster (Mean)	0.035+		0.054***	
	(0.018)		(0.016)	
Distance to the nearest cluster (Mean)	-0.020***			-0.028***
	(0.004)			(0.006)
Firm in industrial district cluster (SD)		-0.061+		-0.080+
		(0.035)		(0.042)
Distance to the nearest cluster (SD)		-0.052*	-0.024*	
		(0.025)	(0.011)	
Number of firms	55784	55784	55784	55784
Number of IVs	69	69	69	69
AR2-pvalue	0.142	0.151	0.155	0.145
Hansen-pvalue	0.312	0.497	0.579	0.471
R-square	0.927	0.922	0.929	0.921

Source: Micro data drawn from the Vietnam Enterprise Survey (2005-2010).
Notes: Results of panel data system GMM are reported. Year and industry effects are included. Robust Standard Errors (SE) are clustered within province (in parentheses). R-square is manually calculated by computing the square of covariance $COV(lnTFP\widehat{lnTFP})$. $\widehat{lnTFP}$ is predicted from OLS estimations of these models.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Impact By Level of Technology Intensity

We divide firms into 4 groups of technical intensity industries (ISIC Rev.3 introduced by OECD), and run the same regression. Regarding to low-density cluster, our results showed in Table 4.4 indicates that firms perform better if they locate inside an industry-district cluster of higher technology intensity industry. The distance to the nearest cluster in the same industry only generates decaying effects in low and medium-low technology intensity groups.

Table 4.4: Firm Performance and Industrial Cluster by Level of Technology Intensity Industries, Low-density Cluster

Firm-level Log(TFP)	Low	Medium-low	Medium-high	High
Firm in industrial cluster	0.036	0.032+	0.035*	0.032*
	(0.029)	(0.017)	(0.015)	(0.016)
Distance to the nearest cluster	-0.028***	-0.014**	0.007	0.001
	(0.004)	(0.005)	(0.012)	(0.003)
Number of firms	22145	14048	8918	10673
Number of IVs	65	65	65	65
AR2-pvalue	0.275	0.448	0.382	0.380
Hansen-pvalue	0.360	0.383	0.606	0.843
R-square	0.932	0.935	0.933	0.921

Source: Micro data drawn from the Vietnam Enterprise Survey (2005-2010).
Notes: Results of panel data system GMM are reported. Robust Standard Errors (SE) are clustered within Province (in parentheses). R-square is manually calculated by computing the square of covariance $COV(lnTFP, \widehat{lnTFP})$. $\widehat{lnTFP}$ is predicted from OLS estimations of these models.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Results for high-density clusters in Table 4.5 only show congestion effects in cluster when firms are in medium-low industries.

Table 4.5: Firm Performance and Industrial Cluster by Level of Technology Intensity Industries, High-Density Cluster

Firm-level Log(TFP)	Low	Medium-low	Medium-high	High
Firm in industrial district cluster	0.041+	-0.041**	0.005	-0.005
	(0.021)	(0.015)	(0.023)	(0.031)
Distance to the nearest cluster	-0.028	-0.004	-0.001	-0.013+
	(0.019)	(0.009)	(0.005)	(0.007)
Number of firms	22145	14048	8918	10673
Number of Ivs	65	65	65	65
AR2 - pvalue	0.243	0.500	0.350	0.400
Hansen - pvalue	0.443	0.411	0.544	0.888
R-square	0.936	0.936	0.932	0.921

Source: Micro data drawn from the Vietnam Enterprise Survey (2005-2010).
Notes: Results of panel data system GMM are reported. Robust Standard Errors (SE) are clustered within Province (in parentheses). R-square is manually calculated by computing the square of covariance $COV(lnTFP, \widehat{lnTFP})$. $\widehat{lnTFP}$ is predicted from OLS estimations of these models.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

Impact By Industry: A Fractional Polynomial Regression Approach

While GMM model (Roodman, 2009) takes an advantage of the dataset with large observations in short time length, it faces the hindrance of smaller number of observations for industry-specific estimation, and the assumption of linearity. To investigate a better trend in the nexus between predicted firm-level TFP $ln(TFP_{ijd})$ and the index of labour density in district d and industry j for each industry, we apply fractional polynomial regression. This method could indicate the best fit for the power of control variable instead of linearity assumption. The results in this section are also a robustness check for the results about agglomeration effects of intra-industry clusters in Section 4.5.1.

Equation 4.8 shows the concept of fractional polynomials following Royston and Altman (1994). Powers p_j is chosen from the set $\{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$, maximum dimension j is m=2 (See Equation 4.8). The best model has the lowest deviance.[22]

$$\beta_0 + \beta_j \sum_{j=1}^{m} X^{p_j} \tag{4.8}$$

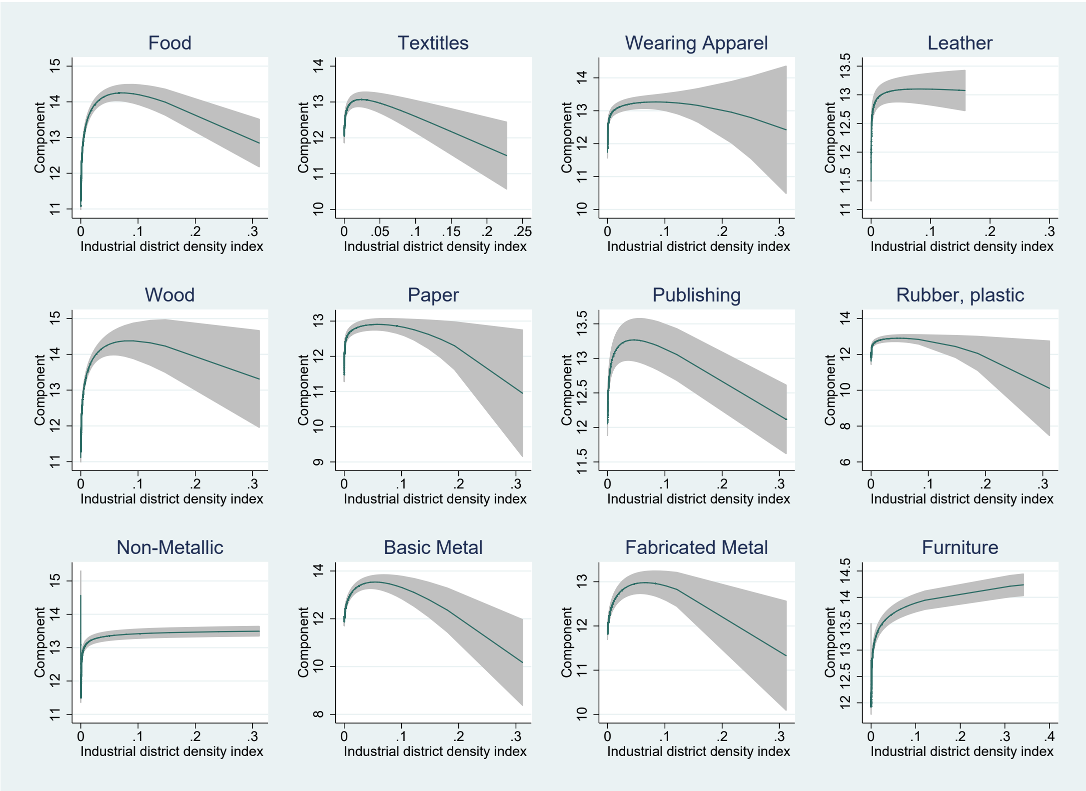
[22]We use fp syntax of Stata package.

Where:

$$X^{(p_j)} = \begin{cases} X^{p_j} & \text{if } p_j \neq 0 \\ ln(X) & \text{otherwise.} \end{cases} \tag{4.9}$$

The following graphs show the optimal curve shape of fitted value $lnTFP_{idj}$ plotted against lagged values of index I_{dj} in the best fit fractional polynomial model for each manufacturing industry in Vietnam (Index $I_{dj,t-1}$ is in Equation 4.1). The fractional polynomials regression controls for lagged values of cluster dummy, and year effects.

Figure 4.3: Fractional Polynomial Regression, Low Technology Intensity



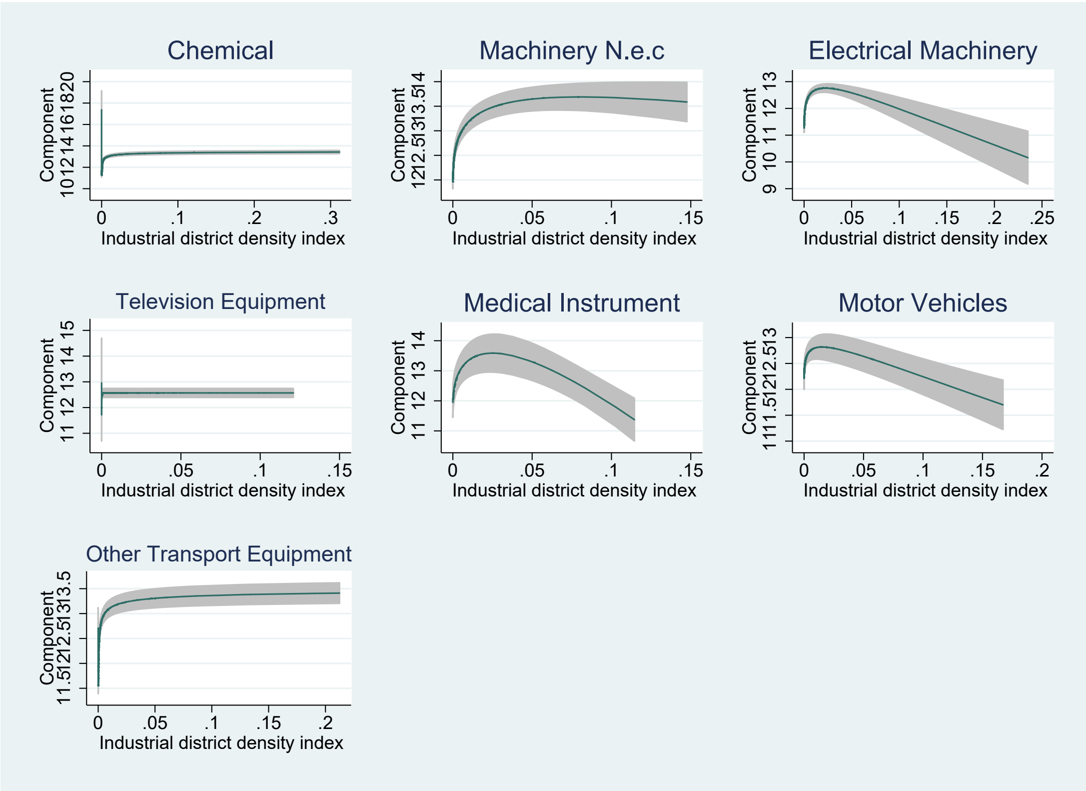
Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2005-2010. Fitted values of $ln(TFP_{idj,t})$ are plotted against index $I_{dj,t-1}$ in the best fit fractional polynomial model. The confidence level is 95%.

Agglomeration effects are shown as an increasing trend of fitted values $ln(TFP_{ijd})$ at low-density labour district, and congestion effects revealed when there is a decreasing trend of the values at higher density labour district are shown in bell-shaped graphs of twelve over twenty industries (See Figure 4.3 and Figure 4.4).[23] Graph of manufacture of Furniture shows an increasing trend while graphs of other industries such as: Leather products, Chemicals, Non-metallic mineral, Television equipments, and other transport equipment observe a flat trend of predicted ln(TFP) at higher values of the index.

Given that the curve shapes chosen are the best fit fractional polynomial model for each industry, the caveat of this approach is that this method cannot be conducted for model with fixed effects due to the complexity of calculation. Hence, we assume random effects when using this approach.

[23]Food, Textiles, Wearing Apparel, Wood products, Paper products, Rubber and plastics products , Basic metal, Fabricated metal products, Machinery and equipment, Electrical machinery and apparatus n.e.c, Medical instruments, and Motor vehicles.

Figure 4.4: Fractional Polynomial Regression, High Technology Intensity



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2005-2010. Fitted values of $ln(TFP_{idj,t})$ are plotted against index $I_{dj,t-1}$ in the best fit fractional polynomial model. The confidence level is 95%.

4.5.2 Cross-industry Clusters (Clusters of Upstream and Downstream Industries)

In this section, we develop the baseline specification for upstream industry and downstream industry clusters. Since the transport system is still underdeveloped, and high-speed transport means are not popular in Vietnam, distance to clusters is important to manufacturers because it influences the transport cost of materials, workers, final goods and knowledges (Marshall, 1920), as well as search cost and labour market matching (Krugman, 1991). First, the nearer the firm stays to its upstream clusters (hubs of intermediate inputs), it is less costly for manufacturers, hence firms gain more productivity when stay closer to upstream clusters. Second, distance between input suppliers to clusters of their downstream manufacturers (clients) is also important to productivity of suppliers as it influences the delivery cost.

Distance to upstream industry is proxied by the weighted total route distance by car to the nearest upstream industries. The proxy of distance to downstream industry is the weighted total route distance by car to the nearest downstream industries. Forward linkage and backward linkage are used as the weights respectively. These linkages are calculated from the domestic linkages in the inter-country input-output table in 2008 (OECD, edition 2015).[24] The coefficients are assumed to be constant during 2005 to 2010. The following table summarize statistics relevant to upstream and downstream clusters:

Table 4.6: Descriptive Statistics for Cross-industry Clusters

	Number of observations	Mean	Min	Max
Log(TFP)	99,677	12.649	9.510	17.864
Lower density cluster				
Log(distance to the nearest intra-industry clusters)	99,677	13.011	8.038	14.548
Log(distance to downstream clusters)	99,677	10.081	4.579	13.575
Log(distance to upstream clusters)	99,677	10.289	6.772	13.131
Higher density cluster				
Log(distance to the nearest intra-industry cluster)	99,677	13.076	8.075	14.618
Log(distance to downstream clusters)	99,677	9.951	1.366	13.758
Log(distance to upstream clusters)	99,677	10.234	0.999	13.248

Table 4.7 shows empirical results of two models that including variables controlling for distances to upstream and downstream industries. Model 1 indicates positive agglomeration effects on firms locating inside an industrial district cluster. It also gives significant evidence of decaying effects of geographical distances to the nearest cluster of the same industry, and weighted average distance to the nearest clusters of upstream industries. In Model 2 the variable “Market size" (log of total real output) was added to absorb the influence from the size of the district, and a dummy variable was also added to indicate whether a firm was foreign invested or not. We also look at the interaction of the foreign firm dummy and the location inside an industry dummy. We found that on average foreign firms performed at higher TFP growth rate than their domestic partners, however, when locating inside a cluster, foreign firms gained less than domestic firms in terms of TFP growth. This could be due to the lower level of integration of FDI firms in local manufacturing clusters in Vietnam when FDI firms in Vietnam depend more on inputs sourced from foreign markets (See similar comments in Ketels et al. (2010)). More investigations would be conducted to answer why the there is no decaying effects of distance to downstream cluster.

[24] http://www.oecd.org/trade/input-outputtables.htm

Apart from that, we consider the competition of import goods proxied by adding log value of imported intermediate inputs, however, the coefficient of this variable is not significant.

In general, results of decaying effects from the nearest cluster in the same industry and the nearest clusters of upstream industries are consistent in these two models.

Table 4.7: Cross - Industry Clusters and Firm Performance

Firm-level Log(TFP)	Model 1	Model 2
Firm in industrial district cluster	0.104**	-0.014
	(0.035)	(0.066)
Distance to the nearest cluster	-0.021*	-0.024**
	(0.010)	(0.008)
Distance to downstream cluster	0.108***	0.082***
	(0.027)	(0.024)
Distance to upstream cluster	-0.067***	-0.060***
	(0.014)	(0.015)
Market size of district		0.083**
		(0.027)
Foreign invested firm		0.504*
		(0.207)
Foreign invested firm in cluster		-0.297*
		(0.145)
Import of intermediate inputs		0.274
		(0.250)
Number of firms	55810	55810
Number of IVs	69	76
AR2 - pvalue	0.097	0.111
Hansen - pvalue	0.197	0.158
R-square	0.93	0.90
Robust standard error cluster	province	province

Source: Micro-data was drawn from the Vietnam Enterprise Survey (2005-2010).
Notes: Results of panel data system GMM are reported. Year and industry effects are included. Robust standard errors are clustered within province (in parentheses). R-square is manually calculated by computing the square of covariance $COV(lnTFP, \widehat{lnTFP})$. $\widehat{lnTFP}$ is predicted from OLS estimations of these models.
+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

4.6 Empirical Result of Zoning Policies and the Formation of Clusters

4.6.1 Marginal Effects of Zoning policies on Clusters Formation

To evaluate the role of EZs establishment in forming industry-district clusters, we estimate Logit model (1) in Equation 4.6 and Logit model (2) in Equation 4.7 for each industry (the two results of these Logit models are presented for robustness check). Table 4.8 indicates the significantly positive sign of the Zone dummy's coefficients for 14 manufacturing industries out of 20 industries (in both models); Industries that have significant results in the Logit model (1) (they are: manufacture of papers, other non-metallic mineral, medical instruments, motor vehicles, and furniture); and only manufacture of computing machineries has no significant results in both models .

The significance of positive coefficients shows that district in which SEZ(s) located within its boundary would gain higher possibility to form a specific industry labour cluster.

In short, results in Table 4.8 gives empirical evidence that the zoning policies were in effects to form the cluster of labour in most of the industries in Vietnam which is oposite to the claim of Ketels et al. (2010). These results are consistent with targeted industries set in the Master plan of industry development in EZs from 2005-2020 in Vietnam, such as: manufacture of chemical products, transport equipment, textile, and footwear (See: Appendix 4.A.3).

Table 4.8: Baseline Results of Logit Models by Industries, 2005-2010

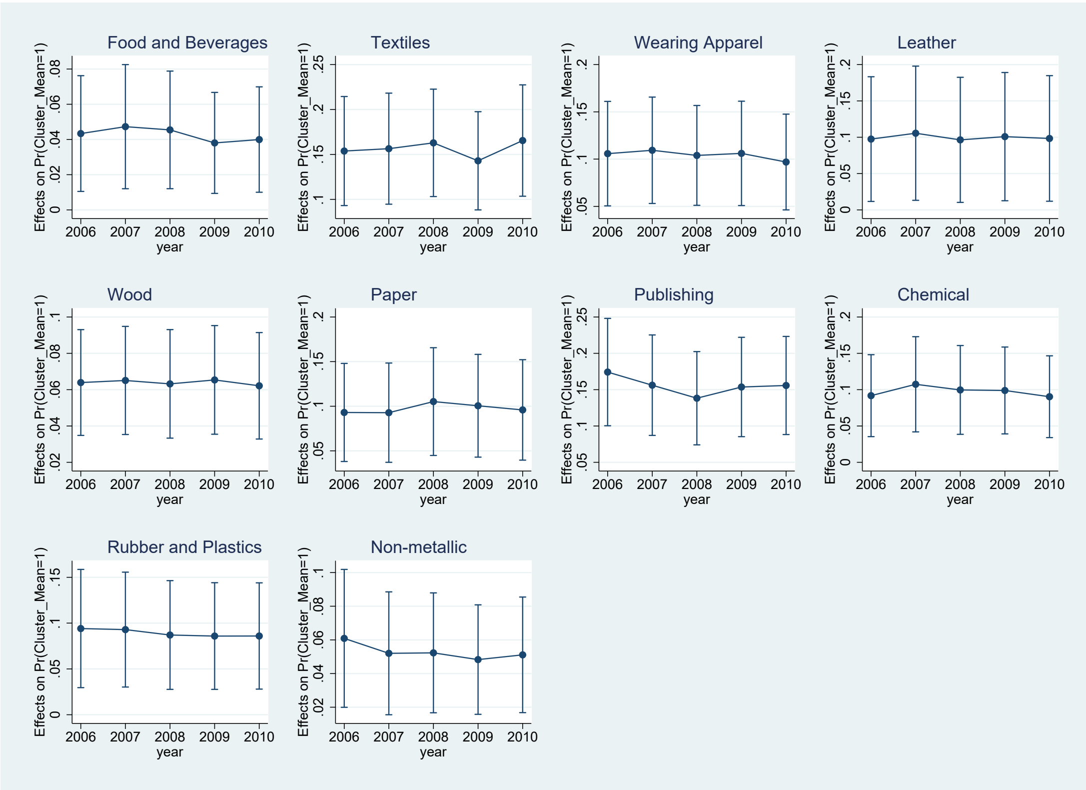
Industry	VSIC1993	Model	Margins	SE	N
Food & Beverages	15	(1)	0.043**	(0.016)	2446
		(2)	0.037***	(0.011)	1974
Textiles	17	(1)	0.156***	(0.029)	854
		(2)	0.079***	(0.023)	686
Wearing apparel	18	(1)	0.104***	(0.027)	1075
		(2)	0.084***	(0.023)	902
Leather	19	(1)	0.100*	(0.044)	542
		(2)	0.097**	(0.036)	467
Wood	20	(1)	0.064***	(0.015)	2028
		(2)	0.048***	(0.012)	1469
Paper	21	(1)	0.098***	(0.029)	924
		(2)	0.049	(0.026)	755
Publishing	22	(1)	0.155***	(0.034)	672
		(2)	0.089**	(0.030)	521
Chemicals	24	(1)	0.097**	(0.030)	978
		(2)	0.070**	(0.025)	896
Rubber and plastics	25	(1)	0.089**	(0.030)	869
		(2)	0.060*	(0.028)	741
Other non-metallic mineral	26	(1)	0.053**	(0.018)	1791
		(2)	0.026	(0.015)	1427
Basic metals	27	(1)	0.134***	(0.039)	606
		(2)	0.125***	(0.034)	493
Fabricated metals	28	(1)	0.081***	(0.021)	1519
		(2)	0.054**	(0.017)	1202
Machinery and equipment n.e.c.	29	(1)	0.173***	(0.033)	662
		(2)	0.162***	(0.033)	590
Computing machinery	30	(1)	-0.020	(0.122)	86
		(2)	0.066	(0.160)	81
Electrical machinery	31	(1)	0.216***	(0.039)	463
		(2)	0.161***	(0.037)	403
Television equipment	32	(1)	0.195***	(0.051)	319
		(2)	0.227***	(0.047)	296
Medical instruments	33	(1)	0.248***	(0.062)	203
		(2)	0.195	(17.715)	193
Motor vehicles	34	(1)	0.116*	(0.048)	418
		(2)	0.071	(0.045)	346
Transport equipment	35	(1)	0.134***	(0.032)	724
		(2)	0.079*	(0.032)	574
Furniture	36	(1)	0.070**	(0.021)	1377
		(2)	0.035	(0.018)	1068

Source: Micro data drawn from the Vietnam Enterprise Survey (2005-2010). Data of EZs location was retrieved in the report by Vietnamese Ministry of Investment and Planning (2011).
Notes: Year fixed effects, Region fixed effects and their interaction term are included. We only consider the districts which had the presence of labour in manufacturing industries (i.e: $I_{ijd} > 0$). SE is Standard Errors by Delta method (in parentheses).
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

4.6.2 Average Marginal Effects of Zoning policies by Year

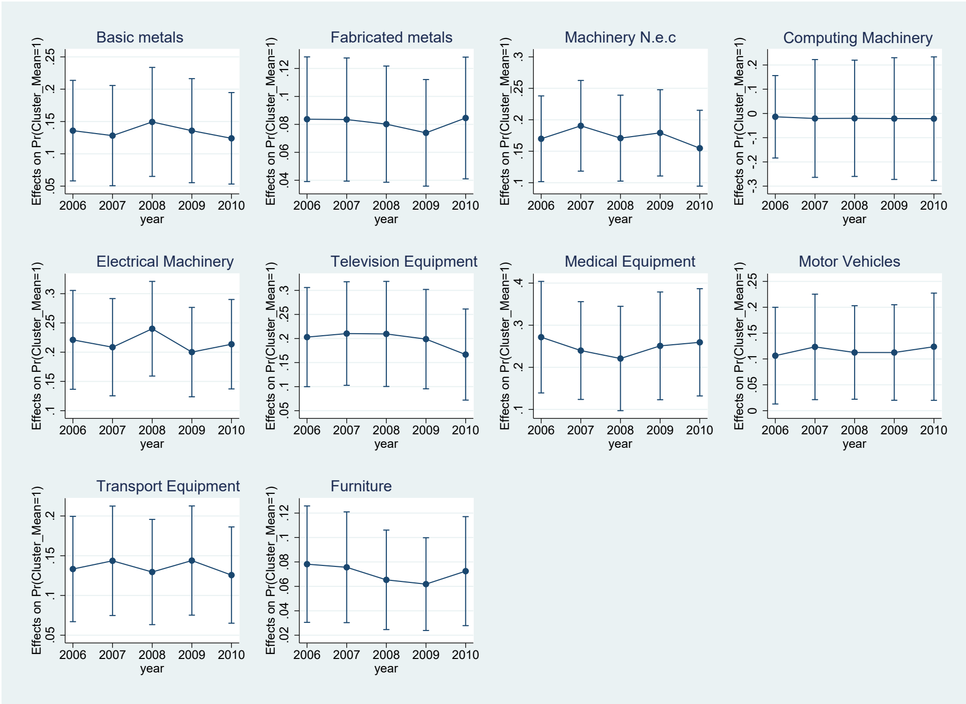
More details of the average marginal effects by year for each industries are visualized in Figure 4.5, Figure 4.6 (Logit Model 1), and Figure 4.7, Figure 4.8 (Logit Model 2). In general, these figures indicate positive average marginal effects of zoning policies on forming the industry-district cluster. The average effects varied differently by industries and by years during the research time frame.

Figure 4.5: Average Marginal Effects of Zoning Policies, Logit Model (1), Industry code 15-26



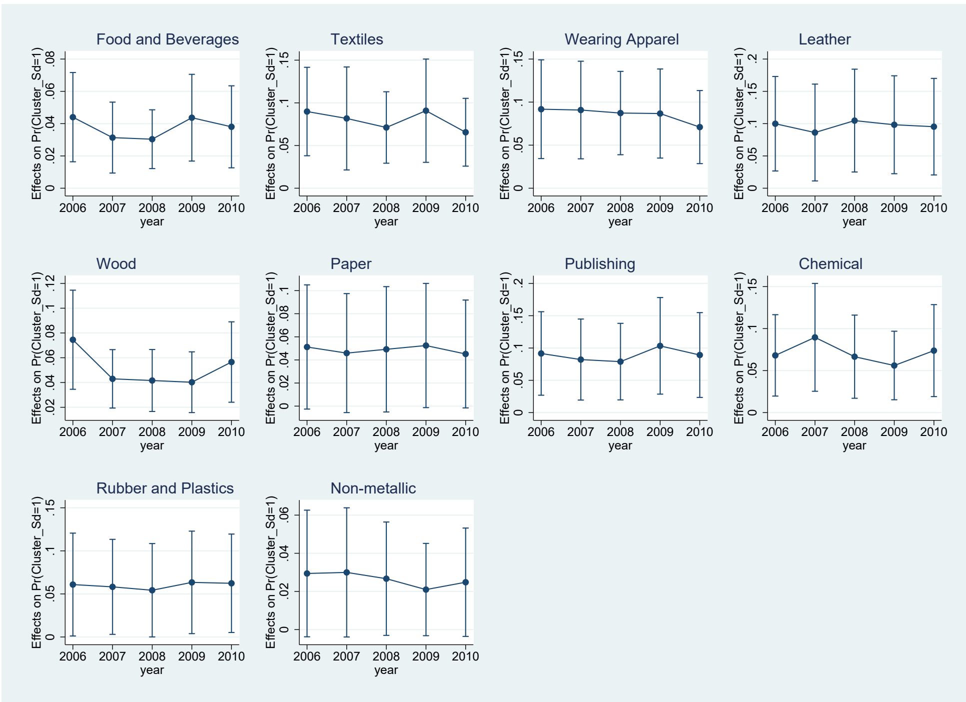
Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2005-2010.

Figure 4.6: Average Marginal Effects of Zoning Policies, Logit Model (1), Industry code 27-36



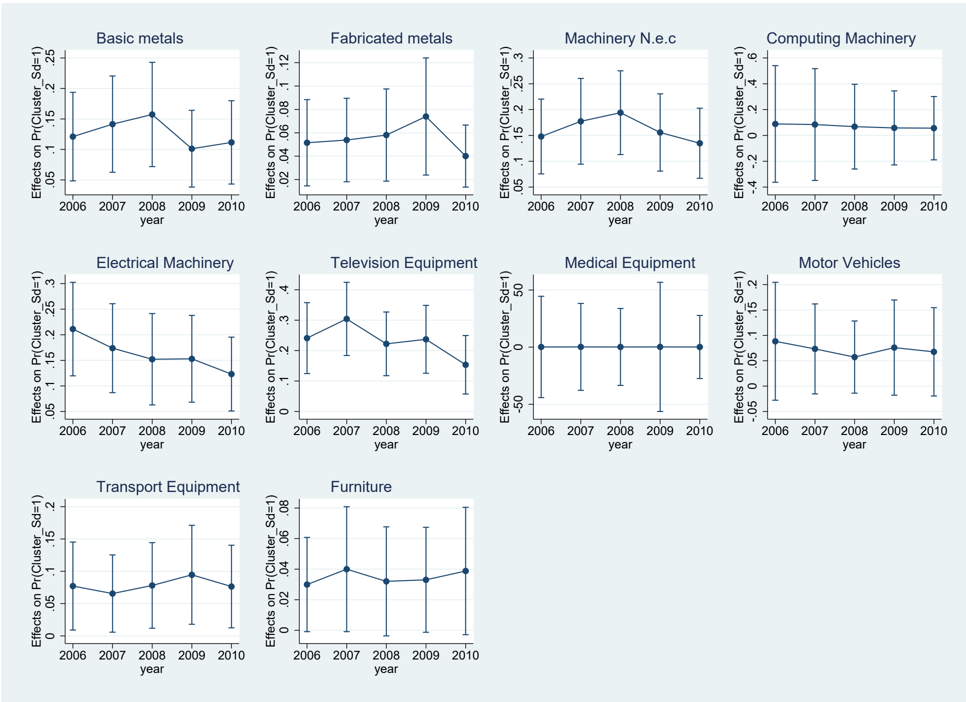
Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2005-2010.

Figure 4.7: Average Marginal Effects of Zoning Policies, Logit Model (2), Industry code 15-26



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2005-2010.

Figure 4.8: Average Marginal Effects of Zoning Policies, Logit Model (2), Industry code 27-36



Source: Author's compilation using data drawn from the Vietnamese Enterprise Survey 2005-2010.

4.7 Conclusion

In this study, the first evidence of two-forces generated by agglomeration effects of industry-district clusters in Vietnam was shown in the bell-shaped relation between productivity and industry-district density index in twelve of twenty industries for the time frame from 2005 to 2010. Clusters of firms in the same industry could stimulate productivity thanks to agglomeration economies, but also possibly pulled the productivity slowdown when congestion costs were dominant. Our results from system GMM estimation also indicated that firms locating inside a low-density industrial cluster gained benefit in terms of productivity, while firms in denser clusters did not. These results are consistent with theoretical literature (Duranton and Puga, 2004 and Fujita and Thisse, 2013), however, they are different from results founded by Newman et al. (2014) which indicated positive agglomeration effects of industrial clusters. Instead of indicating an exact threshold of labour density that determines the peak of bell-shaped productivity growth for manufacturing in general, i.e: medium-density places for the case study of manufacturing in IndiaDesmet et al., 2015, we visualize the estimated threshold of labour density index that stimulates peak value of productivity for each industry. This may give more profound implications for industry-specific policies in Vietnam. Regarding to proximity analysis, although the results are mixed for different linkages, its significance still confirms that geographical proximity from a firm's location to industrial districts matters to its performance.

As a mixed industrial policy, EZs in Vietnam have been established in accordance with the regional economic development plans of the country, and the urban land use of each provinces. Results of Logit model in this study shows that zoning policies stimulated higher possibility of forming industry district clusters in several manufacturing industries in Vietnam for the period from 2005-2010. Hence, the zoning policies were consistent with industrial cluster development plan in the country. It is suggested that zoning policies might help to ease the congestion effects for better performance of firms by setting new clusters, or stimulate the agglomeration economies by providing advanced infrastructure and simplified business regulations in emerging industrial districts. Further studies may assess if (i) land use competition between EZs and non-EZs, (ii) favourable taxation policies as well as "single window" policies in EZs relatively to non-EZs could influence the development of manufacturing in Vietnam and other developing countries.

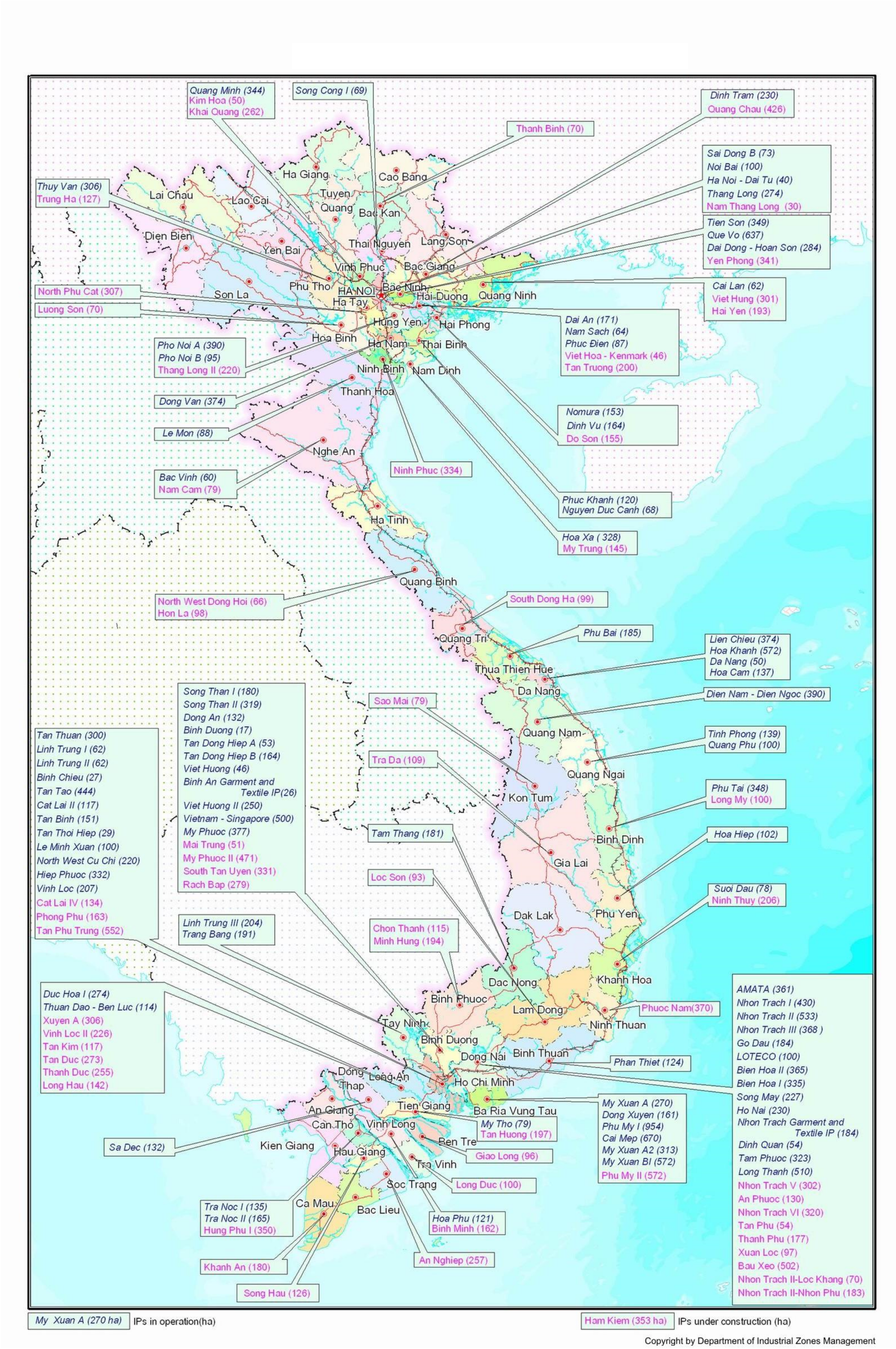
4.A Appendix

4.A.1 EZs and Map of EZs in Vietnam

According to Decree No.29-2008-ND-CP issued by the Vietnamese Government, EZs in Vietnam includes 4 types: “(i) Industrial zones means zones specializing in production of industrial goods and provision of services for industrial production, having fixed geographical boundaries, and being established pursuant to the conditions, order and procedures stipulated in these Regulations. (ii) Export processing zones means industrial zones specializing in production of export goods and provision of services for production of export goods and export activities, having fixed geographical boundaries, and being established pursuant to the conditions, order and procedures applicable to industrial zones stipulated in these Regulations. Industrial zones and export processing zones are all referred to as industrial zones, unless otherwise specified. (iii) “Economic zones” means zones having a separate economic space with an investment and business environment which is specially favourable for investors, and fixed geographical boundaries, and being established pursuant to the conditions, order and procedures stipulated in these Regulations. Economic zones are organized into functional areas including: non-tariff areas, bonded warehouse areas, export processing zones, industrial zones, entertainment areas, resorts, urban areas, residential areas, administrative areas and other functional areas consistent with the characteristics of each economic zone. (iv) Border-gate economic zones means economic zones which are formed in onshore border-gate areas with an international border-gate or main border-gate and which are established pursuant to the conditions, order and procedures stipulated in these Regulations."

The establishment of EZs aims to not only benefit from of local regional advantages, but also to enhance the inter-regional economic activities. In details, EZs attract capital and workers to a specific location, and allow easier mobility of inputs and outputs among districts and provinces (See Appendix 4.A.3 for the summary of target industries in EZs by regions in Vietnam, 2005-2020). Hence, EZs have been built and developed in advanced infrastructure areas, and are especially close to important national routes, and ports (See Appendix 4.A.1 for the map of EZs). Most EZs locating along the Highway No.1A which connects 31 provinces from the North to the South of Vietnam. National routes, which are highways under the direct control of the central government, connect important provinces of the country. For instance: Highway No.1A connects 31 provinces from the North to the South in Vietnam. Highway No.9 from Quang Tri province crossing Laobao town (Vietnam) connecting to Savannakhet (Laos) is an important part of the road system in the East-West corridor. Important seaports of Vietnam are Haiphong, Cailan, Tiensa, Quynhon, Catlai, Saigon seaport. River ports are also crucial in the transportation of cargo in the country. There are 66 ports in the North, 7 ports in the Central and 56 ports in the South. Importantly, the employees working inside the industrial zones could be deducted 50% from their individual income tax payment since 2008. The first SEZ in Vietnam in 1991 was the Tan Thuan Processing Zone in Hochiminh city.

Figure A.1: Industrial Zones in Vietnam



Source: Department of Industrial Zones Management, The Vietnamese Ministry of Investment and Planning.

4.A.2 List of Industries

Table A.1: 2-digit VSIC 1993 Industries

Classification	Name of Industry
15	Food Products & Beverages
17	Textiles
18	Wearing Apparel; Dressing, Dyeing Of Fur
19	Tanning, Dressing Of Leather, Products of Leather & Footwear
20	Wood & Products Of Wood
21	Paper & Paper Products
22	Publishing, Printing & Reproduction of Recorded Media
24	Chemicals & Chemical Products
25	Rubber & Plastics Products
26	Other Non-metallic & Mineral Products
27	Basic Metals
28	Fabricated & Metal Products
29	Machinery & Equipment N.e.c.
30	Office, Accounting & Computing Machinery
31	Electrical Machinery & Apparatus N.e.c.
32	Television & Communication Equipment
33	Medical, Precision & Optical Instruments
34	Motor Vehicles, Trailers & Semitrailers
35	Other Transport Equipment
36	Furniture; &Manufacturing n.e.c.

4.A.3 Key Industries in EZs by Regions, 2005-2020

Table A.2: Key Industries in EZs by Regions, 2005-2020

Regions	Target industries in SEZs
Northern	Hydro electricity
Midland & Mountainous	Manufacture of Agricultural and Forestry products
	Manufacture of Mineral products
	Manufacture of Chemicals and Chemical products
	Manufacturer of Equipment for Agriculture & Manufacturing
Red River Delta	Energy and Fuel
	Manufacture of Transport equipment, electricity equipment
	Informatics and Electronics
	Textile
	Footwear
	Manufacture of Agricultural, Forestry and Fishery products
	Manufacture of Chemicals and Chemical products
	Manufacture of Materials for constructions
Central Coast	Manufacture of Chemicals and Chemical products
	Manufacture of Agricultural, Forestry and Fishery products
	Manufacture of Materials for constructions
	Textile
	Footwear
	Manufacture of other goods for end-users
	Manufacture of goods that import inputs from other regions
Central Highland	Manufacture of Agricultural, Forestry products
	Hydro electricity
	Manufacture of Mineral products
	Manufacture of goods that in the cooperation in the East-West corridor
South-east	High technology manufacture
	Manufacture of High quality materials
	Manufacture of Equipment for local demand, export and import substitutions
	Extraction of crude oils, and Manufacture of Refined petroleum products
	Electricity
	Informatics and Electronics
	Manufacture of Tranport equipment, electricity equipment
	Manufacture of Agricultural and Forestry products
	Textile
	Footwear
	Manufacture of Chemicals and Chemical products
	Manufacture of Fertilizers
Mekong River Delta	Extraction of crude oils, & Manufacture of Refined petroleum products
	Electricity
	Manufacture of Agricultural and Forestry products
	Manufacture of Chemicals and Chemical products
	Manufacture of equipment for agriculture, agricultural, forestry, fishery products

Source: Authors' Summary from Development Plan of EZs, 2005-2020. Online document:
http://www.chinhphu.vn/portal/page/portal/chinhphu/noidungcackhucongnghiepkhuchexuat?categoryId=879&articleId=10001189

4.B Acknowledgement

This paper benefits from discussions with Christophe Degain, Douglas R. Nelson, Sebastian Miroudot, Robert Teh, researchers and officials from the Ministry of Industry and Trade of Vietnam, the General Statistics Office in Vietnam, the Foreign Trade University Vietnam, the WTI, and the WTO. We specially thanks Hanh Pham for sharing the raw micro-data, Khanh Pham for sharing the economic zones data of Vietnam, Stephan Kyburz for sharing the administrative boundary data of Vietnam, and Trang Nguyen for providing relevant legal documents. We also thanks participants in ETSG conference 2017 in Florence (Italy) for their helpful comments.

Huong Nguyen gratefully acknowledges generous financial support to her PhD study from the Swiss State Secretariat for Economics Affairs (Project: SECO/WTI/FTU), the WTI research funding, and the WTO PhD Support Program. Results of this paper do not necessarily reflect the views of those acknowledged individuals and organizations.

5 Does Proximity to Foreign Invested Firms Stimulate Productivity Growth of Domestic Firms? Firm-level Evidence from Vietnam

Joint with Stephan Kyburz

5.1 Introduction

Foreign direct investment is seen as an important driver of technological progress in particular for developing countries due to spillover effects from foreign firms on the domestic industry. Spillovers may foster technological change and thus reduce the productivity gap between advanced multinational firms and incumbent establishments. As technologically superior foreign firms enter a developing country, local manufacturers start imitating products and production processes. People working for foreign firms switch jobs and join local firms transferring valuable know-how on production processes and organisational structures. Entry of foreign firms may also intensify competition in upstream local industries or improve the quality of intermediate goods for downstream industries. Since spillover effects from foreign invested firms are potentially beneficial for technological progress of domestic industries and conducive for economic growth particularly in emerging markets, it is crucial to understand the geographical scope and magnitude of spillover effects.

At least since Marshall (1920) we are aware of the importance of localisation of industries, for which he identifies three sources: labor market pooling, intermediate inputs and technological spillovers. Krugman (1991) further elaborated that geography, hence localisation of industry, clearly matters and that spillovers are much a local phenomenon. Firms benefit from being near other firms. Porter (1990, 2011) points out the importance of geographic concentration in industry clusters. Research on agglomeration economies emphasises the existence of information-based spillovers: The presence of knowledgeable neighbouring people or firms lead to significant learning processes (Glaeser, 2010). The importance of geographical proximity is emphasized e.g. by Orlando (2004) and Lychagin et al. (2016) in the context of R&D activity, Rosenthal and Strange (2003) on agglomerative externalities, Aharonson et al. (2007) on knowledge spillovers, Duranton and Overman (2005) on localization patterns, Rosenthal and Strange (2008) on human capital spillovers, and Greenstone et al. (2010) for openings of large plants in the US.

This paper contributes to the literature by shedding more light on the relevance of spatial proximity with regard to spillover effects from foreign invested firms on the local industries' productivity growth. Essentially, we want to answer the question whether a domestic firm can absorb foreign know-how and increase its productivity from having foreign firms in its close surroundings. We complement the existing literature on FDI spillovers in one important way: we analyse the relevance of geographic proximity for FDI spillovers using exceptionally detailed geo-referenced information on location of domestic and foreign invested firms comprising an almost exhaustive register of Vietnamese firm data. The bulk of studies in this vast literature look at the presence of foreign firms within a certain industry and/or region yet are not able to analyse the spatial scope of spillovers due to existing data limitations – exact location information is usually not extractable from firm level survey data. Studies with a specific focus on localisation of FDI spillovers are scarce. Comparable examples in terms of accuracy of location data are Halpern and Muraközy (2007) who present a similar study setup while employing a much smaller sample of Hungarian firms. Barrios et al. (2012) utilise an Irish plant level survey with detailed information on firms' location, but the sample comprises just 1790 firms. In an influential study, though with less geographic detail, Aitken and Harrison (1999) analyse the presence of foreign invested firms in 220 districts in Venezuela, but find no localised spillovers.[1]

Adoption of advanced technologies from the technology frontier is seen as one of the main drivers of growth in developing countries (Acemoglu et al., 2006). In this respect, it is essential for domestic firms to make use of technology imitation opportunities and copy know-how from foreign direct invested companies. Productivity differences stem from disparities in technological knowledge. These differences may originate either from differences in R&D investment or the ability of adaptation to new technologies (see e.g. König et al., 2012). The intensity of knowledge diffusion and the absorptive capacities of firms is seen as crucial channel for firms to catch up with technologically more advanced firms.

Knowledge diffusion conveyed by spillovers occurs through various channels. *First*, imitation or demonstration effects are an obvious source of knowledge transmission. Reverse-

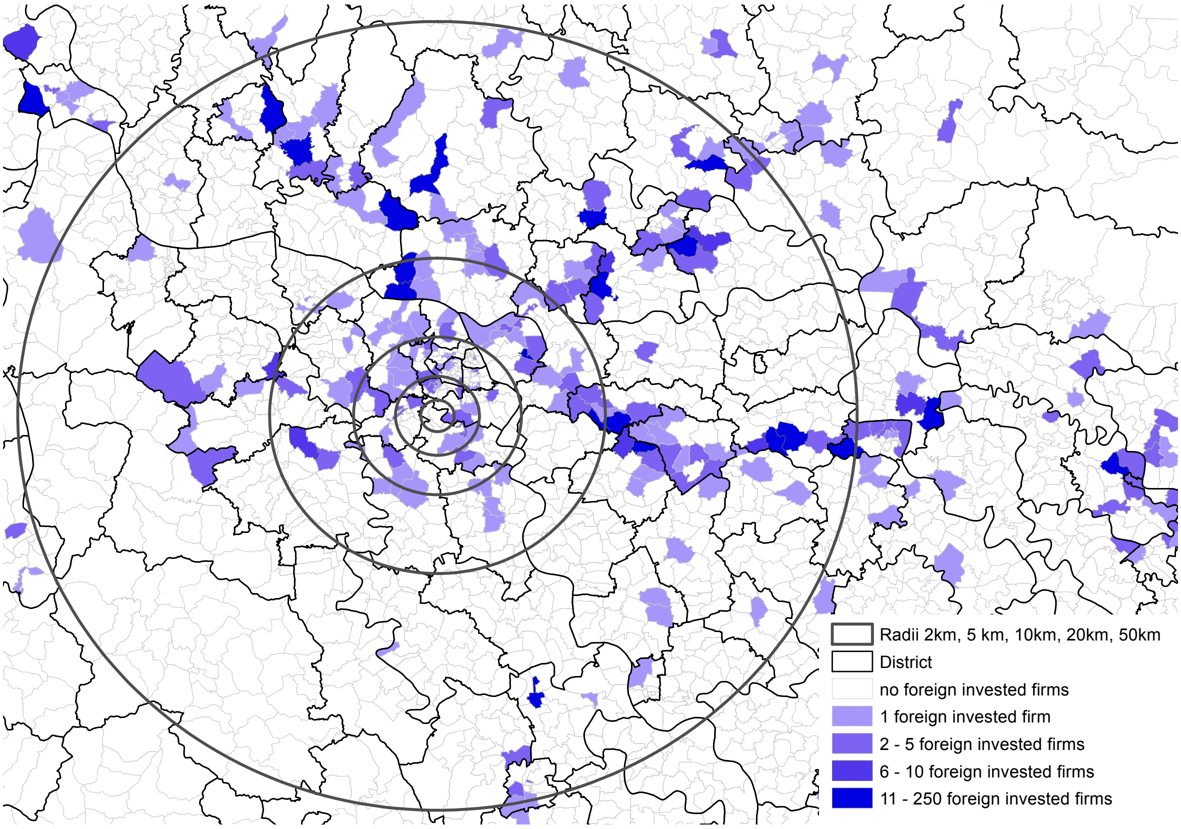
[1] A detailed overview of related literature is provided in the next section.

engineering of production processes by local firms seems a useful way to improve technical know-how (Görg and Greenaway, 2004; Smeets, 2008). *Second*, skill acquisition of local firms through the labor market improves their human capital. Proximity between firms increases the likelihood of physical contact between workers or product developers. Knowledge transition may work through a direct spillover from complementary workers or people switching jobs from a foreign invested firm to a domestic firm (Görg and Greenaway, 2004). This spillover channel that works through the labor market is seen as an important mechanism (Fosfuri et al., 2001) and is supported by empirical work (Görg and Strobl, 2005; Markusen and Trofimenko, 2009). *Third*, competition from foreign firms in the same industry may put pressure on the local firm to use existing technologies more efficiently, when imitation of the foreign company's technology is hardly possible. Competition may also increase the speed of adoption of new technology (Görg and Greenaway, 2004). *Fourth*, spillover effects may materialise through vertical linkages (Smeets, 2008). Javorcik (2004) and Lin and Saggi (2007) emphasise the importance of direct technology transfers through backward linkages. Foreign invested firms are willing to transfer know-how to their suppliers in order to improve the quality and lower the prices of intermediate goods. Blalock and Gertler (2008) argue that a foreign firm even must make the technology widely available to several suppliers in order to avoid hold-up. Domestic firms may also become more productive through forward linkages in case they can buy technologically more advanced, qualitatively better, or less costly intermediate goods produced by multinationals in upstream sectors. Francois and Hoekman (2010) and Arnold et al. (2011) emphasise the crucial linkages between production in goods and services, as services are often direct inputs into economic activities, determining the productivity of the fundamental factors of production, labor and capital.

Our proposition is that proximity between foreign invested and domestic firms crucially facilitates the know-how and technology transmission channels, and hence that geography and the localisation of industry matters for spillover effects. The entry of almost 2000 foreign invested firms to Vietnam in the period 2005 to 2010 presents a unique and suitable natural experiment to assess possible technology spillovers from foreign firms on the incumbent industries. We presume that the nearer a *foreign investment shock* occurs, the more intense potential spillover effects must be and the larger the productivity gains are for domestic firms. Spillover effects decay with increasing distance, as demonstrated by Orlando (2004), Halpern and Muraközy (2007), Greenstone et al. (2010), Lychagin et al. (2016), or Barrios et al. (2012). Our research setup lets us more precisely estimate how localised these spillovers are in the context of foreign direct investment, benefiting from enhanced spatial accuracy over existing studies.

Our data set includes around 67'000 manufacturing firms in Vietnam over the period

Figure 5.1: Circles with different radii and foreign invested firms in 2010 in the Hanoi area



Notes: The map shows a representative example of circles around a firm with radii of 2km, 5km, 10km, 20km, and 50km. For each ward the number of foreign invested firms is indicated. Calculations of firms are based on the Vietnam Enterprise Survey 2010.

2005 to 2010. Most importantly, we can determine the location of all firms by using information on the ward, the smallest administrative unit of Vietnam.[2] Since we know the location of all firms in Vietnam, we are able to observe the presence of foreign invested firms in close proximity to each domestic establishment over time. Figure 5.1 resembles the main idea of our paper for the case of Hanoi metropolitan area: we virtually draw circles of various radii around each domestic firm and measure the presence of foreign firms within its surroundings.

Our empirical strategy to estimate localised spillover effects consists of two stages. The first stage computes total factor productivity of firms from the parameters estimated by an industry specific production function method introduced by Levinsohn and Petrin

[2]Vietnam comprises more than 11'000 wards. The median size of a ward with manufacturing activity is 8 km^2, representing the size of a circle with a radius of about 1.6 km. Measuring the size of wards weighted by the number of firms within a ward, the median size of wards is 4 km^2, a circle with a radius of just 1.12 km. The mean size of wards is then 8.2 km^2. More information on the spatial dimensions of our data and how we geo-reference firms is provided in section 5.4.2.

(2003).[3]

The second stage causally identifies localised spillover effects from foreign invested firms on the local manufacturing firms using an estimation procedure first proposed by Anderson and Hsiao (1981), and controlling for possible confounding factors that may both influence location choice of foreign firms and domestic firm productivity. Additionally to absorbing the firm fixed effect by first differencing, we include industry fixed effects, province-time fixed effects, and time fixed effects. We rigorously examine different aspects relevant to the FDI spillover literature, yet with specific focus on the spatial scope of spillover effects, looking at horizontal – within industry – and vertical linkages working through the supply chain.

Our results affirm that spillover effects are indeed localised and quickly fade with increasing distance. Spillover effects seem to be strongest between 2 and 10 kilometres and attenuate rapidly across geographic space. Due to the great coverage of the data set including a large number of firms with less than 20 employees, we are able to analyse whether the local small and medium enterprises (SME) also benefit from foreign direct investments. Results indeed show strong support in favour of relevant spillover effects from foreign investments on the local SME sector, which is a stark results in terms of business development policy. Relatively unproductive firms also experience large productivity gains through foreign investment in their vicinity, yet there appears to be no within industry absolute convergence in productivity levels. With regard to vertical linkages operating through the supply chain among manufacturing firms, the results suggest that firms benefit from closely located foreign suppliers, but seem to be negatively affected by foreign backward linkages to downstream foreign firms.[4] Furthermore, spillovers occurring through vertical links to service industries seem to be much less geographically limited.

The rest of the paper is organised as follows. Section 2 gives an overview of the existing literature on spillovers from foreign direct investment with a specific look at the studies that scrutinise the localisation of such effects. In addition, we present an overview of studies that look at the case of Vietnam. Section 3 elaborates the trends and patterns of foreign direct investment in Vietnam. In section 4 we present the data used.

[3]The method of Levinsohn and Petrin (2003) is an extension of the framework initiated by Olley and Pakes (1996). Olley and Pakes (1996) decompose the residuals of the production into unobserved firm level productivity and zero-mean measurement errors. They calculate the unobserved productivity of an individual firm by using parameters estimated from the industry's production function. Other researchers such as Halpern and Muraközy (2007), Van Beveren (2012), and De Loecker and Warzynski (2012) define the unobserved firm level productivity (Olley and Pakes, 1996) as the total factor productivity. For consistency of the terminology, we also refer to the firm-productivity as total factor productivity.

[4]We use the same definition of forward and backward linkages as Newman et al. (2015): forward linkages are upstream foreign suppliers; backward linkages are downstream foreign customers. This definition of *forward* and *backward* is implicitly referring to the perspective of the foreign firm.

Section 5 describes the method applied to estimate firm level total factor productivity and presents the corresponding TFP estimates. Section 6 explains the identification strategy to estimate FDI spillover effects within the spatial framework and subsequently discusses the results, including the relevance of horizontal and vertical linkages for spillover effects. Finally, section 7 concludes.

5.2 Related Literature

Our paper contributes to a large international literature on spillover effects of foreign investment on the domestic economy. We present a short overview of the most related studies in two parts. The first part looks at the relevance of horizontal and vertical linkages, and the heterogeneity in effects. The second part sums up the contributions for the case of Vietnam.

5.2.1 Literature on FDI Spillovers

General results: A growing number of theoretical and empirical studies has shown that FDI is a crucial driver stimulating economic growth of the host country through the transfer of knowledge and technologies from advanced multinational enterprises (MNEs) to the domestic firms.[5] Several comprehensive surveys of the literature come to mixed conclusions concerning the importance of FDI for the local economy. Görg and Greenaway (2004) investigate results of 40 studies on FDI spillover effects and attempt to draw general conclusions from the early literature. Their overall corollary is, first of all, that FDI is likely to be a key driver of economic growth by boosting capital formation and the quality of the capital stock in host countries. Multinational companies seem to bring best practice technology and management with them. They deduce that absorptive capacity of domestic firms and geographic proximity to multinationals are important determinants of spillover effects from foreign invested to domestic firms. Governments are hoping to stimulate these external benefits of FDI by offering incentives to foreign companies, suggesting that policy improvements should target the general conditions for doing business instead of particular industries or firms.

Smeets (2008) summarises the literature by concluding that the evidence on the magnitude, direction, and even existence of knowledge spillovers from foreign direct investment is ambiguous. Nonetheless, knowledge spillovers through channels like worker mobility,

[5]See e.g. Lim (2001) for an account of FDI and economic growth; Girma et al. (2001) and Blomstrom and Kokko (2003) on FDI spillovers;Blonigen et al. (2007) indicate spatial interdependence as one of FDI determinants.

demonstration effects, and vertical linkages all seem to matter. Yet, the literature seems to agree on missing evidence for spillovers working through forward linkages, i.e. when foreign firms supply goods to domestic firms. Smeets (2008) emphasises the importance of distinguishing and specifically analysing different channels through which spillovers may occur.

Horizontal spillovers: Horizontal spillovers occur within an industry. Results are mixed. There are two main arguments. On the one hand, firms of the same industry may benefit from each other through face-to-face contacts and imitation of products or processes. On the other hand, firms in the same industry compete with each other. Competitive pressure may lead to more efficient use and quicker adoption of technologies, but it may also drive up the average cost curve due to fewer sales. Using a panel of 4'000 Venezuelan plants between 1976 and 1989, Aitken and Harrison (1999) find that foreign equity participation increases productivity of recipient plants with less than 50 employees, suggesting that plants benefit from productive advantages of foreign owners. Crucially, they also find a negative impact of foreign ownership on wholly domestically owned firms in the same industry. These large significant negative effects are brought by competitive pressures. Overall they conclude that there is no clear evidence of the existence of technology spillovers from foreign firms to domestically owned firms. In contrast to this early influential enquiry, Abraham et al. (2010) is one of the few studies that find positive within industry spillovers. They find that it was beneficial for total factor productivity of domestic firms when there was a certain presence of foreign competitors in analysis of more than 15'000 manufacturing firms in China in the years 2002 to 2004.

Vertical spillovers: Much attention has also been paid to the role of FDI spillovers to domestic firms through vertical linkages in the supply chain. Either a foreign firm supplies intermediate goods to a domestic firm or vice versa. A review of studies on the relevance of vertical linkages is conducted by Smeets (2008). Most studies find positive spillovers through backward linkages, but negative effects in the case of forward linkages. Javorcik (2004) analyses spillovers effects of FDI on productivity through backward and forward linkages using a firm level panel dataset from Lithuania. The author shows that spillovers are associated with projects that are shared between domestic and foreign firms, and not with fully foreign owned projects. Robust evidence for spillovers working through backward linkages is found, while intra-sectoral spillovers are absent in her study. Blalock and Gertler (2008) demonstrate the gain in productivity of Indonesian local suppliers through spillovers from foreign firms in downstream industries over the period of 1988 to 1996. In addition, the significance of inter-industry benefits to upstream domestic firms also seems to depend on the origin of the foreign investment in the downstream industry (Javorcik and Spatareanu, 2011).

Geographic Proximity: Spatial proximity between economic agents has become an important explanation of economies in production since the early stages of economic research and was already investigated by Marshall (1920) in terms of specialised clusters of inputs (e.g. labor, materials, services), and technology spillovers. Halpern and Muraközy (2007) specifically investigate geographical distance as a determinant of FDI spillovers. The novelty in their study is the link between the TFP level (estimated by Levinsohn and Petrin, 2003) and the FDI spillovers in light of the distance from foreign firms to domestic firms in Hungary (1996–2003). They confirm that distance indeed matters for horizontal spillovers and emphasise the local nature of those. In addition, the authors extend the vertical and horizontal linkages proposed by Javorcik (2004) by weighting these variables with a function of distance between a foreign invested firm and a domestic firm. Though, a drawback of Halpern and Muraközy's (2007) approach is the assumption on the functional form of the distance.[6] A priori, using a functional form assumption for the distance $f(d)$ to weight foreign firms, it is unclear whether the effect of a foreign firm which has low output but is close to a domestic firm is similar in magnitude to an other foreign firm with a large output but which is far away. Besides not using a functional form assumption, our study also provides a higher degree of accuracy of firms' locations. Furthermore, they are neither able to control for agglomeration effects such as the size of the labour market in the vicinity of each firm. Girma and Wakelin (2007) emphasise the strong intra-regional dimension of spillover effects in a study on the electronics industry in the UK. However, their study stays at a spatially rather crude level by dividing the UK into 10 regions, finding that horizontal spillovers occur within a region, but not outside (or across) the region.

In an investigation of spillovers from local and global R&D activities of domestic and foreign plants, Barrios et al. (2012) estimate distance decay effects by using the sample of Irish plant-level survey data from 1986 to 1996. Most relevant for our study, they also analyse local spillovers within circular areas around each plant. Considering all plants, they estimate significant local spillovers of R&D activity conducted in Ireland. Effects are strongest and significant within a radius of 10 km around a plant, but decay quickly beyond. Interestingly, domestic firms seem to benefit more from local R&D activities conducted by other domestic firms than those by foreign firms. Compared to Barrios et al. (2012), we employ a much larger data set and our geo-referenced firm location

[6] Halpern and Muraközy (2007) use a variety of functional form assumption to weight the horizontal and vertical linkages. Specifically, they use the following weighting functions of distance: $f_1(d) = 1/(1+d/100)$ (the linkage to a foreign firm that is 100 km away from the domestic firm, is weighted by 0.5); furthermore, they also use two other functions with more pronounced decay patterns: $f_2(d) = 1/(1+d/100)^2$ and $f_3(d) = 1/ln(1+d/100)$ as weighting functions.

information is also more precise.[7]

5.2.2 Literature on FDI Spillover Effects for the Case study of Vietnam

With a fast growing trend of foreign investment into Vietnam, extensive research has been conducted to study the role of FDI inflows for the local economy, whereof we specify the most relevant works.[8]

One of the early studies investigating FDI spillovers effects in Vietnam, Nguyen et al. (2008) use a large firm-level data set for the period 2000 to 2005 to study both horizontal and vertical spillover effects. They find evidence of positive backward spillover effects for the manufacturing sector and positive horizontal spillovers for the service sectors. They do not discover any evidence of backward and forward spillovers for the service sector. The study does not consider any spatial dimension in FDI spillovers. Instead the baseline empirical model merely uses the presence of foreign firms within an industry, incorporated in a Cobb-Douglas production function approach, as applied by most spillover studies.

Thang et al. (2016) are the first to use a spatial econometric model for the case of Vietnam to investigate the importance of proximity for spillovers. Their study is much related to ours in the sense that they specifically examine geographical distance in the context of spillovers. Although they employ a spatial econometric model, they merely use the provinces as geographic unit of analysis. The spatial accuracy is hence considerably lower compared to our study. Thang et al. (2016) find inter-regional spillovers to be four times larger than intra-regional spillovers. Confirming some findings in the literature, they estimate negative horizontal spillovers, positive backward and negative forward spillovers effects. Since their data set is limited to the period 2000 to 2005, the paper cannot provide results for the important period after the first Investment Law (2005) allowed more flexible types of foreign investment and investment license controls were decentralised to the provincial level.

Anwar and Nguyen (2014) analyse the performance of manufacturing firms in the eight

---

[7] We employ a sample of 67'000 Vietnamese manufacturing firms, while Barrios et al. (2012) have a sample of just 1790 plants. In order to geo-reference firms, Barrios et al. (2012) use Irish district electoral divisions (DED) that have a mean size of 21 km^2. The Vietnamese wards that we use in the analysis have a mean size of 17.55 km^2; when weighted by the number of firms within each ward, the mean size is even smaller at 8.21 km^2.

[8] See e.g. Anwar (2011) for the analysis of FDI linkages and local firms' export activities. Another study for FDI in Vietnam (2001-2008) by Kokko and Thang (2014) indicate that the presence of foreign counterparts and foreign suppliers would increase the exit ratio of domestic firms.

regions[9] of Vietnam affected by varying intensity of foreign investment. By applying 2SLS estimations and using manufacturing firm-level data for the period 2000 to 2005, Anwar and Nguyen (2014) suggest that through backward linkages, FDI spillovers positively influence only four of eight regions (i.e. Red River Delta, South Central Coast, South East and Mekong Delta River). They conclude that the impacts of FDI on the domestic economy varies considerably across regions.

Newman et al. (2015) separate out productivity gains along the supply chain through direct transfers of knowledge and technology between linked firms. Importantly, they disentangle the spillovers through direct linkages, real technology transfers and other indirect effects. Their results confirm the importance of vertical linkages versus horizontal linkages with regard to spillover effects. More specifically, considering only direct linkages, they find that domestic firms experience positive productivity spillovers through their direct linkages with upstream FDI suppliers of inputs.

Complementing the existing literature, this paper proposes a simple though intriguing approach to evaluate localised spillover effects. We use the available detailed information on the location of foreign and domestic firms to evaluate how foreign investment in the surrounding area of domestic firms impacts on their performance.

5.3 Foreign Direct Investment in Vietnam

Since the *Doi Moi* (Renovation) in 1986, Vietnam's development policy has sought to promote high economic growth, macroeconomic stability and international integration (Nguyen and Ramachandran, 2006; UNIDO, 2011b). Since the introduction of a new Law on Foreign Investment in 1987 with amendments in 1990, 1992, 1996, 2000, and 2005, Vietnam was constantly expediting foreign direct investment in order to strengthen capital formation and know-how transfer from more advanced economies through foreign firms.[10] In the earlier periods of opening up the country between 1988 and 2001, foreign investors were compelled to form joint-ventures with domestic firms, while after 2001 investments in the form of wholly foreign owned enterprises became more important (UNIDO, 2011b). In 2005, the Law on Foreign Investment and the Law on Domestic Investment were unified into one common Investment Law that in combination with the unified Enterprise Law balanced the rights and treatment between domestic and foreign investors. Importantly,

[9] The government of Vietnam groups the provinces into eight large regions: Northwest, Northeast, Red River Delta, North Central Coast, South Central Coast, Central Highlands, Southeast, and Mekong River Delta.

[10] For a detailed discussion see e.g. Nguyen and Ramachandran (2006).

the 2005 Investment Law and 2005 Enterprise Law decentralised the control of investment license and business registration to provincial-level authorities.

Subsequently, in January 2007, Vietnam made another important step towards remarkable international economic integration by acceding the World Trade Organization (WTO) that brought a further push to foreign investment and eventually resulted in registered USD 198 billion in foreign capital in 2011 (UNIDO, 2011b). Foreign direct investment is concentrated mostly in the manufacturing and real estate sectors, accounting for 77 percent of total registered capital in foreign invested projects in 2011 (UNIDO, 2011b). Manufacturing alone accounts for 58 % of all projects. The share of exports carried out by foreign invested firms jumped from 47 % to 57.2 % in 2007, then slightly decreasing to 54.2 % in 2010.

Table 5.1: Description of Foreign Firms in Process Manufacturing Sectors, Vietnam (2005–2010)

Year	**Number**	**Δ (%)**	**Share of foreign invested firms (%)**		
			Total output	**Total labor**	**Total firms**
2005	2654	14.10	43.51	36.35	11.05
2006	3032	14.24	46.01	39.38	11.29
2007	3516	15.96	45.23	41.41	11.32
2008	3958	12.57	44.42	42.77	10.31
2009	4353	9.98	41.66	42.64	9.74
2010	4587	5.38	44.47	44.73	9.80

Notes: Authors' compilation using the data drawn from the Vietnamese Enterprise Survey (2004–2010). The column "Number" is the number of foreign invested firms. The column Δ(%) is the percentage change in number of FIEs. The column "Share" is the Share of FIEs in Total Manufacturing firms' Total out, Total labor, and Total number.

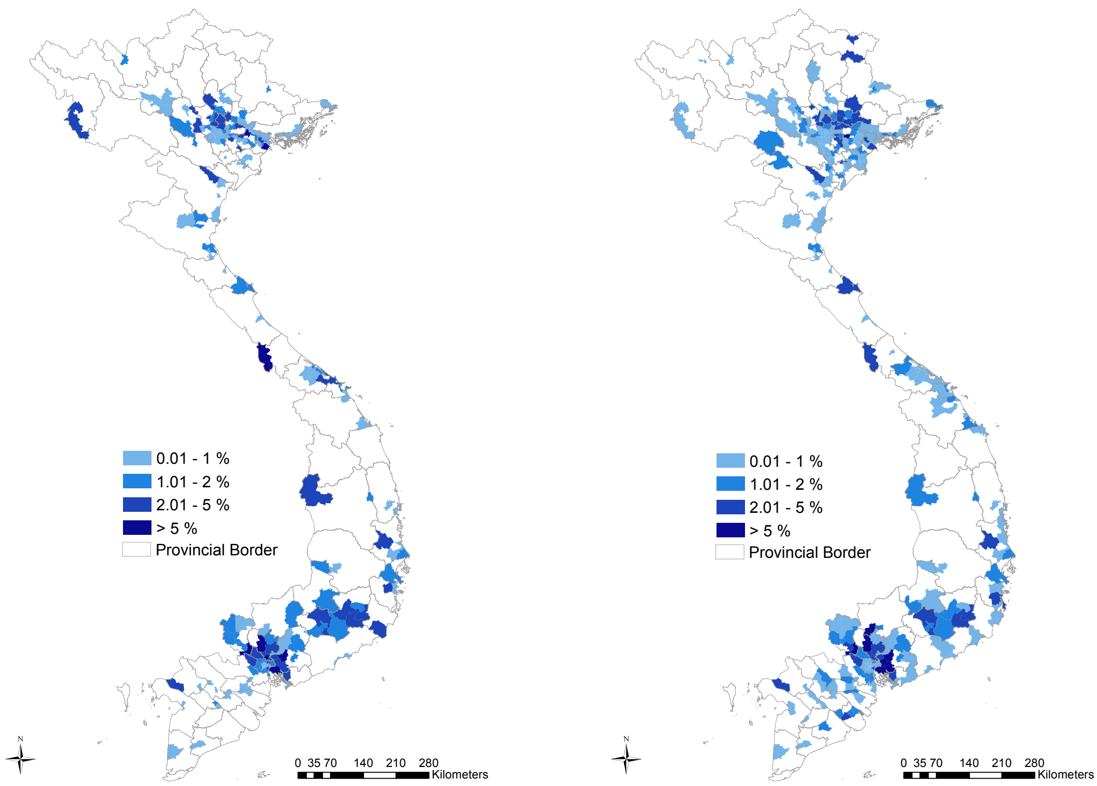
From 2005 to 2010, Vietnam attracted almost 2000 foreign invested firms (net increase) in the process manufacturing sectors. The international financial and economic crisis in 2007 and 2008 probably led many foreign companies to leave the country resulting in a very dynamic pattern of foreign investment during these years. Table 5.1 indicates that the increase in percent of number of foreign firms was 12.57% in 2008, then went down to merely 5.38% in 2010 while the number ranged from 14-16% in years before 2008.

Nevertheless, the crucial role of foreign invested firms in the process manufacturing sectors of Vietnam were still maintained with their significant shares in total output and job creation (table 5.1). From 2005 to 2010, the output share of foreign firms was in the range of 41 to 46%, and hence rather stable over time. The number of workers employed by foreign manufacturers augmented from 36% in the year 2005 to almost 45% percent in the year 2010. This highlights the growing foreign presence in the labor market, where potential spillover channels are at play.

Figure 5.2 presents two maps with the regional allocation of foreign invested firms in the years 2005 and 2010 at the district level. It gives a clear indication of the dispersion of foreign presence across provinces. While in the year 2005 most of the foreign firms were located in the economic core areas around Hanoi Capital and Ho Chi Minh City, foreign activity moved more into suburban and rural areas over time up to 2010. This change in the regional distribution is possibly due to the more favourable investment environment mentioned above, for instance, the simpler licensing process and the more decentralised authority control at provincial level. Figure B.1 in the appendix presents two similar maps using the share of revenue accruing to foreign firms in each district in the years 2005 and 2010. Figure B.2 in the appendix shows the same pattern for the labor force working for foreign invested firms.

Figure 5.2: Shares in Number of Foreign Invested Firms by District in Vietnam 2005 & 2010

(a) 2005 (b) 2010



Notes: Share of foreign invested firms per province is equal to the number of foreign invested firms in the province over the total number of firms. The maps are based on authors' calculations using the Vietnam Enterprise Survey 2005 & 2010. Administrative boundaries are based on Global Administrative Areas data (www.gadm.org). Several Vietnamese islands (e.g. Hoang Sa and Truong Sa) are not displayed due to the limitation of the GADM administrative boundaries data.

5.4 Data Description

This section gives a brief overview over the Vietnamese firm level data used (section 5.4.1) and explains in more detail the process of geo-referencing firms using the lowest Vietnamese administrative units (section 5.4.2).[11]

5.4.1 Firm Level Data

To investigate the relevance of proximity for FDI spillovers in Vietnam, we use firm level data of process manufacturing industries in Vietnam drawn from the Vietnam Enterprise Survey for six consecutive years (2005–2010).[12] The survey is a rich firm level database reporting yearly information on the legally registered enterprises that were in operation on the 31^{st} of December each year. The data provides information about the establishment year, the location (at province, district, and ward level), the revenue, the profit before taxes, the total cost, the total wages, the number of workers, and the value of net fixed assets of each firm. The information on foreign investment is a dummy variable that is 0 for firms with no foreign investment and 1 for firms that are partly or fully foreign invested. Every firm in the data features a unique identification number and is compiled in an unbalanced panel over six years.

In order to exploit the location specific information of firms in the survey to examine spillovers at the local level, the firm level data is merged with the ward level administrative boundary data retrieved from the Global Administrative Areas database (GADM)[13], and combined with the national input-output table of Vietnam (2007) assembled by the Vietnam General Statistics Office.[14] Industry codes in the firm level data are specified by the 2-digit and 4-digit Vietnamese Standard Industrial Classification 1993 (VSIC 1993)[15].

[11]See Nguyen (2016) for a detailed description of the Vietnamese manufacturing firm-level dataset from 2000-2010. See also Ha and Kiyota (2014) and Newman et al. (2015) for the descriptions of similar datasets respectively in the time frames from 2000-2009 and from 2009-2012.

[12]The census is annually conducted by the Vietnam General Statistics Office (GSO; www.gso.gov.vn) since 2000 till the current year. The data is published at an aggregated level in the Statistical Yearbook of Vietnam. We restrict our analysis to the years 2005 to 2010 because the new Law on Investment was introduced in Vietnam in 2005, and came into effects in 2006. Dataset after 2010 is not available to us. We would like to thank Pham Hanh at the Middlesex University (UK) for sharing the raw data with us, and Doan Thi Thanh Ha and Doan Hung at the Foreign Trade University (Vietnam) for discussing and sharing related documents.

[13]GIS shapefiles of administrative boundaries for Vietnam are available at www.gadm.org.

[14]The input-output table is available at:
http://www.gso.gov.vn/default.aspx?tabid=512&idmid=5&ItemID=10752.

[15]VSIC 1993 is provided by the GSO, and is similar to the International Standard Industrial Classification (ISIC Rev.3) provided by the United Nations. We use only the sample of industries for which

To merge the firm-level data with the input-output table, we convert the 4-digit VSIC 1993 in the firm-level data to the 2-digit industrial classifications of the input-output table using the concordance table provided by the GSO. After merging the firm-level data with the input-output table, only firms in industries considered in the input-output data are included. It is assumed that the cost coefficients in the input-output table do not change over the studied periods.[16]

The yearly number of firms in each industry is presented in Table B.2. The resulting panel is unbalanced including 67'275 firms. Table B.2 in the appendix shows the number of firms in each industry.

5.4.2 Information on Firms' Location

To the best of our knowledge, we are the first to geo-reference the lowest administrative unit of the Vietnamese governing system in the Vietnam Enterprise Survey.[17] Vietnam comprises more than 11'000 wards (communes). Since the Vietnam Enterprise Survey provides information on the province, the district, and the ward for each firm since 2005, we are able to geo-locate all of the 67'000 firms in our data set to their respective wards.[18]

The geographical scope of Vietnamese wards is remarkably small-scale. If we summarise the spatial dimensions of wards with at least one manufacturing firm domiciled, the median size is 7.94 km^2, while the mean size is 17.55 km^2 with a standard deviation of 34.78. Looking at the whole sample of firms and weighting the extent of wards by the number of firms based within, we receive a median size of just 3.99 km^2, and a mean of 8.21 km^2 with a standard deviation of 16.79. This is equal to a circle of radius 1.12 km^2. These numbers convey the geographic specifics at which we can perform the analysis and highlight the exceptionally small spatial scale used.

For each ward we determine the dyad wards within a certain radius and are thus able to calculate distances between firms with high accuracy.[19] This enables us to model the

2-digits industry classification ranges from 15 to 37.

[16]This assumption follows Javorcik (2004).

[17]In order to map the ward information in the Vietnam Enterprise Survey on the Global Administrative Areas boundary shapefiles, we used the geocode command in Stata and mapped the wards according to the information on province, district, and ward. We then manually checked all 11'043 wards in ArcGIS for the correct geolocation. As the position of a ward we use the geographic centroid. For wards that were not located automatically by the geocode command, we extracted the coordinates by the use of Google Maps (http://maps.google.com).

[18]The survey data of manufacturing sectors records 5'662 unique codes of wards, 664 unique codes of districts, and 63 unique codes of provinces. These administrative units incorporate at least one observation, resulting in about 5'300 wards with no registered manufacturing firm in operation.

[19]Since we cannot determine the exact location of firms within wards, we assume that they are all

entire agglomeration of firms and to calculate industry specific statistics at various spatial dimensions for each firm in the data set.

5.5 Total Factor Productivity Estimation

Our empirical strategy consists of two steps that are prevalently applied in the spillover literature (see e.g. Combes and Gobillon, 2015, Newman et al., 2015, Anwar and Nguyen, 2014, and Barrios et al., 2012). In the first step discussed in this section, we estimate a production function within each industry and use the parameters estimated to impute firm level productivity. We then proceed to the second step of estimating spillover effects with a pronounced focus on spatial proximity between firms in the subsequent section.

5.5.1 TFP Estimation Methodology

We compute the firm-level productivity from the estimation of parameters in industry specific production functions as proposed and documented by Levinsohn and Petrin (2003) and Petrin et al. (2004), which are extensions of the Olley and Pakes (1996) methodology. More technical details on the methodological framework are explained in the Appendix 5.A.

One important problem with firm-level productivity estimation is data related. Missing or non-positive values of investment flow reported in or imputed from micro data is a prevalent challenge in manufacturing firm data (Levinsohn and Petrin, 2003). The issue of lumpy investments is simply due to the typical high fixed cost in manufacturing sectors (i.e. start-up expenditures for machines and infrastructure) and does not allow for the inversion of the investment demand as the function of unobserved productivity. Firms in these sectors tend to invest large amounts of capital for expensive fixed assets when starting their business, but then delay the investment in the next year while the capital stock continues to depreciate. A feasible solution is to use intermediate inputs (materials and services) instead, which are demanded yearly, and can be observed or calculated from the information available in our data (see the description in Table A.1). Modifying the model of Olley and Pakes (1996) that uses investment as the proxy for unobserved productivity, Levinsohn and Petrin (2003) suggest to use the observed yearly smooth demand of the intermediate inputs as an alternative proxy.

---

located at the geographical centroid of each ward. For firm dyads within wards we determine a minimal distance below 2km. Some studies randomly allocate firms within an administrative unit (e.g. Barrios et al., 2012). We abstain from this procedure since Vietnamese wards are sufficiently small units, and due to limited computing power.

We apply the method of Levinsohn and Petrin (2003) and Petrin et al. (2004) to estimate the coefficients of the production function and impute them to calculate firm level productivity, as it provides several advantages over ordinary least square, fixed effects and instrumental variables estimation (Van Beveren, 2012). First, the framework solves for simultaneity issues, and produces a consistent estimator.[20] Second, the data required for intermediate input used as the proxy for unobserved productivity fits well with our data set, as discussed above.[21] Levinsohn and Petrin (2003) apply the approximation in the third order polynomials for the unknown form of productivity shocks. Olley and Pakes (1996) note that either third or fourth polynomials show identical result in their estimations. Assumptions about the timing of the intermediate input choice may be applied to prevent the multi-collinearity of inputs (Ackerberg et al., 2006). We check the multicollinearity among inputs and non-parametric productivity in the actual data and the results reject the hypothesis of Ackerberg et al. (2006).

5.5.2 Results of TFP Estimation

The estimation results of total factor productivity are shown in Table 5.2 in logarithmic form. The results reveal that the mean value of log(TFP) is higher in foreign firms than domestic firms, hence foreign firms feature higher productivity than domestic firms. Additionally, annual growth in TFP differs between foreign and domestic firms: while domestic firms' productivity grew by just 1.4 percent, it was 3.3 percent among foreign firms.

Table 5.2: Summary of Covariates and Estimated Total Factor Productivity

Variables	**Domestics Firms**			**Foreign Invested Firms**			**All Observations**			**Unit**
	N	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	
Value Added	156922	4815	64628	20035	35271	172258	176957	8264	84596	Million VND
Capital Stock	158145	8842	183939	20127	71070	341877	178272	15868	208799	Million VND
No. of Workers	163214	71	285	20255	452	1596	183469	113	606	Workers
Material Inputs	150924	25104	364197	18904	159615	779192	169828	40077	432717	Million VND
Log(TFP)	150301	8.344	2.017	19541	9.369	2.054	169842	8.462	2.048	
Growth in TFP	88,150	.014	.803	14398	.033	.822	102548	.017	.806	%

Notes: Authors' compilation and estimation using data drawn the Vietnam Enterprise Survey 2005–2010. Variables (except for estimated log(TFP)) are in nominal values.

For further investigation of the difference between foreign and domestic firms' total factor productivity distributions, Kernel densities of log(TFP) by year and by firms'

[20]Levinsohn and Petrin (2003) show in detail the advantages of the method over OLS and FE methods.

[21]The method of Levinsohn and Petrin (2003) and Petrin et al. (2004) has been widely applied in the literature. For a review of applications in international trade, see e.g. **?**; a review of applications in research of agglomeration effects is provided in Combes and Gobillon (2015); **?** shows the similarity of TFP calculations by the method of Levinsohn and Petrin (2003), Olley and Pakes (1996) and others; a very recent application of the method is conducted by **?** who uses electricity consumption as a proxy for unobserved productivity.

ownership are presented in Figure 5.3. The figure reveals that for all the years from 2005 through 2010 foreign firms' productivity distribution was consistently shifted towards the right tail, hence higher productivity levels, compared to their domestic counterparts. The mean in log(TFP) (average over all years in Table 5.2) was also consistently higher in all years for foreign firms compared to domestic firms. While the mean of domestic firm's log(TFP) is steadily increasing over time from 8.1 to 8.3, the foreign firms' productivity does not show a steady increase over time (not shown in the figure). This trend might indicate a TFP catch-up of domestic firm towards foreign firms in our study period. Figure 5.3 supports the notion that foreign firms are more productive so that technology and knowledge are more likely to spill over from foreign to domestic firms. In the following part, we explore the causal link between the temporal variation of the presence of foreign firms in proximity of domestic firms and domestic firms' subsequent TFP growth.

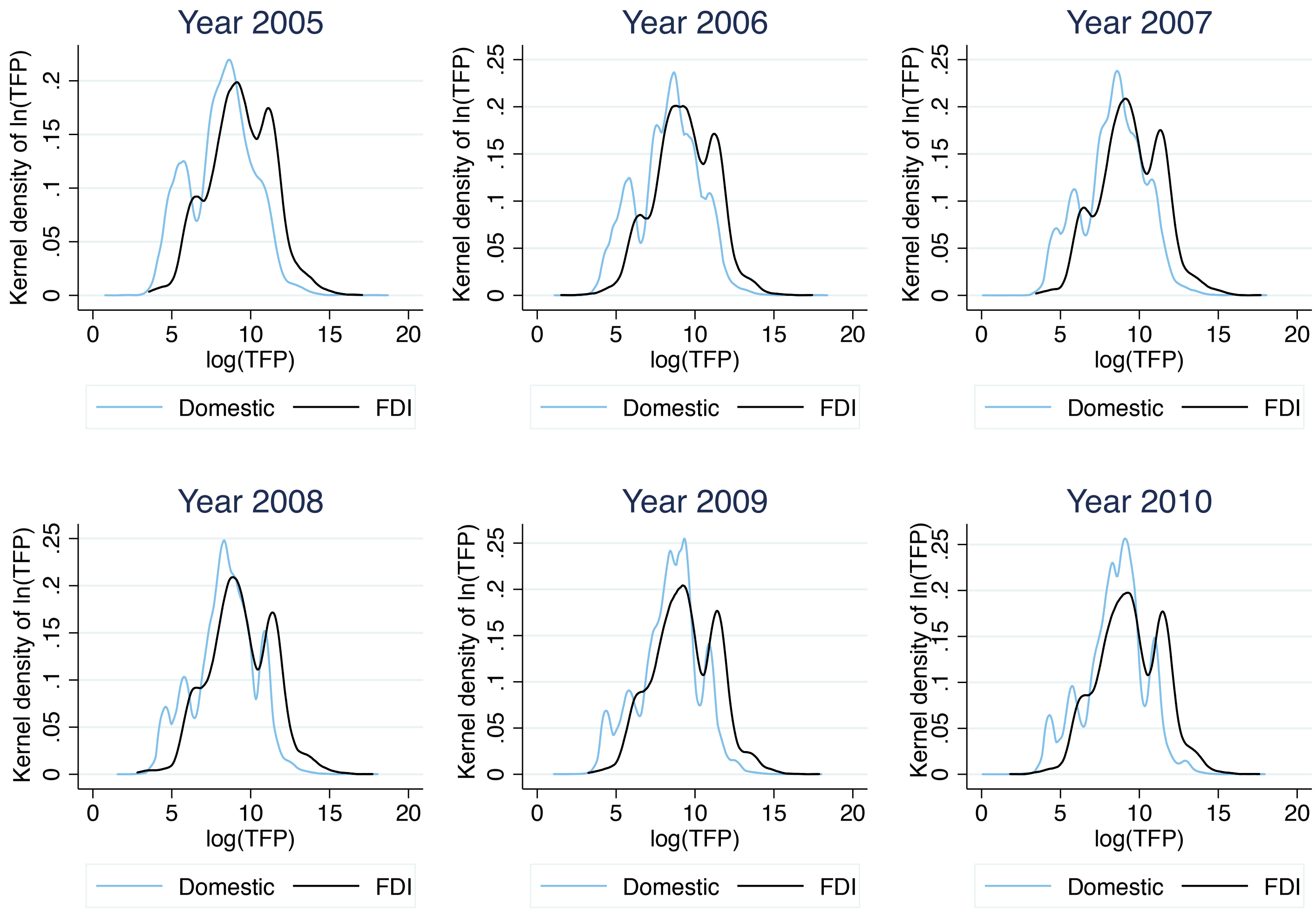


Figure 5.3: Kernel Distribution of Log(TFP) by Year and by Ownership

Notes: The figure displays a Kernel density estimator using a Epanechnikov Kernel function for foreign and domestic firms for each year. Log(TFP) is estimated by the methodology of Levinsohn and Petrin (2003) and based on data of the Vietnam Enterprise Survey, 2005–2010.

5.6 Localised FDI Spillover Effects: Identification Strategy and Results

In this section, we present our identification strategy for FDI spillover effects and the main results. The identification strategy isolates spillover effects of foreign investment on the local manufacturing firms. We apply a 2-stage-least-squares regression at the firm level utilising time variation in the presence of foreign investment in the near surrounding of each domestic firm, and control for possible confounders that may influence both the location of foreign investment and domestic firms' productivity growth. In section 5.6.2 we present a series of results focusing on within industry spillovers. We then disentangle the heterogeneity in effects according to firm size, productivity levels and the productivity gap of local firms to the foreign firms. Spillover effects working through the supply chain, called vertical linkages, are discussed in a separate section 5.6.6. Robustness checks confirm the main results.

5.6.1 Baseline Specification

Our baseline specification presents a causal estimation of the effect of foreign direct investment on total factor productivity growth of domestic firms in Vietnam. We presume that the influence of a foreign firm on a domestic firm is constraint to a geographic space around each local firm. We assess whether the change in presence of foreign invested firms within a specific perimeter of a domestic firm i in year t has a positive (or negative) spillover effect on the local firm's productivity. By varying the spatial extent of the radius around each firm – 2km, 5km, 10km, 20km, and 50km –, we investigate the intensity of spillovers with regard to geographical proximity. The inquiry of location specific spillovers effects restricted to a given radius around each firm i is similarly applied by e.g. Rosenthal and Strange (2008) on human capital spillovers in the US, Halpern and Muraközy (2007) on horizontal and vertical spillovers in Hungary, and Barrios et al. (2012) on R&D spillovers in Ireland.[22]

The most simple assessment of within industry FDI spillover effects on productivity of domestic firms is to estimate the following specification by ordinary least squares:

[22]Classifications of radii by Barrios et al. (2012) are 10km, 20km, 50km, 100km, 200km, and 300 km, respectively. In our study, the maximum radius for which we present results is 50km. Due to the peculiar shape of Vietnam, the support of the data gets unreliable beyond 50km.

$$log(TFP_{ik,t}) \quad = \quad \alpha_i + \delta log(FDI^{RD}_{ik,t}) + \beta log(X^{RD}_{ik,t}) + \varphi HHI_{k,t} + \varepsilon_{ik,t} \qquad (5.1)$$

where the dependent variable is the logarithm of $TFP_{ik,t}$ of domestic firm i in industry k at time t. The variable of interest is measuring the presence of foreign firms of the same industry k in a circle of radius RD around each domestic firm i, denominated $FDI^{RD}_{ik,t}$. We measure the presence of foreign invested firms either by the number or the total output of firms within a circle. In order to interpret the estimated coefficient as elasticity of foreign direct investment on a local firm's productivity, we use the logarithm of $FDI_{ik,t}$. a_i is a firm fixed effect. $X^{RD}_{ik,t}$ is a vector of time varying control variables in logarithms measured for each firm within a circle of radius RD. It includes the local presence of domestic firms in the same industry k, the presence of foreign firms in all other industries, and the presence of domestic firms in all other industries. We hence control for all possible agglomeration economies and spillovers that are not attributable to foreign firms of the same industry. Furthermore, it also contains a variable that measures the size of the labor market, summing up the number of employees within the circle of radius RD. $HHI_{k,t}$ is an indicator for the concentration of an industry, the Herfindahl-Hirschman Index. $\varepsilon_{ik,t}$ is a an error term.

The above specification has one important caveat. According to Olley and Pakes (1996), total factor productivity follows the Markov rule: its current value depends on its past and hence forms an autocorrelation process. Therefore, a simple OLS estimation of the coefficients in the specification above omits one crucial variable, the lagged dependent variable (LDV) of total factor productivity. In the existing literature on spillovers from foreign investment, this Markov process in total factor productivity is often ignored, as e.g. in Barrios et al. (2012) or Anwar and Nguyen (2014). Incorporating the LDV accounts for the AR(1) structure in the data generating process of dynamic total factor productivity at the firm level.

Including the LDV $log(TFP_{ik,t-1})$ in a panel fixed effect estimation with a short time dimension yields, however, a downward bias (Nickell, 1981). By construction, the LDV correlates with the error term. In order to solve this estimation issue, we propose two steps following **?**. First, we estimate the specification in first differences, which eliminates the unobserved firm fixed effect. Moreover we can get rid of the persistent characteristic of the log(TFP) and reduce the problem of serial correlation. Second, we use $ln(TFP_{ik,t-2})$ as an internal instrument for $\triangle ln(TFP_{ik,t-1})$ and estimate the specification by 2-stage-least-squares. Obviously, the following conditions need to hold to consistently estimate this instrumental variable approach:

$$E[\triangle ln(TFP_{ik,t-1})|ln(TFP_{ik,t-2})] \neq 0 \quad (5.2)$$

and

$$E[\triangle \varepsilon_{ik,t}|ln(TFP_{ik,t-2})] = 0 \quad (5.3)$$

The enhanced specification in first differences, our baseline specification, hence is:

$$\begin{aligned} \triangle log(TFP_{ik,t}) &= \rho \triangle log(TFP_{ik,t-1}) + \delta \triangle log(FDI^{RD}_{ik,t}) \\ &+ \beta \triangle log(X^{RD}_{ik,t}) + \varphi \triangle HHI_{k,t} \\ &+ \triangle \phi_t + \triangle \eta_p \times \phi_t + \triangle \varepsilon_{ik,t} \end{aligned} \quad (5.4)$$

where we added time fixed effects, province fixed effects and time-province fixed effects. Industrial policy regulations are mostly determined at the national or provincial government level. By including province-time fixed effects $\rho \times \phi_t$, we take account of the regulatory environment that may change year on year, and regional business cycles. General annual shocks are absorbed by the time fixed effect ϕ_t. Naturally, to obtain a consistent estimation in equation 5.4, the control variables from the equation also need to be orthogonal to the error term $\triangle \varepsilon_{ik,t}$.

Our identification assumption with regard to the main regressor of interest $\triangle log(FDI^{RD}_{ik,t})$ is that a single domestic firm is not decisive for the location choice of foreign invested firms. In other words, we assume that the yearly change of a single domestic firm's TFP is not affecting the change in foreign presence in the surrounding area of a firm. We argue that the problem of endogeneity is unlikely, since it is not possible for a foreign firm to observe the yearly change in a domestic firm's productivity (our dependent variable), and for that reason to select a specific location. Specifically, our assumption is that the change in productivity is only observed by the firm itself but not by other firms. When making investment decisions, foreign firms can investigate the general conditions of the location. The location choice first of all depends on local production conditions such as the local labour market, access to transportation infrastructure, and proximity to forward and backward linked industries. Further discussion about this issue is presented in Section 5.6.3.

Since there is no possibility to run a random experiment by assigning location choices to foreign firms and see how it affects TFP of domestic incumbent firms, we need to determine the factors that are correlated with the location choice of foreign firms and at the same time influence TFP of the domestic firms. By including likely confounding variables, we address these concerns. We control for the change in the presence of other domestic firms, foreign firms of all other industries, and the size of labor market. We are thus able to adjust our coefficient estimates for the attractiveness of a specific location for foreign investment. Furthermore, we present a placebo test in our baseline specification by including the lead of our variable measuring foreign investment within the close surrounding of domestic firms. If there was a selection problem in our specification, then change in productivity should already be higher before foreign firms enter the location, hence show up in the lead, the year before foreign investment takes place. An additional potential confounding factor could be the development of local infrastructure that may both attract foreign investment and improve a local firm's productivity. One may think of new roads or improved internet access that makes an area more attractive for investment. While the province time fixed effect should absorb large scale changes in accessibility, changes in local infrastructure is hard to capture. In order to dispel such concerns, we provide a variant of the basic specification using ward time fixed effects.

5.6.2 Baseline Results

This section presents the results of our estimation of local spillover effects of foreign direct investment on domestic firms' productivity. The section is organised so as to cover various aspects of spillover effects discussed in the literature.

Table 5.3 presents our baseline estimations, by building up step by step our preferred specification. In these first series of regressions we consistently use the number of foreign firms in the vicinity of a domestic firm as underlying measure for our main explanatory variable. To construct it, we simply count the number of foreign firms of the same industry as the domestic firm within a circle of radius RD. Since we are estimating our specification in log differences, we can interpret this variable as growth rate in the presence of foreign firms within a certain area. We are convinced that the number of firms, while not containing any information on the size of firms, is a good indicator of the presence of foreign firms, because it is a rather neutral measure. A priori, one does not really know whether a few large firms convey more spillovers than a large number of small firms.

Panel A presents an ordinary least squares regression of the log of total factor productivity of domestic firms on the log of the number of foreign invested firms in the same industry in first differences, leaving aside any controls. Two important points are revealed.

First, these raw results – while by first differencing is corrected for the unobserved firm fixed effect – show that there is a significant positive correlation between TFP growth of domestic firms and the change in the presence of foreign firms in the close surroundings of these domestic firms. Second, the relationship is strongest for circles with radii of 2 to 10 kilometres, and there seems also to be a clear decaying pattern of spillovers with increasing distance beyond 5 kilometres.

In panel B we add the control variables, accounting for agglomeration forces and factors influencing the location choice of foreign firms. The estimated coefficients slightly decrease in size, while keeping the decaying pattern and their significance. Panel C instruments the lagged dependent variable by the internal instrument $ln(TFP_{ik,t-2})$. Estimated with a 2-stage-least-squares procedure it corresponds to the Anderson-Hsiao estimator (Anderson and Hsiao, 1981) and is our preferred specification.[23] The estimated coefficients are, again, highly significant and still show the pattern of strong within industry localised spillover effects, and the weakening of spillovers beyond 5 to 10 kilometres. The coefficients are only significantly different from zero up to a circle with a radius 20 kilometres. While the effect within a circle of 5 kilometres is substantial at almost 0.3 percentage points higher growth in TFP by an additional percent in the number of foreign firms.

Panel C additionally presents the results for the four most relevant control variables. Remarkably, foreign invested firms other than those of the same industry do not have positive impact on the local economy. In contrast, having more foreign firms close by does have a significantly negative impact on TFP growth of domestic firms. Yet the effect is much smaller and is more than compensated by the positive effect of within industry foreign investment. This interesting results proposes that foreign firms absorb resources when settling into an area. In case the foreign firm is from a different industry there are no positive spillovers and only the negative impact on domestic firms' TFP remains. Other domestic firms do not show spillover effects, both within and across industries. Although we should be able to detect agglomeration spillovers, this restrictive estimation seems to absorb them.

The lagged dependent variable $\triangle Log(TFP_{i,t-1})$ is strongly affecting current TFP growth, supporting our concern of a dynamic autocorrelation process in our dependent variable. The high value of the first stage F-test suggests that the internal instrument is working well.

[23] Adding further lags as instruments for the lagged dependent variable in differences in a GMM framework would increase efficiency. However, due to the unbalanced structure of our data, adding further lags results in losing numerous observations. Since our first stage estimation confirms the strength of the instrument, we stick with the simple version with only one lag. We checked the results using GMM, but the loss in observations due to using additional lags is actually worse than the increase in efficiency. Results are available on request.

Table 5.3: Number of FDI Firms and TFP Growth of Domestic Firms, Baseline Results

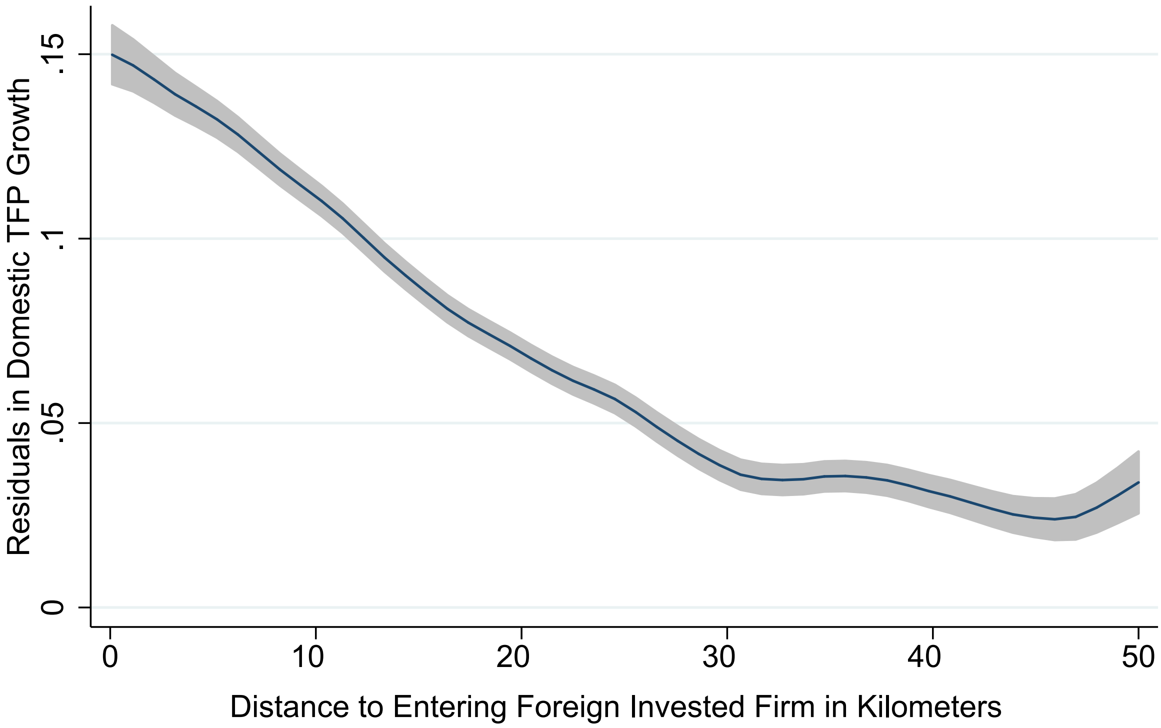
Dep. Var.: $\triangle Log(TFP_{ik,t})$	Circle Radius				
	2km	5km	10km	20km	50km
	(1)	(2)	(3)	(4)	(5)
Panel A: Ordinary Least Squares without controls					
△ Log No. of FDI firms within industry	0.276^{*}	0.296^{**}	0.267^{*}	0.234^{*}	0.188^{+}
	(0.106)	(0.100)	(0.102)	(0.101)	(0.098)
R^2	0.07	0.08	0.08	0.08	0.07
Observations	88150	88150	88150	88150	88150
Panel B: Ordinary Least Squares with controls					
△ Log No. of FDI firms within industry	0.246^{**}	0.273^{**}	0.247^{**}	0.217^{*}	0.165^{+}
	(0.089)	(0.085)	(0.087)	(0.090)	(0.086)
R^2	0.07	0.08	0.08	0.08	0.07
Observations	88150	88150	88150	88150	88150
Panel C: 2SLS, Instrumented Lagged Dependent Variable					
△ Log No. of FDI firms within industry	0.248^{*}	0.297^{**}	0.287^{*}	0.215^{+}	0.154
	(0.101)	(0.101)	(0.108)	(0.108)	(0.100)
△ Log No. of FDI firms in other industries	-0.101^{**}	-0.079^{*}	−0.032	−0.041	-0.072^{*}
	(0.031)	(0.036)	(0.036)	(0.031)	(0.032)
△ Log No. of dom. firms within industry	0.037	0.002	−0.018	−0.005	−0.003
	(0.051)	(0.052)	(0.060)	(0.076)	(0.101)
△ Log No. of dom. firms in other industries	0.000	0.026	0.013	0.042	0.041
	(0.021)	(0.019)	(0.027)	(0.037)	(0.044)
$\triangle Log(TFP_{i,t-1})$	0.531^{***}	0.523^{***}	0.523^{***}	0.529^{***}	0.533^{***}
	(0.035)	(0.036)	(0.037)	(0.037)	(0.037)
Observations	52461	52461	52461	52461	52461
First Stage F-statistic	965.82	916.38	879.93	891.04	903.15
Panel D: Placebo Test, Lead of Change in Foreign Firms					
Lead △ Log No. of FDI firms	0.003	0.002	−0.011	0.002	0.024
	(0.024)	(0.030)	(0.027)	(0.024)	(0.022)
$\triangle Log(TFP_{i,t-1})$	0.443^{***}	0.442^{***}	0.441^{***}	0.442^{***}	0.443^{***}
	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Observations	31162	31162	31162	31162	31162
First Stage F-statistic	474.91	473.66	470.79	470.97	473.01
Time FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Province-Time FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent Variable is annual TFP growth, defined as $log(TFP_{ik,t}) - log(TFP_{ik,t-1})$. Our main explanatory variable is △ *Log No. of FDI firms*, defined as the annual change of the log of number of foreign invested firms within a circle of radius RD. Estimations include a lagged dependent variable $\triangle log(TFP_{i,t-1})$ that is instrumented with its lagged value in levels $log(TFP_{ik,t-2})$. The control variables (Panels B, C, and D) are measured for each firm within a circle of radius RD over time, and calculated as annual change in logarithms. They include the number of domestic firms in the same industry k, the number of foreign firms in all other manufacturing industries, the number of domestic firms in all other manufacturing industries, and the labor force (total number of workers) of all manufacturing industries. Further included is the annual change of the Herfindahl-Index at the industry level. All estimations include time fixed effects, province-time fixed effects, and industry fixed effects. Huber-White standard errors in parentheses are clustered at the industry level. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$.

Panel D presents a placebo test. Instead of the contemporaneous value of FDI, its lead $\triangle log(NOF\ TFP_{ik,t+1})$ is included as main regressor. There seems to be no selection problem in the sense that foreign firms move to places where TFP growth of domestic firms is high in the previous year. This placebo test affirms our well specified estimation procedure.

To illustrate the pattern of spillover effects, Figure 5.4 depicts a local polynomial regression of the residual in TFP growth of domestic firms on their distance to entering foreign invested firms. The residual is based on a regression of TFP growth on industry fixed effects, interaction between time and province dummies, and time fixed effects in order to account for the location specific factors that influence TFP growth. The figure impressively depicts the spillover effects that attenuate with increasing distance. It resembles our baseline results in Table 5.3. The decay in spillovers is very regular and approaches zero at larger distances beyond 30 kilometres.

Figure 5.4: Entering Foreign Invested Firms and Domestic TFP Growth



Notes: The figure presents a kernel-weighted local polynomial regression of domestic firm level TFP growth on the distance to entering foreign invested firms. In gray is a 95% level confidence band. Each observation in the regression is a domestic firm–foreign firm dyad. The residual is based on a regression of TFP growth on industry fixed effects, time-province fixed effects, and time fixed effects. The local polynomial uses an Epanechnikov kernel of degree 0, a bandwidth of 2.96, and pilot bandwidth for calculating the standard errors of 4.44.

Table 5.4 presents the exactly same series of regressions, though using total revenue of foreign firms as the underlying measurement of foreign direct investment. The overall pattern in the results is highly similar. However, estimated spillover effects are weaker and limited to a circle size of radius 5 kilometres. A one percent increase in the change of presence of foreign firms measured by their revenue within a 5 kilometre radius translates into an increase in TFP growth of 0.015 percentage points.

Drawing a preliminary conclusion from our main results, within industry or horizontal spillover effects of foreign direct invested firms seem to be a distinctly local phenomenon. They only occur within limited spatial scope, quickly fading out beyond 5 to 10 kilometres. These results may suggest why most studies that analyse within industry foreign investment spillovers do not find significant results, as e.g. Newman et al., 2015 for the case of Vietnam: their data simply does not allow them to detect such localised horizontal spillover effects.

5.6.3 Does Domestic Firm Size Matter for Spillover Effects?

In this subsection we provide more evidence on the heterogeneity of effects with respect to firm size of local establishments. Table 5.5 presents the results with regard to size of domestic firms measured by the number of workers. The regressions are also based on our baseline specification in first differences and instrumenting the lagged dependent variable. The whole sample of domestic firms is divided into three brackets according to the definition of firm size by the Vietnamese Statistical Office: micro firms with up to 10 workers, small firms have between 10 and 200 workers, and medium and large firms have more than 200 workers.[24]

Remarkably, micro firms seem to especially benefit from the presence of foreign firms in close proximity: Firms with less than 10 employees exhibit the largest coefficients at 0.4 percentage points additional growth in TFP as they are exposed to an additional one percent of foreign firms within 5 kilometres (panel A). Again, one observes a distinct decay of spillovers beyond a 5 kilometre radius. In panel B, the effects are similar although somewhat smaller for firms with 11 to 200 workers employed. For this group of small firms, spillover effects are strongest within 10 kilometres, restricted to a circle radius of 20 kilometres, and they fade out with increasing distance.

The group of medium and large firms is the smallest bracket as there are about 6'000 such firms in our sample. Also for the medium and large firms, the pattern of spillover effects is localised, affirming the robustness in spillover pattern. The effects are slightly increasing up to 10 kilometres, and fading out thereafter. Yet, the effect is only significant within a distance of 5 kilometres, at the 10 percent level. The estimated size of the spillover effect appears to be smaller for these large firms than for the small firms.[25]

To figure out whether this larger effects for small firms actually leads to an absolute convergence in productivity of small and large firms, it is insightful to have a closer look at the size of effects over time. The average TFP level for micro firms is around 2'600

[24] Spillover effects for medium and large firms are jointly estimated since the sample becomes small.

[25] A direct comparison of effects is not possible based on these results since the table presents an separate estimation for each subsample.

Table 5.4: Total Revenue of FDI Firms and TFP Growth of Domestic Firms, Baseline Results

Dep. Var.: $\triangle Log(TFP_{ik,t})$	Circle Radius				
	2km	5km	10km	20km	50km
	(1)	(2)	(3)	(4)	(5)
Panel A: Ordinary Least Squares without controls					
△ Log Tot. Rev. of FDI firms within industry	0.018*	0.020**	0.014+	0.009	0.002
	(0.008)	(0.007)	(0.008)	(0.010)	(0.009)
R^2	0.06	0.06	0.06	0.06	0.06
Observations	88150	88150	88150	88150	88150
Panel B: Ordinary Least Squares with controls					
△ Log Tot. Rev. of FDI firms within industry	0.017*	0.019**	0.012	0.008	0.003
	(0.008)	(0.006)	(0.008)	(0.009)	(0.008)
R^2	0.07	0.07	0.07	0.06	0.06
Observations	88150	88150	88150	88150	88150
Panel C: 2SLS, Instrumented Lagged Dependent Variable					
△ Log Tot. Rev. of FDI firms within industry	0.015+	0.017*	0.012	0.004	0.001
	(0.008)	(0.007)	(0.009)	(0.010)	(0.008)
△ Log Tot. Rev. of FDI firms in other industries	−0.007**	−0.003	0.002	−0.006+	−0.008**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
△ Log Tot. Rev. of dom. firms within industry	−0.004	−0.005	0.000	−0.002	−0.023
	(0.004)	(0.007)	(0.010)	(0.013)	(0.024)
△ Log Tot. Rev. of dom. firms in other industries	−0.004	−0.002	−0.001	0.004	0.024
	(0.004)	(0.004)	(0.005)	(0.008)	(0.021)
$\triangle Log(TFP_{i,t-1})$	0.536***	0.535***	0.537***	0.538***	0.537***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Observations	52461	52461	52461	52461	52461
First Stage F-statistic	992.34	985.62	984.51	992.89	1005.53
Panel D: Placebo Test, Lead of Change in Foreign Firms					
Lead △ Log Tot. Rev. of FDI firms within industry	0.002	0.005+	0.003	0.000	0.004
	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
$\triangle Log(TFP_{i,t-1})$	0.444***	0.444***	0.444***	0.444***	0.443***
	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Observations	31162	31162	31162	31162	31162
First Stage F-statistic	476.74	476.47	475.40	474.97	477.90
Time FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Province time FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent Variable is annual TFP growth, defined as $log(TFP_{ik,t}) - log(TFP_{ik,t-1})$. Our main explanatory variable is △ *Log Tot. Rev. of FDI firms*, defined as the annual change of the log of the total revenue of foreign invested firms within a circle of radius RD . Estimations include a lagged dependent variable $\triangle log(TFP_{i,t-1})$ that is instrumented with its lagged value in levels $log(TFP_{ik,t-2})$. The control variables (Panels B, C, and D) are measured for each firm within a circle of radius RD over time, and calculated as annual change in logarithms. They include the sum of total revenue of all other domestic firms in the same industry k , the sum of total revenue of foreign firms in all other manufacturing industries, the sum of total revenue of domestic firms in all other manufacturing industries, and the labor force (total number of workers) of all manufacturing industries. Further included is the annual change of the Herfindahl-Index at the industry level. All estimations include year fixed effects, province-year fixed effects, and industry fixed effects. Huber-White standard errors in parentheses are clustered at the industry level. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$.

Table 5.5: Number of FDI firms and TFP growth of Domestic Firms, Heterogeneity in Firm Size

Dep. Var.: $\triangle Log(TFP_{ik,t})$	Circle Radius				
	2km	5km	10km	20km	50km
Panel A: Micro firms: Labor force up to 10 workers					
$\triangle$ Log No. of FDI firms within industry	0.386**	0.410**	0.374**	0.251+	0.165
	(0.139)	(0.127)	(0.134)	(0.141)	(0.133)
Observations	14071	14071	14071	14071	14071
First Stage F-statistic	580.33	586.14	556.05	545.06	524.58
Panel B: Small firms: Labor force 11-200 workers					
$\triangle$ Log No. of FDI firms within industry	0.198*	0.255*	0.262*	0.206+	0.147
	(0.097)	(0.099)	(0.102)	(0.105)	(0.104)
Observations	32362	32362	32362	32362	32362
First Stage F-statistic	1486.18	1408.13	1351.90	1356.09	1348.06
Panel C: Medium/large firms: Labor force more than 200 workers					
$\triangle$ Log No. of FDI firms within industry	0.155	0.193+	0.212	0.184	0.184
	(0.101)	(0.106)	(0.151)	(0.134)	(0.120)
Observations	6004	6004	6004	6004	6004
First Stage F-statistic	210.40	209.26	209.68	208.25	209.35
Controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Province time FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent Variable is annual TFP growth, defined as $log(TFP\_ik,t) - log(TFP\_ik,t-1)$. Our main explanatory variable is $\triangle$ *Log No. of FDI firms within industry*, defined as the annual change of the log of number of foreign invested firms in the same industry within a circle of radius RD . Estimations include a lagged dependent variable $\triangle log(TFP\_i,t-1)$ that is instrumented with its lagged value in levels $log(TFP\_ik,t-2)$. The control variables are measured for each firm within a circle of radius RD over time, and calculated as annual change in logarithms. They include the local number of domestic firms in the same industry k , the number of foreign firms in all other manufacturing industries, the number of domestic firms in all other manufacturing industries, and the labor force (total number of workers) of all manufacturing industries. Further included is the annual change of the Herfindahl-Index at the industry level. All estimations include time fixed effects, province-time fixed effects, and industry fixed effects. Huber-White standard errors in parentheses are clustered at the industry level. + $p<0.10$, * $p<0.05$, ** $p<0.01$ *** $p<0.001$.

in the year 2006, while the average for small firms with 11 to 200 workers is 4'300. The average TFP of medium and large firms was 15'700 in 2006. An additional 10 percent in the presence of foreign firms would increase TFP growth for micro firms by 4.1 percentage points, which would yield a TFP level of 3'183 after five years, and hence an absolute gain of 580. For a small firm, the same increase in the presence of foreign firms would yield a TFP level of 5'345. The gain in TFP is 632 and slightly larger in absolute terms after 5 years. We hence cannot expect a closing gap in productivity between micro and small firms due to the presence of foreign firms within 5 kilometres in the medium term. For medium and large firms, a 10 percent increase in the presence of foreign firms would lead to an additional TFP of 1'578, an absolute increase higher than the gain for micro and small firms. We hence conclude that the presence of foreign firms in the close surroundings of domestic firms does not lead to an absolute convergence process in TFP levels between micro, small and large firms in the medium run of 5 years

Overall, the analysis of spillover effects for different domestic firm sizes reveals that micro and small firms appear to benefit from foreign investment even more than medium and large firms in relative terms. The within industry spillover effects are restricted to a small distances, both for large and small firms.

5.6.4 Does the Productivity Level of Domestic Firms Matter for Spillovers Effects?

In this section we look at the heterogeneity in effects with respect to productivity levels of domestic firms within each industry. For each industry, we divide the sample of firms into three groups: below median productivity, third quartile of productivity, and fourth quartile of productivity.[26]

The results in Table 5.6 are astounding. Relatively unproductive firms within each industry seem to specifically benefit from the presence of foreign firms in their vicinity (panel A). Firms at the upper end of the productivity distribution, in contrast, show less pronounced signs of spillover effects. Yet the pattern of decaying spillovers is still detectable (panel C). Domestic firms in the third quartile of the TFP distribution also experience large and significant spillover effects, which are yet somewhat smaller (panel B). These results indicate a convergence process in productivity levels between low and high productivity firms. Unproductive firms indeed seem to be able to absorb know-how from their foreign counterparts, but only if they are sufficiently close-by.[27]

While one would need to look at each industry individually to see if there is absolute convergence in productivity levels, we can still make some calculation at the average of each group of firms. A 10 percent increase in the presence of foreign firms would lead to an absolute growth of TFP of 604 for low productivity firms, 610 for medium productive firms, and 812 for the productive firms over a 5 year period. While, at first glance, these back of the envelope calculations do not point to a quick convergence of productivity levels, unproductive and medium productive firms may increase their productivity almost at the same rate. Furthermore, while the effects for the productive quartile of firms is large in absolute terms, the effects are only marginally significant.

---

[26]Instead of dividing the group into three equally large groups, we decided to separate the unproductive lower half of firms from the third and fourth quartile. It seems more interesting to have a more pronounced picture in the upper half of the productivity distribution.

[27]A direct comparison of effects is not possible based on these results since the table presents an separate estimation for each subsample.

Table 5.6: Spillover Effects for Unproductive and Productive Domestic Firms, 2SLS estimation

Dep. Var.: $\triangle Log(TFP_{ik,t})$	Circle Radius				
	2km	5km	10km	20km	50km
Panel A: Unproductive Firms (below median productivity)					
△ Log No. of FDI firms within industry	0.376**	0.380**	0.323**	0.250*	0.172+
	(0.132)	(0.117)	(0.115)	(0.112)	(0.096)
Observations	23818	23818	23818	23818	23818
First Stage F-statistic	673.31	668.06	650.32	650.39	643.13
Panel B: Medium productive firms (third quartile)					
△ Log No. of FDI firms within industry	0.216*	0.291*	0.348**	0.265*	0.224+
	(0.098)	(0.110)	(0.124)	(0.120)	(0.113)
Observations	14232	14232	14232	14232	14232
First Stage F-statistic	498.40	493.79	499.14	497.48	503.31
Panel C: Productive firms (fourth quartile)					
△ Log No. of FDI firms within industry	0.116	0.205+	0.198	0.134	0.077
	(0.091)	(0.111)	(0.123)	(0.132)	(0.149)
Observations	14398	14398	14398	14398	14398
First Stage F-statistic	430.80	433.84	424.49	420.36	423.06
Controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Province time FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent Variable is annual TFP growth, defined as $log(TFP_{ik,t}) - log(TFP_{ik,t-1})$. Our main explanatory variable is △ *Log No. of FDI firms*, defined as the annual change of the log of number of foreign invested firms within a circle of radius RD. Estimations include a lagged dependent variable $\triangle log(TFP_{i,t-1})$ that is instrumented with its lagged value in levels $log(TFP_{ik,t-2})$. The control variables are measured for each firm within a circle of radius RD over time, and calculated as annual change in logarithms. They include the local number of domestic firms in the same industry k, the number of foreign firms in all other manufacturing industries, the number of domestic firms in all other manufacturing industries, and the labor force (total number of workers) of all manufacturing industries. Further included is the annual change of the Herfindahl-Index at the industry level. All estimations include year fixed effects, province-year fixed effects, and industry fixed effects. Huber-White standard errors in parentheses are clustered at the industry level. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$.

5.6.5 Does the Productivity Gap between Domestic and Foreign Firms Matter?

In this section, we specifically look at how the productivity level of foreign firms is affecting domestic firms' productivity of diverse productivity levels. Do domestic firms benefit more from foreign firms of similar productivity levels or from foreign firms of much higher productivity? To answer this question, we divide the sample into groups of firms according to their TFP level for each industry. We define three groups: below median (low productivity), third quartile (medium productivity), and fourth quartile (high productivity). Both domestic firms and foreign firms are divided under the same TFP distribution for each industry. Since we want to look at the productivity gap within a certain circle of radius RD, we cannot directly calculate a TFP gap to the foreign productivity leader since in many cases, there are no foreign firms within a certain radius at all. The objective of this analysis is to figure out whether the technology gap is important in determining the size of spillover effects.

We perform a series of our baseline regression while including only certain subsamples with specific TFP levels. Figure 5.5 depicts 3x3 graphs with combinations of TFP levels of foreign and domestic firms. The top row shows low productivity domestic firms, while the TFP level of foreign firms increases from left to right: lower half, third quartile, and fourth quartile. The middle row shows domestic firms with medium level (third quartile) TFP, while, again, varying the level of foreign firms' level of TFP. And logically, the third row shows high productivity domestic firms, with increasing levels of foreign firms' productivity level from left to right.

Two results stand out. First, unproductive local firms (row 1) seem to benefit both from rather unproductive foreign firms, but also from very productive firms. Firms are able to learn both from other firms in the same industry that are similar in technology levels, but even more so from firms that are at a advanced technology level. The size in spillover effects is smallest for foreign firms with intermediate productivity levels (middle column). The patterns is similar for medium (row 2) and highly productive (row 3) domestic firms, although with reduced clarity. The productive domestic firms absorb the smallest spillover effects overall, and less significantly so. Second, and more importantly, a small productivity gap leads to a relatively lower learning ability of domestic firms, compared to a large productivity gap. A larger productivity gap between foreign and domestic firms within the same industry appears to facilitate the learning aptitude of domestic firms.

What is common to all combinations of productivity levels of domestic and foreign firms is the diminishing pattern of spillover effects with increasing distance. This, again, supports the robustness of the pattern in our baseline results. A further investigation of the productivity gap could reveal a more clear picture of effects, yet this is beyond the scope of this paper.

5.6.6 Localised Spillover Effects Through Vertical Linkages (Manufacturing and Services): Specification and Results

Vertical linkages between foreign and domestic firms to upstream (forward linked) and downstream (backward linked) industries are another important channel through which spillover effects may work. With regard to FDI presence in vertical linkages, spillover

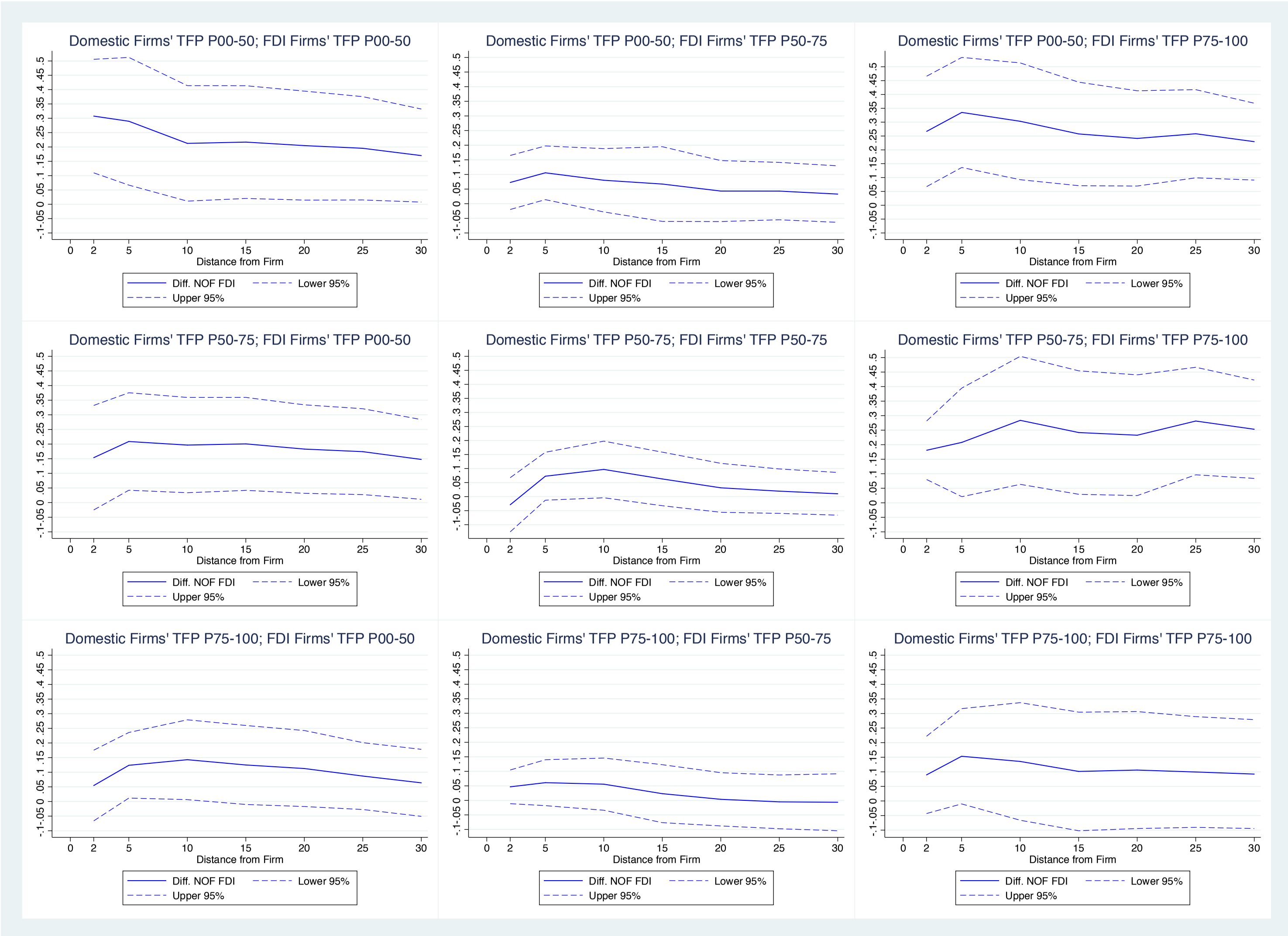


Figure 5.5: Varying Productivity of Foreign Firms Affecting Low and High Productivity Domestic Firms

Notes: The 9 graphs show our baseline regressions of growth in TFP on growth of number of foreign firms within a circle. Each graph shows a distinct combination of TFP levels of domestic and foreign firms. "Domestic firms' TFP P00-50" means the firm the estimation was based on a sample selection of domestic firms with TFP level below the median. TFP P50-75 is the third quartile of TFP levels, and P75-100 is the fourth quartile TFP level of all firms. Denotation for TFP levels of foreign firms is identical. The TFP quartiles are based on the whole sample of domestic and foreign firms.

effects have been extensively studied by various scholars (see e.g. Javorcik, 2004, Halpern and Muraközy, 2007, Anwar, 2011, and Newman et al., 2015). In the following, we assess the relevance of foreign presence in forward and backward linkages, yet in our established spatial framework, by adding these linkages to our baseline regression. Instead of just considering vertical linkages within an industry as done in most existing studies, we investigate vertical linkages to FDI firms in the vicinity of each domestic firms. Due to the decentralised structure of our data, we calculate the absolute value of deflated revenue produced by foreign firms in forward and backward linked industries instead of the output share (Javorcik, 2004) or value added share (Francois and Woerz, 2008). We enhance our baseline specification with the vertical forward and backward linkages as follows:

$$\begin{aligned}\triangle log(TFP_{ik,t}) \;=\; & \rho\triangle log(TFP_{ik,t-1}) + \delta\triangle log(FDI^{RD}_{ik,t}) \\ & +\kappa_1\triangle log(FWL\ Manu^{RD}_{i,fk,t}) + \kappa_2\triangle log(BWL\ Manu^{RD}_{i,kb,t}) \\ & +\kappa_3\triangle log(FWL\ Serv^{RD}_{i,fk,t}) + \kappa_4\triangle log(BWL\ Serv^{RD}_{i,kb,t}) \\ & +\lambda\triangle log(DOM\ Links^{RD}_{ik,t}) + \beta\triangle log(X^{RD}_{ik,t}) + \varphi\triangle HHI_{k,t} \\ & +\triangle\phi_t + \triangle\eta_p * \phi_t + \triangle\varepsilon_{ik,t}\end{aligned} \quad (5.5)$$

where we construct the four variables measuring vertical linkages to foreign firms as follows. The first differenced vertical linkage to foreign invested manufacturing firms in forward linked industries is defined as

$$\triangle log(FWL\ Manu^{RD}_{i,fk,t}) = log(\sum_{j=1}^{N}\alpha_{fk}TR^{RD}_{jf,t}) - log(\sum_{j=1}^{N}\alpha_{fk}TR^{RD}_{jf,t-1}) \quad (5.6)$$

reflecting the annual change in the forward linked foreign manufacturing firms in industries f within a circle of radius RD around each domestic firm i. Each dyad foreign firm j's total revenue TR is weighted by α_{fk}, the coefficient measuring the forward link in the input-output table. α_{fk} measures the amount of goods supplied by forward linked industries f (upstream) to downstream industries k.[28]

[28]The indices reflect three types of industries: k is the industry of domestic firm i itself. Industries f are forward linked industries (upstream), and industries b are backward linked industries. The index fk represents goods or services supplied by industry f to industry k; kb represents goods or services supplied by industry k to industry b. As mentioned before, we use the same definition of forward and backward linkages as Newman et al. (2015): forward linkages are upstream foreign suppliers; backward linkages are downstream foreign customers. This definition of forward and backward is implicitly referring to the perspective of the foreign firm.

The first differenced vertical linkage to foreign invested manufacturing firms in backward linked industries is defined as

$$\triangle log(BWL\ Manu^{RD}_{i,kb,t}) = log(\sum_{i=1}^{N} \beta_{kb} TR^{RD}_{jb,t}) - log(\sum_{i=1}^{N} \beta_{kb} TR^{RD}_{jb,t-1}) \qquad (5.7)$$

where total revenue $TR_{jb,t}$ of each downstream foreign firms j is weighted by β_{kb}, measuring the amount of goods supplied by upstream industry k to downstream industry b.

The remaining two linkages to forward and backward linked service firms are calculated identically as:

$$\triangle log(FWL\ Serv^{RD}_{i,fk,t}) = log(\sum_{i=1}^{N} \alpha_{fk} TR^{RD}_{jf,t}) - log(\sum_{i=1}^{N} \alpha_{fk} TR^{RD}_{jf,t-1}) \qquad (5.8)$$

$$\triangle log(BWL\ Serv^{RD}_{i,kb,t}) = log(\sum_{i=1}^{N} \beta_{kb} TR^{RD}_{jb,t}) - log(\sum_{i=1}^{N} \beta_{kb} TR^{RD}_{jb,t-1}) \qquad (5.9)$$

Specification 5.5 also includes a set of control variables $\triangle log(DOM\ Links^{RD}_{ik,t})$, measuring the presence of forward and backward linked domestic firms. Identically as for the vertical links, we calculate these four types of vertical links for domestic firms.

Table 5.7 presents the results of our baseline specification, but now including the four variables measuring the vertical linkages to forward and backward linked foreign firms, as described above. In general, spillover effects, whether positive or negative, seem to be much more locally restricted in the manufacturing sector compared to the service sector. The coefficients on the measures of linkages to foreign firms in manufacturing sectors are only significant within 10 kilometres. This result confirms the spatially bounded spillover effects among manufacturing industries and is robust for horizontal and vertical linkages. Besides, spillover effects from foreign service firms are quite stable across space and significant also across larger distances.

Our preferred 2SLS estimates are presented in panel C in Table 5.7, including the horizontal linkages from our baseline regression and all control variables. Regarding vertical linkages from FDI manufacturers, our estimations show positive spillover effects from forward linked industries (i.e. foreign manufacturing firms are suppliers to domestic firms). The spillover effects are spatially restricted to within 10 kilometres. This finding is partly in line with the results of Newman et al. (2015) who also find positive spillovers from FDI forward linkages for Vietnamese manufacturing during the period 2006 to 2012 (the research period is comparable to ours), but only when they consider direct forward linkages of upstream foreign to downstream domestic producers.[29] The positive spillovers from upstream foreign firms may be explained by the know-how transfer through the products supplied to domestic downstream firms.

Spillover effects from backward linked foreign firms are negative (foreign manufacturing firms as customers of domestic firms). The foreign firms in downstream industries might have substantial bargaining power and drive down sale prices for domestic firms. As foreign firms in downstream sectors enter the market and choose locations close to the domestic firms, they also potentially absorb a lot of resources, as e.g high skilled employees join technologically more advanced foreign firms in downstream industries. This result is, however, in contrast to the literature: Thang et al. (2016), for instance, find positive backward and negative forward spillovers in their case study of FDI spillovers in the Vietnamese manufacturing sectors for the period 2000 to 2005. Our results are, however, not fully comparable to other studies as we analyse the vertical linkages in our spatial framework, where different mechanisms are supposably at play.

Interestingly, spillovers through vertical linkages of foreign service companies seem much less spatially constraint. Our results record negative spillover effects from forward service linkages, and positive spillover effects for backward linked foreign service firms. Both effects do not provide evidence that spillover effects through vertical service linkages do fade out with increasing distance. This finding is not in line with results in the literature which indicate evidence of positive impacts from foreign service suppliers on the performance of downstream manufacturing (e.g. Francois and Hoekman, 2010). Arnold et al. (2011) find a positive correlation between liberalisation in the service sector and the productivity of downstream manufacturing firms.

Again, our results are not easily comparable to the results in the literature as we look specifically at the presence of foreign firms in vertical linkages within a certain area. The interpretation of our results is, therefore, different from most studies. Most importantly, however, is the fact that the horizontal spillovers are stable and keep being significant after controlling for vertical linkages in panel C of Table 5.7. This supports the robustness of our results on the spatially restricted horizontal spillover effects of foreign direct investment.

[29] Newman et al. (2015) distinguish between direct and indirect vertical linkages. As they have information on direct supplier-customer relationships, they are able to measure direct links between firms, although for a much smaller sample. We construct the vertical linkages by the input-output table and cannot distinguish between direct and indirect linkages. As we look at vertical linkages in close proximity of domestic firms, the probability that a vertical linkage is actually a supplier-customer relationship is increasing, if we assume that closer firms are more likely to trade.

Table 5.7: Vertical Linkages, Total Revenue (nom.) of FDI Firms and TFP Growth of Domestic Firms

Dep. Var.: $\triangle Log(TFP_{ik,t})$	Circle Radius				
	2km	5km	10km	20km	50km
	(1)	(2)	(3)	(4)	(5)
Panel A: Ordinary Least Squares without controls					
△ Log FWL Manufacturing	0.042***	0.057***	0.044**	0.017	0.007
	(0.009)	(0.014)	(0.015)	(0.016)	(0.020)
△ Log FWL Services	−0.082**	−0.100*	−0.116*	−0.120*	−0.103+
	(0.025)	(0.038)	(0.046)	(0.054)	(0.057)
△ Log BWL Manufacturing	−0.032***	−0.033*	−0.025	−0.006	0.003
	(0.009)	(0.013)	(0.016)	(0.020)	(0.027)
△ Log BWL Services	0.101***	0.137**	0.165**	0.163*	0.173*
	(0.029)	(0.047)	(0.060)	(0.066)	(0.072)
R^2	0.07	0.08	0.08	0.08	0.09
Observations	83907	86939	87809	88095	88150
Panel B: Ordinary Least Squares with Horizontal Linkage and Control Variables					
△ Log FWL Manufacturing	0.032***	0.045***	0.042**	0.017	0.022
	(0.008)	(0.012)	(0.013)	(0.016)	(0.018)
△ Log FWL Services	−0.053***	−0.064**	−0.081***	−0.080***	−0.073**
	(0.012)	(0.019)	(0.022)	(0.023)	(0.022)
△ Log BWL Manufacturing	−0.040***	−0.039**	−0.030*	−0.010	−0.012
	(0.009)	(0.012)	(0.014)	(0.015)	(0.019)
△ Log BWL Services	0.064***	0.089***	0.114***	0.110***	0.121***
	(0.015)	(0.024)	(0.029)	(0.028)	(0.027)
△ Log TR of FDI firms (horiz.)	0.013*	0.013**	0.005	0.003	0.000
	(0.006)	(0.005)	(0.005)	(0.008)	(0.008)
R^2	0.09	0.09	0.09	0.09	0.10
Observations	83907	86939	87809	88095	88150
Panel C: 2SLS, Vertical and Horizontal Linkages					
△ Log FWL Manufacturing	0.030**	0.045**	0.041**	0.022	0.036
	(0.009)	(0.014)	(0.015)	(0.020)	(0.025)
△ Log FWL Services	−0.061***	−0.077**	−0.094**	−0.077**	−0.078***
	(0.014)	(0.024)	(0.027)	(0.027)	(0.022)
△ Log BWL Manufacturing	−0.043***	−0.039**	−0.031*	−0.015	−0.020
	(0.011)	(0.013)	(0.015)	(0.015)	(0.021)
△ Log BWL Services	0.071***	0.106**	0.130***	0.102**	0.118***
	(0.017)	(0.031)	(0.036)	(0.033)	(0.029)
△ Log TR of FDI firms (horiz.)	0.013+	0.012*	0.005	−0.001	−0.001
	(0.007)	(0.006)	(0.006)	(0.009)	(0.009)
Observations	50059	51800	52273	52436	52461
First Stage F-statistic	1012.91	1016.88	1021.97	1064.57	1065.70
Time FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Province time FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent Variable is annual TFP growth, defined as $log(TFP_{ik,t}) - log(TFP_{ik,t-1})$. Our main explanatory variables are measuring the presence of foreign firms in forward and backward linked manufacturing and service sectors within a circle of radius RD around a domestic firm. The linkages are calculated as the weighted sum of total revenue of foreign firms in forward and backward linked industries within a circle of radius RD, and are calculated as annual changes in logarithms. The weights are α_{jk}, measuring the supply goods of forward linked industry j to industry k of the domestic firm i, and β_{kj} measuring the supply of goods of industry k of domestic firm i to backward linked industry j. Also included is the horizontal linkage $\triangle Log$ *Tot. Rev. of FDI firms*, defined as the annual change of the log of the total revenue of foreign invested firms in industry k within a circle of radius RD. 2SLS estimations include a lagged dependent variable $\triangle log(TFP_{i,t-1})$ that is instrumented with its lagged value in levels $log(TFP_{ik,t-2})$. The control variables are measured for each firm within a circle of radius RD over time, and calculated as annual change in logarithms. They include the sum of total revenue of all other domestic firms in the same industry k, the sum of total revenue of foreign firms in all other manufacturing industries, the sum of total revenue of domestic firms in all other manufacturing industries, and the labor force (total number of workers) of all manufacturing industries. Further included is the annual change of the Herfindahl-Index at the industry level. All estimations include year fixed effects, province-year fixed effects, and industry fixed effects. Huber-White standard errors in parentheses are clustered at the industry level. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

5.7 Conclusion

This paper investigates spillover effects of foreign direct investment on the local economy with specific focus on the role of proximity between foreign and local firms. A unique data set with detailed geo-referenced firm data allows us go one step beyond existing studies conducted at the regional level. We provide a spatially precise analysis of FDI spillover effects using a large sample of Vietnamese manufacturing firms over the period 2005 to 2010. Exploiting variation in the presence of foreign firms in the vicinity of local establishments, we contribute interesting new evidence of remarkably localised foreign investment spillover effects to the literature.

We find positive and highly significant within industry (horizontal) spillover effects of a highly localised type. Spillovers are strongest within distances of 2 to 10 kilometres and fade out beyond. Our 2-stage least squares estimations, accounting for the dynamics in TFP, estimate elasticities of 0.25–0.4 in TFP growth with respect to changes in FDI presence in the surrounding of domestic firms within a distance of 10 kilometres. Analysing the presence of foreign firms in terms of revenue, elasticities of TFP growth with respect to foreign investment is about 0.017 within a five kilometre radius. These results reflect studies in the agglomeration literature that emphasise the very localised characteristics of agglomeration externalities (e.g. Rosenthal and Strange, 2003). This main result is robust to a variety of concerns. Placebo tests show that our specification is able to control for possible local selection effects of foreign investment. Results are also stable both in the northern and southern regions of Vietnam, and even when ward-time fixed effects are included.

Furthermore, our results indicate that spillover effects are in fact largest for small and relatively unproductive firms. Estimated elasticities are larger and more significant for micro and small firms with a labor force of up to 200 workers. The lower half of firms in terms of industry productivity benefit more than the more productive half. Notably, our results also show that vertical spillovers from foreign manufacturing firms are localised while the spillovers from foreign service firms are not spatially constraint.

5.A Total Factor Productivity Estimation Appendix

As noted by Olley and Pakes (1996), coefficients of capital stocks estimated in the Cobb-Douglas production function by simply using OLS are biased upwards due to the correlation between capital stock and unobserved productivity shocks (TFP). In addition, the authors also indicate that using balanced panel data to estimate TFP, which ignores the entry and exit of firms in the industries, causes a selection bias problem. They argue that the efficient firm, which maximises its "expected discounted value of future net cash flow" in a framework of the Bellman equation, stays in the industry and invests more if its TFP level exceeds a certain threshold. A less efficient firm that has a TFP level below the threshold in contrast, exits the market (Olley and Pakes, 1996).

Olley and Pakes (1996) solve for the selection bias and simultaneity issues in dynamic TFP estimation by using unbalanced panel data and including the survival ratio of a firm in the industry in their estimation. Importantly, they use investment as the proxy for unobserved productivity. They argue that there is a correlation between the choice of the capital stock, investment demand and TFP. Capital stocks are determined at period $t-1$ such that:

$$K_{i,t} = (1-\delta)(K_{i,t-1}) + I_{i,t-1} \tag{5.10}$$

Where δ is the depreciation ratio, $K_{i,t}$ and $K_{i,t-1}$ are respectively the capital stocks in year t and year $t-1$, and $I_{i,t-1}$ is the investment of firm i in year $t-1$.

The investment demand is assumed to be monotonically increasing in TFP. Thus, the demand function of investment can be inverted and investment can be used as proxy for the productivity shock ω. The investment demand is defined as

$$I_t = f_t(K_t, \omega_t) \tag{5.11}$$

With the assumption that $I_t > 0$, after being inverted, we get

$$\omega_t = f_t^{-1}(I_t, K_t) \tag{5.12}$$

Due to the possibility of non-availability or negative values of investment reported in many data sets, Levinsohn and Petrin (2003) develop a theoretical framework based on Olley and Pakes (1996) and suggest to use intermediate inputs as alternative proxies for unobserved productivity shocks instead of investment. Important assumptions made by

Olley and Pakes (1996) are kept in Levinsohn's (2003) model. Moreover, the demand of intermediate inputs chosen needs to be strictly increasing in productivity.

The Cobb-Douglas production function is assumed to be similar among firms in the same industry.

$$VA_{it} = A_{it}K_{it}^{\beta_k}L_{it}^{\beta l} \tag{5.13}$$

Taking logarithms of both sides we have

$$log(VA_{it}) = \beta_0 + \beta_k log(K_{it}) + \beta_l log(L_{it}) + \omega_{it} + \varepsilon_{it} \tag{5.14}$$

In equation (5.14), $log(VA_{it})$ is the logarithm of deflated value added, while $log(L_{it})$ is the logarithm of number of labourers, and $log(K_{it})$ is the logarithm of the real capital stock.[30] ω_{it} is the productivity shock (TFP) we need to estimate. ε_{it} is the error term that is unknown to the firm and the econometrician. ω_{it} is known by the firm when it makes the choice on intermediate inputs and the capital stock, but it is also unobserved by the econometricians.

Rewriting equation (5.14) in lower case, we have

$$va_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it} \tag{5.15}$$

The assumptions implied in equation (5.15) follow Levinsohn and Petrin (2003) and Petrin et al. (2004), such that

(i) The choice of intermediate input m_{it} response to k_{it} and ω_{it}:

$$m_{it} = m(k_{it}, \omega_{it})$$

When a firm gains higher productivity than the threshold and stays in the market, it expands the demand for intermediate inputs, so $m_{it} > 0$, which allows for $m_{it}(k_{it}, \omega_{it})$ to be inverted. Therefore $\omega_{it} = \omega(k_{it}, m_{it})$.

(ii) Labour is not a state variable which means it is demanded when the productivity is realised. In this case, we choose the number of employees as labor input, as we

[30] The Value added production function is popularly utilised in the literature, for instance in Petrin and Levinsohn (2012), De Loecker and Warzynski (2012), and Newman et al. (2015).Petrin et al. (2004) introduce two cases that apply the method by Levinsohn and Petrin (2003): a production function using value added and a production function using gross output.

do not have information on the wage or working hours.[31]

(iii) The first-order Markov process is applied to productivity shocks:

$$\omega_{it} = E[\omega_{i,t}|\omega_{i,t-1}] + \xi_{i,t} \tag{5.16}$$

where $\xi_{i,t}$ is the innovation to productivity.

(iv) Firms are assumed to face the same input and output prices. Hence, in our paper, we estimate TFP by each industry, and assume that within the same industry this assumption holds.

Following Levinsohn and Petrin (2003), we estimate equation (5.15) in two steps by using levpet which is a Stata command written by Petrin et al. (2004). The explanation of the algorithm is as follows:

In the **first step**, making the assumption that

$$\omega_{it} = \omega(k_{it}, m_{it}) \tag{5.17}$$

We have

$$\phi(k_{it}, m_{it}) = \beta_0 + \beta_k k_{it} + \omega(k_{it}, m_{it}) \tag{5.18}$$

Since the form of $\phi_{it}(k_{it}, m_{it})$ is unknown, $\phi_{it}(k_{it}, m_{it})$ is estimated by using a third order polynomial approximation in k_{it} and m_{it}: $\phi(k_{it}, m_{it}) = \Sigma_{n=0}^{3}\Sigma_{j=0}^{3-n}\sigma_{nj}k_t^n m_t^j$.

We rewrite equation (5.15)

$$va_{it} = \beta_l l_{it} + \phi_{it}(k_{it}, m_t) + \varepsilon_{it} \tag{5.19}$$

This first step aims to estimate the consistent coefficient of $_{it}$ in the no-intercept OLS (equation 5.19). It is assumed that $E[\varepsilon_{it}|l_{it}, k_{it}, m_{it}] = 0$.[32]

In the **second step**, coefficients estimated in the first step are used to identify β_k. From equation (5.18), we see that $\widehat{\omega_{it}}$ can also be expressed as

[31] We also check for collinearity of labour with material and capital stock by using the STATA user-written command *collin*. The results of the variance inflation factor ($VIF = \frac{1}{1-R^2}$) which is less than 3 in our case does not indicate a problem of multicollinearity).

[32] Being different from Olley and Pakes (1996), Levinsohn and Petrin (2003) accumulate capital stock by using current investment value $K_{i,t} = (1-\delta)(K_{i,t-1}) + I_{i,t}$.

$$\widehat{\omega_{it}} = \widehat{\phi_{it}} - \beta_k^* k_{it} \tag{5.20}$$

With the grid search, for each β_k^* we can define the appropriate $\widehat{\omega_{it}}$. Using the value $\widehat{\omega_{it}}$ from equation 5.20, Levinsohn and Petrin (2003) approximate $E[\widehat{\omega_{it}|\omega_{i,t-1}}]$ with a third-degree polynomial. With $\widehat{\beta_l}$ derived in the first step, $E[\widehat{\omega_{it}|\omega_{i,t-1}}]$ and β_k^*, rearranging equation 5.13 and combining it with the first-order Markov process, the sample residual of the production function is equal to

$$\widehat{\varepsilon_{it} + \xi_{it}} = va_{it} - \widehat{\beta_l} l_{it} - \beta_k^* k_{it} - E[\widehat{\omega_{it}|\omega_{i,t-1}}] \tag{5.21}$$

The solution to find $\widehat{\beta_k}$ is

$$min_{\beta_k^*} \Sigma_{it} (va_{it} - \widehat{\beta_l} l_{it} - \beta_k^* k_{it} - E[\widehat{\omega_{it}|\omega_{i,t-1}}])^2 \tag{5.22}$$

This yields a consistent estimate of β_k since $E[(\varepsilon_{it} + \xi_{it})|k_{it}] = 0$, and because k_{it} was chosen at time $t-1$ by the accumulation of $k_{i,t-1}$ and $i_{i,t}$ (Levinsohn and Petrin, 2003).

After obtaining consistent coefficients β_k and β_l, the log(TFP), $\widehat{\omega_{it}}$ can be computed as follows (Newman et al., 2015; Olley and Pakes, 1996; Van Beveren, 2012)

$$\widehat{\omega_{it}} = va_{it} - \widehat{\beta_k} k_{it} - \widehat{\beta_l} l_{it} \tag{5.23}$$

Table A.1 specifies how the main variables are constructed using available firm level information from the Vietnamese Enterprise Survey. Value added is calculated by the addition method using firm-level records on profit, wage bills, and indirect tax[33] and depreciation (see Ha and Kiyota, 2014). The depreciation ratio is assumed to be 10%. In addition, different deflators are used to convert the nominal values in the current prices to the base year price which is the year 2000.[34]

[33] The indirect tax is the difference between the total tax paid by the firm and its income tax. See Ha and Kiyota (2014).

[34] Specifically, the producer price index of each industry is the deflator for output and value added. We calculate the index by using the annual producer price index (PPI) by industry provided by the General Statistic Offices of Vietnam (GSO; www.gso.gov.vn). Capital stocks are converted to the base year price by the gross fixed capital formation deflators which are calculated using the annual nominal gross fixed capital formation values of Vietnam provided by the World Bank country database available at www.worldbank.org. Nominal values of materials and services are deflated using the annual GDP deflators downloaded from the World Economic Outlook database available at www.imf.org.

Table A.1: Measurement of Main Variables

Variables	**Measurement**
Total output (Y_{it})	Total revenue (TR_{it}) at the end of year t
Wage(W_{it})	Total wage paid to employees at the end of year t
Labor (L_{it})	Total employees at the end of year t
Capital Stocks (K_t)	Net booked values of fixed assets at the end of year t,
Profit (Π_{it})	Total profit before taxes at the end of year t
Value Added (VA_{it})	$\Pi_{it} + W_{it} + indirecttax_{it} + depreciation_{it}$
Materials and Services (MS_{it})	TR_{it}-$\Pi_{it} - W_{it} - (K_{it} - K_{i,t-1})$
Total Cost(TC_{it})	$TR_{it} - \Pi_{it}$
Depreciation ($Depre_{it}$)	$K_{it} * \frac{depreciationratio}{1-depreciationratio}$

Notes: Authors' compilation using data drawn from the Vietnamese Enterprise Survey 2005–2010.

5.B Data Appendix

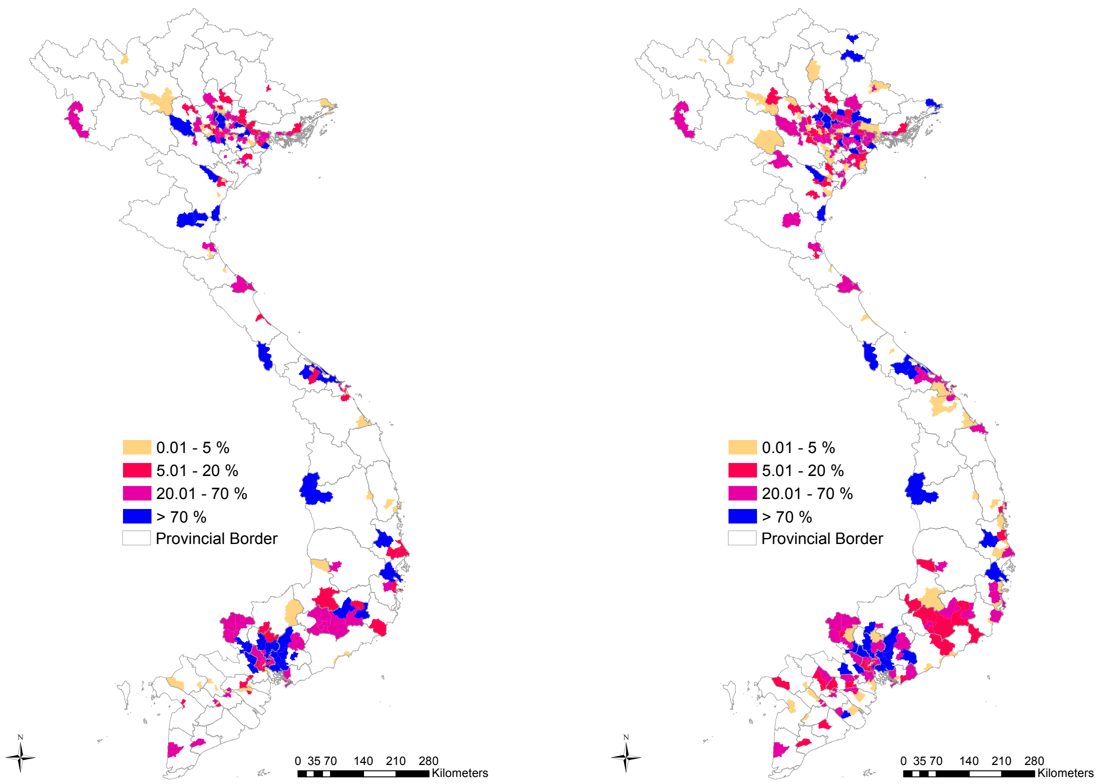
Table B.2: Number of Firms by Industry

Code	Industry	2005	2006	2007	2008	2009	2010	Code	Industry	2005	2006	2007	2008	2009	2010
23	Meat products	94	155	141	179	215	231	54	Medicine, chemical prophylaxis & pharmacy	196	194	210	264	273	291
24	Fishery products	663	762	858	1024	1039	1032	55	By-product rubber	212	218	276	303	346	355
25	Products of vegetables and fruit	185	347	317	407	386	410	56	By-product plastic	1211	1394	1661	2019	2341	2533
26	Vegetable and animals oils and fats	34	34	45	54	60	64	57	Glass and by-product glass	83	84	111	112	119	147
27	Milk products	39	48	57	88	115	106	58	Cements	117	117	122	141	165	188
28	Rice	1207	1177	1115	1128	1145	1012	59	Other non-metallic mineral products	114	143	127	168	198	200
29	Flour (all kinds)	83	97	108	120	142	136	60	Iron, steel, iron	226	243	311	446	483	515
30	Sugar	41	41	37	45	48	51	61	Other metal products	130	138	184	264	320	424
31	Cocoa, chocolate and candy, cake	253	269	321	452	492	530	62	Electronic device, computer and peripheral	23	27	36	62	82	69
33	Other remaining food	1052	1108	1185	1343	1343	1392	63	Machinery & equipment for broadcasting	186	196	228	261	294	332
34	Animal feed	288	320	359	469	472	468	64	Electrical household appliance	74	75	106	180	242	264
35	Alcohol	52	46	58	102	130	130	65	Other electronic & optical products	192	242	251	270	340	361
36	Beer	17	20	17	18	20	17	66	Motor, electric generator, transformers	106	114	137	199	218	244
37	Non-alcohol water and soft drinks	720	756	1081	1380	1551	1586	67	Cell and battery	26	28	32	34	41	33
38	Cigarettes	25	24	25	26	25	24	68	Electric conductor	99	131	142	173	182	181
39	Fiber (all kinds)	370	322	385	470	628	612	69	Electric light equiptment	36	61	51	61	88	109
40	Textile products (all kinds)	292	399	439	536	694	733	70	Consumer electronic equipment	209	231	317	410	458	450
41	Costume (all kinds)	1809	2161	2545	3444	3711	4207	71	Other electric equipments	128	115	92	120	177	181
42	Leather products	202	192	239	292	375	457	72	General-purpose machinery	213	225	254	280	331	377
43	Shoes, sandal (all kinds)	361	366	413	523	554	636	73	Special-purpose machinery	1358	1758	2028	2680	2982	3132
44	Wood products	1489	1851	2158	3094	3493	3558	74	Cars (all kinds)	163	221	220	261	250	261
45	Paper products	943	1075	1190	1509	1650	1734	75	Car engines with tractor (not automotive)	21	41	34	38	42	41
46	Products of printing activities	1176	1650	1803	2253	2854	3338	76	Ships and boats	153	292	237	299	380	343
47	Coke & coal products	4	9	9	12	12	19	77	Motor vehicles, motor bikes	143	152	181	203	201	203
48	Gasoline, lubricants	10	21	15	21	28	37	78	Other transport means	54	65	69	83	70	74
49	Other products from oil, gas	11	17	15	16	15	17	79	Bed, cabinet, tables, chairs	1284	1438	1735	2397	2441	2636
50	Basic organic chemicals	75	74	97	116	159	155	80	Jewelry; instruments; sports, games	146	154	174	230	252	299
51	Fertilizer and nitrogen compound	122	133	169	221	257	285	81	Medical equipments	82	113	123	150	159	166
52	Plastic and primary synthetic rubber	26	77	82	98	106	121	82	Others	1338	1010	1562	1991	3047	3167
53	Other chemical products; fibers	236	277	301	373	389	436		**Total**	**20202**	**23048**	**26595**	**33912**	**38630**	**41110**

Notes: Authors' compilation using data drawn from the Vietnamese Enterprise Survey 2005–2010. Names of industries are shortened, further details are provided on: http://www.gso.gov.vn/default.aspx?tabid=512&idmid=5&ItemID=10752.

Figure B.1: Shares of Total Revenue of Foreign Invested Firms by District in Vietnam 2005 & 2010

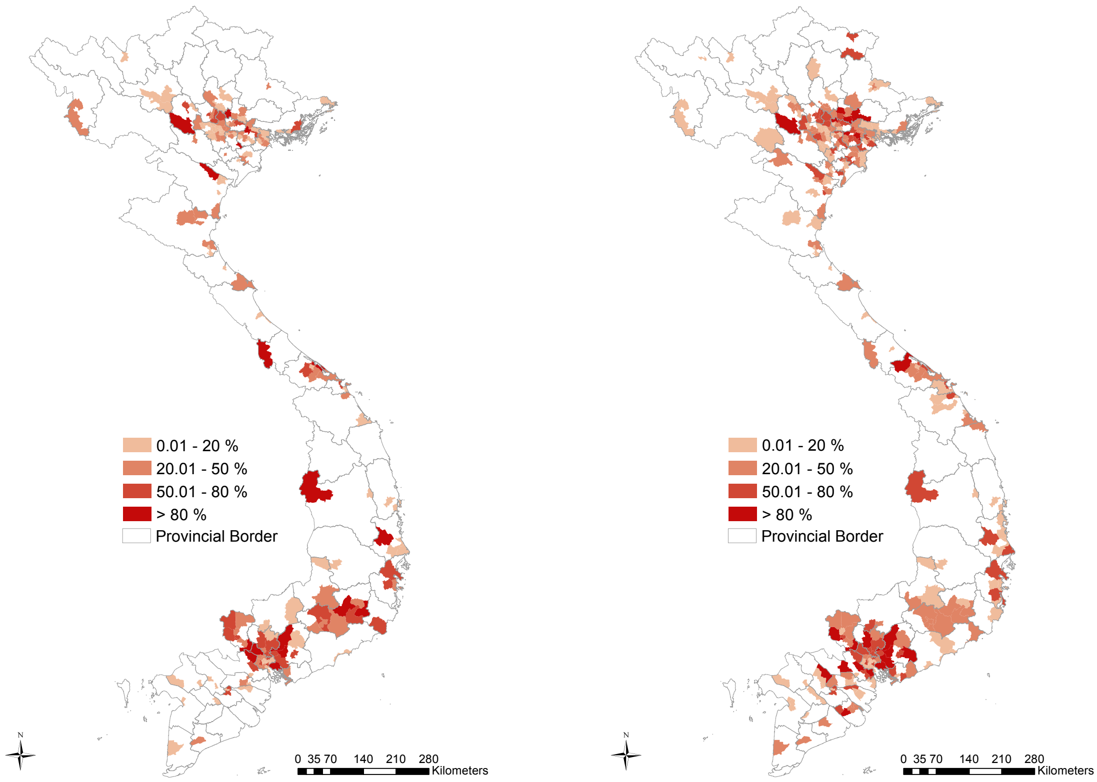
(a) 2005 (b) 2010



Notes: The share of total revenue of foreign invested firms per district is equal to the revenue of foreign invested firms in the district over the total revenue of all firms. The maps are based on authors' calculations using the Vietnam Enterprise Survey 2005 & 2010. Administrative boundaries are based on Global Administrative Areas data (www.gadm.org). Several Vietnamese islands (e.g. Hoang Sa and Truong Sa) are not displayed due to the limitation of the GADM administrative boundaries data.

Figure B.2: Shares of Total Labor Force of Foreign Invested Firms by District in Vietnam 2005 & 2010

(a) 2005 (b) 2010



Notes: The share of total labor force of foreign invested firms per district is equal to the number of workers of foreign invested firms in the district over the total number of workers of all firms. The maps are based on authors' calculations using the Vietnam Enterprise Survey 2005 & 2010. Administrative boundaries are based on Global Administrative Areas data (www.gadm.org). Several Vietnamese islands (e.g. Hoang Sa and Truong Sa) are not displayed due to the limitation of the GADM administrative boundaries data.

Table B.3: Variable Description & Data Sources

Variable	Description	Source
Dependent Variable		
$\triangle log(TFP_{ik,t})$	$log(TFP_{ik,t}) - log(TFP_{ik,t-1})$ Annual growth in total factor productivity	Own calculations based on Levinsohn and Petrin (2003) and Petrin et al. (2004). Source: Vietnamese Enterprise Survey 2005–2010.
Main Explanatory Variables		
Horizonal Linkages:		
$\triangle$Log No. of FDI Firms	$log(\sum_{j=1}^{N} firm_{jk,t}^{RD}) - log(\sum_{j=1}^{N} firm_{jk,t-1}^{RD})$ Annual change in the number of foreign firms j in industry k within radius RD around each firm i	Own calculations based on sample of geo-referenced firms. Source: Vietnamese Enterprise Survey 2005–2010.
$\triangle$Log Tot. Rev. of FDI Firms	$log(\sum_{i=1}^{N} TR_{jk,t}^{RD}) - log(\sum_{i=1}^{N} TR_{jk,t-1}^{RD})$ Annual change in the sum of total revenue of foreign firms j in the same industry k within circle of radius RD around each domestic firm i	Own calculations based on sample of geo-referenced firms. Source: Vietnamese Enterprise Survey 2005–2010.
Vertical Linkages:		
$\triangle$Log FWL Manufacturing	$log(\sum_{i=1}^{N} \alpha_{fk} TR_{jf,t}^{RD}) - log(\sum_{i=1}^{N} \alpha_{fk} TR_{jf,t-1}^{RD})$ Annual change in the forward linked foreign manufacturing firms in industries f within circle of radius RD around each domestic firm i; α_{fk} is the coefficient measuring the link in the input-output table.	Own calculations based on geo-referenced firms of the Vietnamese Enterprise Survey 2005–2010 and the GSO input-output table (2007).
$\triangle$Log BWL Manufacturing	$log(\sum_{i=1}^{N} \beta_{kb} TR_{jb,t}^{RD}) - log(\sum_{i=1}^{N} \beta_{kb} TR_{jb,t-1}^{RD})$ Annual change in the forward linked foreign manufacturing firms in industries b within circle of radius RD around each domestic firm i; β_{fk} is the coefficient measuring the backward link from upstream industry k to downstream industry b in the input-output table.	Own calculations based on geo-referenced firms of the Vietnamese Enterprise Survey 2005–2010 and the GSO input-output table (2007).
$\triangle$Log FWL Services	$log(\sum_{i=1}^{N} \alpha_{fk} TR_{jf,t}^{RD}) - log(\sum_{i=1}^{N} \alpha_{fk} TR_{jf,t-1}^{RD})$ Annual change in the forward linked foreign service firms in industries f within circle of radius RD around each domestic firm i; α_{fk} is the coefficient measuring the forward link from upstream industry f to downstream industry k in the input-output table.	Own calculations based on geo-referenced firms of the Vietnamese Enterprise Survey 2005–2010 and the GSO input-output table (2007).
$\triangle$Log BWL Services	$log(\sum_{i=1}^{N} \beta_{kb} TR_{jb,t}^{RD}) - log(\sum_{i=1}^{N} \beta_{kb} TR_{jb,t-1}^{RD})$ Annual change in the forward linked foreign service firms in industries b within circle of radius RD around each domestic firm i; β_{kb} is the coefficient measuring the backward link from upstream industry k to downstream industry b in the input-output table.	Own calculations based on geo-referenced firms of the Vietnamese Enterprise Survey 2005–2010 and the GSO input-output table (2007).
Control Variables		
$\triangle$Log No. of Dom. Firms	$log(\sum_{d=1}^{D} firm_{dk,t}^{RD}) - log(\sum_{d=1}^{D} firm_{dk,t-1}^{RD})$ Annual change in the number of domestic firms d in the same industry k within radius RD around each firm i	Own calculations based on sample of geo-referenced firms. Source: Vietnamese Enterprise Survey 2005–2010.
$\triangle$Log No. of FDI Firms in Oth. Ind.	$log(\sum_{j=1}^{J} firm_{j,-k,t}^{RD}) - log(\sum_{j=1}^{J} firm_{j,-k,t-1}^{RD})$ Annual change in the number of foreign firms j in all other industries $-k$ within radius RD around each firm i	Own calculations based on sample of geo-referenced firms. Source: Vietnamese Enterprise Survey 2005–2010.
$\triangle$Log No. of Dom. Firms in Oth. Ind.	$log(\sum_{d=1}^{D} firm_{d,-k,t}^{RD}) - log(\sum_{d=1}^{D} firm_{d,-k,t-1}^{RD})$ Annual change in the number of domestic firms d in all other industries $-k$ within radius RD around each firm i	Own calculations based on sample of geo-referenced firms. Source: Vietnamese Enterprise Survey 2005–2010.
$\triangle$Log Labor Force	$log(\sum_{i=1}^{N} L_{i,t}^{RD}) - log(\sum_{i=1}^{N} L_{i,t-1}^{RD})$ Annual change in the labor force measured as number of employees of all firms i in all industries within radius RD around each firm i	Own calculations based on sample of geo-referenced firms. Source: Vietnamese Enterprise Survey 2005–2010.
$\triangle$HHI	$\sum_{i=1}^{N} RS_{ik,t}^2 - \sum_{i=1}^{N} RS_{ik,t-1}^2$ Annual change the Herfindahl-Hirschman-Index, measuring the concentration in industry k; RS is the revenue share of firm i in industry k	Own calculations. Source: Vietnamese Enterprise Survey 2005–2010.

Notes: Additional control variables not shown in the table are forward and backward linkages to domestic firms. They are constructed identically as the forward and backward links to foreign firms.

Table B.4: Descriptive Statistics, Variables 2nd Stage

Variable	Mean	SD	Min	Max
Log TFP	8.50	2.04	-12.73	18.70
No. of FDI Firms same industry within 2km	1.04	3.13	0	41
No. of FDI Firms same industry within 5km	4.19	9.53	0	71
No. of FDI Firms same industry within 10km	12.29	24.58	0	151
No. of FDI Firms same industry within 20km	33.31	60.89	0	307
No. of FDI Firms same industry within 50km	66.74	104.26	0	400
No. of FDI Firms in other industries w. 2km	13.40	29.32	0	250
No. of FDI Firms in other industries w. 5km	57.90	88.00	0	549
No. of FDI Firms in other industries w. 10km	174.93	202.93	0	895
No. of FDI Firms in other industries w. 20km	511.98	540.11	0	1792
No. of FDI Firms in other industries w. 50km	1162.10	1078.68	0	2824
No. of Dom. Firms same industry w. 2km	19.32	39.88	0	374
No. of Dom. Firms same industry w. 5km	82.09	163.78	0	1083
No. of Dom. Firms same industry w. 10km	189.61	342.47	0	1839
No. of Dom. Firms same industry w. 20km	311.16	466.98	0	2177
No. of Dom. Firms same industry w. 50km	442.99	529.75	0	2305
No. of Dom. Firms other industries w. 2km	174.87	256.31	0	1477
No. of Dom. Firms other industries w. 5km	871.89	1215.05	0	5509
No. of Dom. Firms other industries w. 10km	2255.41	2816.27	0	10629
No. of Dom. Firms other industries w. 20km	4198.82	4362.99	0	14007
No. of Dom. Firms other industries w. 50km	6533.28	5449.62	0	16315
Tot. Rev. of FDI Firms same industry w. 2km	100093.29	626292.73	0	31232608
Tot. Rev. of FDI Firms same industry w. 5km	328837.22	1135717.29	0	39137876
Tot. Rev. of FDI Firms same industry w. 10km	957467.25	2242035.04	0	46199108
Tot. Rev. of FDI Firms same industry w. 20km	2882964.91	5081311.32	0	72626352
Tot. Rev. of FDI Firms same industry w. 50km	6484466.72	9947152.11	0	79454896
Tot. Rev. of Dom. Firms same industry w. 2km	266918.53	879812.54	0	56786672
Tot. Rev. of Dom. Firms same industry w. 5km	1085313.87	2491800.48	0	59184680
Tot. Rev. of Dom. Firms same industry w. 10km	2531337.59	4690283.07	0	60782644
Tot. Rev. of Dom. Firms same industry w. 20km	4421958.43	6606477.32	0	63819400
Tot. Rev. of Dom. Firms same industry w. 50km	6845696.40	8274545.71	0	65283704
Forward Link Manufacturing within 2km	37617.66	271070.67	0	17667644
Forward Link Manufacturing within 5km	122793.13	442392.01	0	18044086
Forward Link Manufacturing within 10km	357022.36	755365.05	0	18044148
Forward Link Manufacturing within 20km	1133495.24	1640861.77	0	19821206
Forward Link Manufacturing within 50km	3133920.07	3938335.80	0	22581422
Backward Link Manufacturing within 2km	58218.98	361762.98	0	15738135
Backward Link Manufacturing within 5km	192268.78	629518.34	0	18898828
Backward Link Manufacturing within 10km	594485.27	1336049.96	0	35745540
Backward Link Manufacturing within 20km	1802038.94	3028895.42	0	56167772
Backward Link Manufacturing within 50km	4454358.90	6277958.30	0	61560848
Forward Link Service within 2km	18809.66	100777.43	0	3231617
Forward Link Service within 5km	120686.37	354250.81	0	4279436
Forward Link Service within 10km	371498.34	680569.41	0	4805906
Forward Link Service within 20km	691641.92	924463.45	0	5518592
Forward Link Service within 50km	1127771.62	1222494.12	0	6451718
Backward Link Service within 2km	19549.29	167056.40	0	9920492
Backward Link Service within 5km	127136.05	601172.86	0	13172639
Backward Link Service within 10km	397668.17	1229185.76	0	14228132
Backward Link Service within 20km	700315.62	1750618.28	0	16847164
Backward Link Service within 50km	1058043.97	2302171.41	0	19576840
Total Labor force within 2km	15793.16	20774.29	0	121229
Total Labor force within 5km	71830.16	79005.63	0	357626
Total Labor force within 10km	205165.73	209403.77	0	783254
Total Labor force within 20km	491031.33	477715.79	0	1409818
Total Labor force within 50km	957948.55	806019.73	0	2071620
HHI	429.69	709.47	33	7802
Observations	164349			

5.C Empirical Appendix

Table C.5: Number of FDI Firms and TFP Growth of Domestic Firms, only southern Vietnam

Dep. Var.: $\triangle Log(TFP_{ik,t})$	Circle Radius				
	2km	5km	10km	20km	50km
	(1)	(2)	(3)	(4)	(5)
Panel A: Ordinary Least Squares without controls					
△ Log No. of FDI firms within industry	0.304*	0.331**	0.326**	0.282*	0.217+
	(0.119)	(0.116)	(0.121)	(0.119)	(0.113)
R^2	0.08	0.09	0.10	0.09	0.08
Observations	55500	55500	55500	55500	55500
Panel B: Ordinary Least Squares with controls					
△ Log No. of FDI firms within industry	0.267**	0.316**	0.338**	0.313**	0.210+
	(0.097)	(0.096)	(0.105)	(0.108)	(0.105)
R^2	0.08	0.10	0.10	0.09	0.08
Observations	55500	55500	55500	55500	55500
Panel C: 2SLS, Instrumented Lagged Dependent Variable					
△ Log No. of FDI firms within industry	0.277*	0.376**	0.418**	0.339*	0.226+
	(0.116)	(0.121)	(0.136)	(0.136)	(0.131)
△ Log No. of FDI firms in other industries	−0.122**	−0.117*	−0.109+	−0.078	−0.130*
	(0.042)	(0.059)	(0.063)	(0.055)	(0.060)
△ Log No. of dom. firms within industry	0.039	−0.055	−0.116+	−0.118	−0.072
	(0.064)	(0.063)	(0.069)	(0.087)	(0.132)
△ Log No. of dom. firms in other industries	−0.012	0.035	0.038	0.013	0.024
	(0.033)	(0.030)	(0.041)	(0.049)	(0.053)
$\triangle Log(TFP_{i,t-1})$	0.535***	0.524***	0.519***	0.528***	0.538***
	(0.043)	(0.044)	(0.045)	(0.045)	(0.046)
Observations	32797	32797	32797	32797	32797
First Stage F-statistic	717.16	668.89	630.20	643.14	663.84
Panel D: Placebo Test, Lead of Change in Foreign Firms					
Lead △ Log No. of FDI firms within industry	−0.005	−0.011	−0.011	0.007	0.033
	(0.026)	(0.033)	(0.031)	(0.027)	(0.027)
$\triangle Log(TFP_{i,t-1})$	0.429***	0.428***	0.427***	0.429***	0.430***
	(0.065)	(0.065)	(0.064)	(0.065)	(0.065)
Observations	19750	19750	19750	19750	19750
First Stage F-statistic	324.06	320.94	321.37	320.97	323.58
Time FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Province-Time FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent Variable is annual TFP growth, defined as $log(TFP_{ik,t}) - log(TFP_{ik,t-1})$. Our main explanatory variable is △ *Log No. of FDI firms*, defined as the annual change of the log of number of foreign invested firms within a circle of radius RD. Estimations include a lagged dependent variable $\triangle log(TFP_{i,t-1})$ that is instrumented with its lagged value in levels $log(TFP_{ik,t-2})$. The control variables (Panels B, C, and D) are measured for each firm within a circle of radius RD over time, and calculated as annual change in logarithms. They include the number of domestic firms in the same industry k, the number of foreign firms in all other manufacturing industries, the number of domestic firms in all other manufacturing industries, and the labor force (total number of workers) of all manufacturing industries. Further included is the annual change of the Herfindahl-Index at the industry level. All estimations include year fixed effects, province-year fixed effects, and industry fixed effects. Huber-White standard errors in parentheses are clustered at the industry level. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$.

Table C.6: Number of FDI Firms and TFP Growth of Domestic Firms, only northern Vietnam

Dep. Var.: $\triangle Log(TFP_{ik,t})$	Circle Radius				
	2km	5km	10km	20km	50km
	(1)	(2)	(3)	(4)	(5)
Panel A: Ordinary Least Squares without controls					
$\triangle$ Log No. of FDI firms within industry	0.144**	0.190***	0.131*	0.132+	0.138
	(0.047)	(0.053)	(0.055)	(0.073)	(0.085)
R^2	0.06	0.06	0.06	0.06	0.06
Observations	32650	32650	32650	32650	32650
Panel B: Ordinary Least Squares with controls					
$\triangle$ Log No. of FDI firms within industry	0.135**	0.171***	0.103*	0.093	0.111
	(0.044)	(0.048)	(0.046)	(0.066)	(0.078)
R^2	0.06	0.07	0.06	0.07	0.06
Observations	32650	32650	32650	32650	32650
Panel C: 2SLS, Instrumented Lagged Dependent Variable					
$\triangle$ Log No. of FDI firms within industry	0.138**	0.159**	0.109+	0.076	0.078
	(0.052)	(0.054)	(0.060)	(0.075)	(0.083)
$\triangle$ Log No. of FDI firms in other industries	−0.074**	−0.042+	0.028	0.003	−0.014
	(0.027)	(0.024)	(0.040)	(0.028)	(0.022)
$\triangle$ Log No. of dom. firms within industry	0.026	0.079	0.090	0.120	0.079
	(0.038)	(0.050)	(0.059)	(0.080)	(0.094)
$\triangle$ Log No. of dom. firms in other industries	0.010	0.020	−0.001	0.070	0.055
	(0.021)	(0.024)	(0.033)	(0.046)	(0.059)
$\triangle Log(TFP_{i,t-1})$	0.507***	0.504***	0.506***	0.507***	0.507***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.043)
Observations	19663	19663	19663	19663	19663
First Stage F-statistic	595.35	591.44	589.63	584.41	579.18
Panel D: Placebo Test, Lead of Change in Foreign Firms					
Lead $\triangle$ Log No. of FDI firms within industry	0.039	0.037	−0.015	−0.009	0.005
	(0.047)	(0.038)	(0.036)	(0.044)	(0.038)
$\triangle Log(TFP_{i,t-1})$	0.453***	0.454***	0.455***	0.454***	0.454***
	(0.083)	(0.082)	(0.083)	(0.083)	(0.083)
Observations	11411	11411	11411	11411	11411
First Stage F-statistic	275.87	278.08	275.26	273.90	276.24
Time FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Province-Time FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent Variable is annual TFP growth, defined as $log(TFP_{ik,t}) - log(TFP_{ik,t-1})$. Our main explanatory variable is $\triangle$ *Log No. of FDI firms*, defined as the annual change of the log of number of foreign invested firms within a circle of radius RD. Estimations include a lagged dependent variable $\triangle log(TFP_{i,t-1})$ that is instrumented with its lagged value in levels $log(TFP_{ik,t-2})$. The control variables (Panels B, C, and D) are measured for each firm within a circle of radius RD over time, and calculated as annual change in logarithms. They include the number of domestic firms in the same industry k, the number of foreign firms in all other manufacturing industries, the number of domestic firms in all other manufacturing industries, and the labor force (total number of workers) of all manufacturing industries. Further included is the annual change of the Herfindahl-Index at the industry level. All estimations include year fixed effects, province-year fixed effects, and industry fixed effects. Huber-White standard errors in parentheses are clustered at the industry level. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$.

Table C.7: Total Revenue FDI Firms and TFP Growth of Domestic Firms, including Ward-Time FE

Dep. Var.: $\triangle Log(TFP_{ik,t})$	Circle Radius				
	2km	5km	10km	20km	50km
	(1)	(2)	(3)	(4)	(5)
Panel A: Ordinary Least Squares without controls					
△ Log Tot. Rev. of FDI firms	0.020*	0.022**	0.016	0.012	0.001
	(0.009)	(0.008)	(0.009)	(0.013)	(0.011)
R^2	0.18	0.18	0.18	0.18	0.18
Observations	81640	81640	81640	81640	81640
Panel B: Ordinary Least Squares with controls					
△ Log Tot. Rev. of FDI firms	0.019*	0.020**	0.014	0.011	0.002
	(0.009)	(0.007)	(0.009)	(0.012)	(0.011)
R^2	0.19	0.19	0.18	0.18	0.18
Observations	81640	81640	81640	81640	81640
Panel C: 2SLS, Instrumented Lagged Dependent Variable					
△ Log Tot. Rev. of FDI firms within industry	0.016+	0.019*	0.013	0.003	−0.004
	(0.010)	(0.009)	(0.010)	(0.013)	(0.011)
△ Log Tot. Rev. of FDI firms in other industries	−0.012*	0.003	0.005	0.007	−0.002
	(0.006)	(0.008)	(0.011)	(0.011)	(0.011)
△ Log Tot. Rev. of dom. firms within industry	−0.007	−0.012	−0.005	−0.007	−0.035
	(0.005)	(0.009)	(0.012)	(0.016)	(0.029)
△ Log Tot. Rev. of dom. firms in other industries	−0.034*	−0.074***	−0.077***	−0.021	0.037
	(0.014)	(0.016)	(0.020)	(0.034)	(0.118)
$\triangle Log(TFP_{i,t-1})$	0.543***	0.543***	0.545***	0.546***	0.545***
	(0.043)	(0.044)	(0.043)	(0.043)	(0.043)
Observations	47670	47670	47670	47670	47670
First Stage F-statistic	865.96	864.81	873.16	884.12	896.28
Panel D: Placebo Test, Lead of Change in Foreign Firms					
Lead △ Log Tot. Rev. of FDI firms	0.004	0.006	0.003	0.000	0.005
	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)
$\triangle Log(TFP_{i,t-1})$	0.470***	0.471***	0.470***	0.470***	0.470***
	(0.058)	(0.058)	(0.058)	(0.059)	(0.059)
Observations	27832	27832	27832	27832	27832
First Stage F-statistic	537.74	537.95	535.70	535.98	536.12
Time FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Ward-Time FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent Variable is annual TFP growth, defined as $log(TFP_{ik,t}) - log(TFP_{ik,t-1})$. Our main explanatory variable is △ *Log Tot. Rev. of FDI firms*, defined as the annual change of the log of the total revenue of foreign invested firms within a circle of radius RD . Estimations include a lagged dependent variable $\triangle log(TFP_{i,t-1})$ that is instrumented with its lagged value in levels $log(TFP_{ik,t-2})$. The control variables (Panels B, C, and D) are measured for each firm within a circle of radius RD over time, and calculated as annual change in logarithms. They include the sum of total revenue of all other domestic firms in the same industry k , the sum of total revenue of foreign firms in all other manufacturing industries, the sum of total revenue of domestic firms in all other manufacturing industries, and the labor force (total number of workers) of all manufacturing industries. Further included is the annual change of the Herfindahl-Index at the industry level. All estimations include time fixed effects, ward-time fixed effects, and industry fixed effects. Huber-White standard errors in parentheses are clustered at the industry level. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$.

Bibliography

Abraham, F., Konings, J., and Slootmaekers, V. (2010). FDI spillovers in the Chinese manufacturing sector. *Economics of Transition*, 18(1):143–182.

Acemoglu, D., Aghion, P., and Zilibotti, F. (2006). Distance to Frontier, Selection, and Economic Growth. *Journal of the European Economic Association*, 4(1):37–74.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, 1:385–472.

Acemoglu, D. and Robinson, J. (2008). The role of institutions in growth and development. *World Bank, Washington DC*.

Ackerberg, D., Benkard, L., Berry, S., and Pakes, A. (2007). *The Handbook of Econometrics*, volume 6A. North-Holland, Amsterdam.

Ackerberg, D., Caves, K., and Frazer, G. (2006). Structural identification of production functions. MPRA Paper 38349, University Library of Munich, Germany.

Aharonson, B. S., Baum, J. A. C., and Feldman, M. P. (2007). Desperately seeking spillovers? Increasing returns, industrial organization and the location of new entrants in geographic and technological space. *Industrial and Corporate Change*, 16(1):89–130.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The Economics of Density: Evidence From the Berlin Wall. *Econometrica*, 83(6):2127–2189.

Ahrend, R., Farchy, E., Kaplanis, I., and Lembcke, A. C. (2014). What makes cities more productive? Evidence on the role of urban governance from five OECD countries.

Aitken, B. J. and Harrison, A. E. (1999). Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela. *American Economic Review*, 89(3):605–618.

Alder, S., Shao, L., and Zilibotti, F. (2012). The Effect of Economic Reform and Industrial Policy in a Panel of Chinese Cities. DEGIT Conference Paper, DEGIT, Dynamics, Economic Growth, and International Trade.

Amiti, M. and Davis, D. R. (2012). Trade, Firms, and Wages: Theory and Evidence. *The Review of Economic Studies*, 79(1):1–36.

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

Anderson, T. W. and Hsiao, C. (1981). Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association*, 76(375):598–606.

Anh, N. T. T., Duc, L. M., and Chieu, T. D. (2014). The evolution of Vietnamese industry. WIDER Working Paper Series 076, World Institute for Development Economic Research (UNU-WIDER).

Anwar, S. (2011). Foreign direct investment and export spillovers: Evidence from Vietnam. *International Business Review*, 20(2):177–193.

Anwar, S. and Nguyen, L. P. (2014). Is foreign direct investment productive? A case study of the regions of Vietnam. *Journal of Business Research*, 67(7):1376–1387.

Arnold, J. M., Javorcik, B. S., and Mattoo, A. (2011). Does services liberalization benefit manufacturing firms?: Evidence from the Czech Republic. *Journal of International Economics*, 85(1):136–146.

Arnold, M. (2015). Industrial parks, special economic zones, eco industrial parks, innovation districts as strategies for industrial competitiveness. Technical report, UNIDO Country Office in Vietnam.

Bank, T. W. (2006). Vietnam infrastructure strategy: Cross-sectoral issues. Technical report, The World Bank.

Barrios, S., Bertinelli, L., Heinen, A., and Strobl, E. (2012). Exploring the Existence of Local and Global Knowledge Spillovers: Evidence from Plant-Level Data*. *The Scandinavian Journal of Economics*, 114(3):856–880.

Baum, C. F., Schaffer, M. E., and Stillman, S. (2007). Enhanced routines for instrumental variables/GMM estimation and testing. Boston College Working Papers in Economics 667, Boston College Department of Economics.

Blalock, G. and Gertler, P. J. (2008). Welfare gains from Foreign Direct Investment through technology transfer to local suppliers. *Journal of International Economics*, 74(2):402–421.

Blomstrom, M. and Kokko, A. (2003). The Economics of Foreign Direct Investment Incentives. Working Paper 9489, National Bureau of Economic Research.

Blonigen, B. A., Davies, R. B., Waddell, G. R., and Naughton, H. T. (2007). FDI in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review*, 51(5):1303–1325.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Blundell, R. and Bond, S. (2000). GMM Estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19(3):321–340.

Bodman, P. and Le, T. (2013). Assessing the roles that absorptive capacity and economic distance play in the foreign direct investment-productivity growth nexus. *Applied Economics*, 45(8):1027–1039.

Brandt, L., Van Biesebroeck, J., Wang, L., and Zhang, Y. (2017). WTO Accession and Performance of Chinese Manufacturing Firms. *American Economic Review*, 107(9):2784–2820.

Breinlich, H., Ottaviano, G. I. P., and Temple, J. R. W. (2014). Chapter 4 - Regional Growth and Regional Decline. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 2 of *Handbook of Economic Growth*, pages 683–779. Elsevier.

Brunetti, A., Kisunko, G., and Beatrice, W. (1999). *Institutional Obstacles to Doing Business: Region-by-Region Results from a Worldwide Survey of the Private Sector*. Policy Research Working Papers. The World Bank.

Callois, J.-M. (2008). The two sides of proximity in industrial clusters: The trade-off between process and product innovation. *Journal of Urban Economics*, 63(1):146–162.

Cameron, A. C. and Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372.

Carbonara, E., Santarelli, E., and Tran, H. T. (2016). De jure determinants of new firm formation: How the pillars of constitutions influence entrepreneurship. *Small Business Economics*, 47(1):139–162.

Chhair, S. and Newman, C. (2014). Clustering, competition, and spillover effects: Evidence from Cambodia. Working Paper 2014/065, WIDER Working Paper.

Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, 28(3):591–605.

Clark, G. L., Gertler, M. S., and Feldman, M. P. (2003). *The Oxford Handbook of Economic Geography*. OUP Oxford.

Collard-Wexler, A. (2011). Productivity dispersion and plant selection in the ready-mix concrete industry. *US Census Bureau Center for Economic Studies Paper No. CES-WP-11-25*.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2):723–742.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica*, 80(6):2543–2594.

Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2010). Estimating Agglomeration Economies with History, Geology, and Worker Effects. *Agglomeration Economics*, pages 15–66.

Combes, P.-P. and Gobillon, L. (2015). Chapter 5 - The Empirics of Agglomeration Economies. In Gilles Duranton, J. V. H. a. W. C. S., editor, *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 247–348. Elsevier.

Comin, D. (2010). Total factor productivity. In *Economic Growth*, pages 260–263. Springer.

Crafts, N. and Klein, A. (2015). Agglomeration Economies and Productivity Growth: U.S. Cities, 1880-1930. Technical Report 10673, C.E.P.R. Discussion Papers.

Cung, D. N. (2008). Implementation process of Vietnam's Enterprise Law and Investment Law from Perspective of Regulatory Reform.

De Loecker, J. (2013). Detecting Learning by Exporting. *American Economic Journal: Microeconomics*, 5(3):1–21.

De Loecker, J. and Warzynski, F. (2012). Markups and Firm-Level Export Status. *American Economic Review*, 102(6):2437–71.

Delgado, M., Porter, M. E., and Stern, S. (2015). Defining clusters of related industries. *Journal of Economic Geography*, page lbv017.

Desmet, K., Ghani, E., O'Connell, S., and Rossi-Hansberg, E. (2015). The Spatial Development of India. *Journal of Regional Science*, 55(1):10–30.

Desmet, K. and Rossi-Hansberg, E. (2009). Spatial growth and industry age. *Journal of Economic Theory*, 144(6):2477–2502.

Desmet, K. and Rossi-Hansberg, E. (2010). On Spatial Dynamics*. *Journal of Regional Science*, 50(1):43–63.

Dinh, H. T. (2013). Light Manufacturing in Vietnam. *The World Bank*, page 155.

Dixit, A. (2009). Governance institutions and economic activity. *The American Economic Review*, 99(1):3–24.

Dixit, A. (2015). Governance Reforms and Growth: Some Ideas from Economic Theory. *Frontiers of Economics in China*, 10(4):567–584.

Djankov, S., McLiesh, C., and Ramalho, R. M. (2006). Regulation and growth. *Economics Letters*, 92(3):395–401.

Duranton, G. and Overman, H. G. (2005). Testing for localization using micro-geographic data. *Review of Economic Studies*, 72(4):1077–1106.

duranton, G. and Puga, D. (2000). Diversity and specialisation in cities: Why, where and when does it matter? *Urban studies*, 37(3):533–555.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of regional and urban economics*, 4:2063–2117.

Ellison, G. and Glaeser, E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5):889–927.

Ellison, G. and Glaeser, E. L. (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *The American Economic Review*, 89(2):311–316.

Ellison, G., Glaeser, E. L., and Kerr, W. R. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*, 100(3):1195–1213.

Farole, T. (2011). *Special Economic Zones in Africa: Comparing Performance and Learning from Global Experiences*. The World Bank.

Fazio, G. and Maltese, E. (2015). Agglomeration Externalities and the Productivity of Italian Firms. *Growth and Change*, 46(3):354–378.

Fosfuri, A., Motta, M., and Rønde, T. (2001). Foreign direct investment and spillovers through workers' mobility. *Journal of International Economics*, 53(1):205–222.

Fox, J. T. and Smeets, V. (2011). Does Input Quality Drive Measured Differences in Firm Productivity? Working Paper 16853, National Bureau of Economic Research.

Francois, J. and Davies, R. (2015). Special Tax Treatment as Trade Policy: A Database on Export Processing and Special Economic Zones.

Francois, J. and Hoekman, B. (2010). Services Trade and Policy. *Journal of Economic Literature*, 48(3):642–692.

Francois, J. and Woerz, J. (2008). Producer Services, Manufacturing Linkages, and Trade. *Journal of Industry, Competition and Trade*, 8(3-4):199–229.

Fujita, M. and Thisse, J.-F. (2013). *Economics of Agglomeration: Cities, Industrial Location, and Globalization*. Cambridge University Press.

Gaubert, C. (2018). Firm Sorting and Agglomeration. Working Paper 24478, National Bureau of Economic Research.

Ghosh, S. (2013). Do economic reforms matter for manufacturing productivity? Evidence from the Indian experience. *Economic Modelling*, 31:723–733.

Girma, S., Greenaway, D., and Wakelin, K. (2001). Who Benefits from Foreign Direct Investment in the UK? *Scottish Journal of Political Economy*, 48(2):119–133.

Girma, S. and Wakelin, K. (2007). Local productivity spillovers from foreign direct investment in the U.K. electronics industry. *Regional Science and Urban Economics*, 37(3):399–412.

Giuliano, G., Redfearn, C., Agarwal, A., Li, C., and Zhuang, D. (2005). Not all sprawl: Evolution of employment concentrations in Los Angeles, 1980-2000.

Giuliano, G. and Small, K. A. (1991). Subcenters in the Los Angeles region. *Regional science and urban economics*, 21(2):163–182.

Glaeser, E. L. (1999). Learning in Cities. *Journal of Urban Economics*, 46(2):254–277.

Glaeser, E. L. (2010). Introduction to" Agglomeration Economics". In *Agglomeration Economics*, pages 1–14. University of Chicago Press.

Glaeser, E. L. and Kerr, W. R. (2009). Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain? *Journal of Economics & Management Strategy*, 18(3):623–663.

Görg, H. and Greenaway, D. (2004). Much Ado About Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment? *The World Bank Research Observer*, 19(2):171–197.

Görg, H. and Strobl, E. (2005). Spillovers from Foreign Firms through Worker Mobility: An Empirical Investigation. *Scandinavian Journal of Economics*, 107(4):693–709.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, 118(3):536–598.

Ha, D. T. T. and Kiyota, K. (2014). Firm-Level Evidence on Productivity Differentials and Turnover in Vietnamese Manufacturing. *Japanese Economic Review*, 65(2):193–217.

Halpern, L. and Muraközy, B. (2007). Does distance matter in spillover? *Economics of Transition*, 15(4):781–805.

Heblich, S., Redding, S. J., and Sturm, D. M. (2018). The Making of the Modern Metropolis: Evidence from London. Working Paper 25047, National Bureau of Economic Research.

Howard, E. (2017). Social networks, geographic proximity, and firm performance in Viet Nam. Technical report, World Institute for Development Economic Research (UNU-WIDER).

Howard, E., Newman, C., and Tarp, F. (2015). Measuring industry coagglomeration and identifying the driving forces. *Journal of Economic Geography*, page lbv037.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Huber, S. and Rust, C. (2015). Osrmtime: Calculate Travel Time and Distance with OpenStreetMap Data Using the Open Source Routing Machine (OSRM). *Available at SSRN 2691551*.

Jacobs, J. (1969). The city. The economy of the cities. By Jane Jacobs. Random House, 201 East 50th Street, New York 10022, 1969. 268 pp. *National Civic Review*, 58(9):447–448.

Javorcik, B. S. (2004). Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages. *American Economic Review*, 94(3):605–627.

Javorcik, B. S. and Spatareanu, M. (2011). Does it matter where you come from? Vertical spillovers from foreign direct investment and the origin of investors. *Journal of Development Economics*, 96(1):126–138.

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

Ketels, C., Nguyen, D. C., Nguyen, T. T. A., and Do, H. H. (2010). Vietnam Competitiveness Report. Technical report, Vietnam Centre for Economic Management, Asia Competitiveness Institute, Lee Kuan Yew School of Public Policy (National University of Singapore).

Kokko, A. and Thang, T. T. (2014). Foreign Direct Investment and the Survival of Domestic Private Firms in Viet Nam. *Asian Development Review*, 31(1):53–91.

König, M., Lorenz, J., and Zilibotti, F. (2012). Innovation vs Imitation and the Evolution of Productivity Distributions.

Krugman, P. (1990). Increasing Returns and Economic Geography. Working Paper 3275, National Bureau of Economic Research.

Krugman, P. R. (1991). *Geography and Trade*. MIT press.

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

Lim, E.-G. (2001). Determinants of, and the Relation Between, Foreign Direct Investment and Growth A Summary of the Recent Literature. SSRN Scholarly Paper ID 880230, Social Science Research Network, Rochester, NY.

Lin, P. and Saggi, K. (2007). Multinational firms, exclusivity, and backward linkages. *Journal of International Economics*, 71(1):206–220.

Lychagin, S., Pinkse, J., Slade, M. E., and Reenen, J. V. (2016). Spillovers in Space: Does Geography Matter? *The Journal of Industrial Economics*, 64(2):295–335.

Maiti, D. (2013). Market imperfections, trade reform and total factor productivity growth: Theory and practices from India. *Journal of Productivity Analysis*, 40(2):207–218.

Malesky, E. and Taussig, M. (2009). Out of the gray: The impact of provincial institutions on business formalization in Vietnam. *Journal of East Asian Studies*, 9(2):249–290.

Malesky, E. J. (2010). Provincial Governance and Foreign Direct Investment in Vietnam.

Markusen, J. R. and Trofimenko, N. (2009). Teaching locals new tricks: Foreign experts as a channel of knowledge transfers. *Journal of Development Economics*, 88(1):120–131.

Marshall, A. (1890). *Principles of Economics*. London: Macmillan and Co., Ltd.

Marshall, A. (1920). *Principles of Economics*. London: Macmillan and Co., Ltd.

Mayer, T., Mayneris, F., and Py, L. (2017). The impact of Urban Enterprise Zones on establishment location decisions and labor market outcomes: Evidence from France. *Journal of Economic Geography*, 17(4):709–752.

McCulloch, N. and Malesky, E. (2011). Does Better Local Governance Improve District Growth Performance in Indonesia? *IDS Working Papers*, 2011(369):1–48.

McCulloch, N., Malesky, E., and Duc, N. N. (2013). Does better provincial governance boost private investment in Vietnam? *IDS Working Papers*, 2013(414):1–27.

Neumark, D. and Kolko, J. (2010). Do enterprise zones create jobs? Evidence from California's enterprise zone program. *Journal of Urban Economics*, 68(1):1–19.

Neumark, D. and Simpson, H. (2015). Place-Based Policies. Handbook of Regional and Urban Economics, Elsevier.

Newman, C., Rand, J., Talbot, T., and Tarp, F. (2015). Technology transfers, foreign investment and productivity spillovers. *European Economic Review*, 76:168–187.

Newman, C., Rand, J., Tarp, F., and Howard, E. (2014). Productivity-enhancing manufacturing clusters: Evidence from Vietnam. Working Paper Series UNU-WIDER Research Paper WP2014/071, World Institute for Development Economic Research (UNU-WIDER).

Nguyen, A. N., Nguyen, T., Le, T. D., Pham, Q. N., Nguyen, D. C., and Nguyen, D. N. (2008). Foreign Direct Investment in Vietnam: Is There Any Evidence of Technological Spillover Effects. SSRN Scholarly Paper ID 1117202, Social Science Research Network, Rochester, NY.

Nguyen, H. (2016). Reforms for the Ease of Doing Business in Vietnam: Implications for Total Factor Productivity in Manufacturing Industries. SECO-WTI Working Papers 2016/9, The World Trade Institute, University of Bern.

Nguyen, H. Q. (2017a). Business reforms and total factor productivity in Vietnamese manufacturing. *Journal of Asian Economics*, 51:33–42.

Nguyen, P. A., Phan, T. H., and Simioni, M. (2016). Productivity Convergence in Vietnamese Manufacturing Industry: Evidence Using a Spatial Durbin Model. In Huynh, V.-N., Kreinovich, V., and Sriboonchitta, S., editors, *Causal Inference in Econometrics*, number 622 in Studies in Computational Intelligence, pages 603–619. Springer International Publishing.

Nguyen, T. (2017b). Impact of Trade Liberalization on Poverty in Vietnam.

Nguyen, T. D. K. and Ramachandran, N. (2006). Capital Structure in Small and Medium-sized Enterprises: The Case of Vietnam. *ASEAN Economic Bulletin*, 23(2):192–211. <p>Volume 23, Number 2, August 2006</p>.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6):1417–1426.

North, D. C. (1994). Institutions and productivity in history. *Economics working paper archive at WUSTL*.

Olley, G. S. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6):1263–97.

Orlando, M. J. (2004). Measuring Spillovers from Industrial R&D: On the Importance of Geographic and Technological Proximity. *The RAND Journal of Economics*, 35(4):777–786.

Ottaviano, G. I. and Puga, D. (1998). Agglomeration in the global economy: A survey of the 'new economic geography'. *World Economy*, 21(6):707–731.

Ozimek, A. and Miles, D. (2011). Stata utilities for geocoding and generating travel time and travel distance information. *Stata Journal*, 11(1):106.

Petrin, A. and Levinsohn, J. (2012). Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics*, 43(4):705–725.

Petrin, A., Poi, B. P., and Levinsohn, J. (2004). Production function estimation in Stata using inputs to control for unobservables. *Stata journal*, 4:113–123.

Porter, M. (1990). Competitive Advantage of Nations. *Competitive Intelligence Review*, 1(1):14–14.

Porter, M. E. (2011). *Competitive Advantage of Nations: Creating and Sustaining Superior Performance*. Simon and Schuster.

Redding, S. J. (2010). The Empirics of New Economic Geography*. *Journal of Regional Science*, 50(1):297–311.

Romer, P. M. (1990). Human capital and growth: Theory and evidence. *Carnegie-Rochester Conference Series on Public Policy*, 32(1):251–286.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system gmm in stata. Stata Journal, 9 (1), 86-136. *FOR ONLINE PUBLICATION*.

Rosenthal, S. and Strange, W. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, 64(2):373–389.

Rosenthal, S. S. and Strange, W. C. (2003). Geography, Industrial Organization, and Agglomeration. *Review of Economics and Statistics*, 85(2):377–393.

Royston, P. and Altman, D. G. (1994). Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43(3):429–467.

Smeets, R. (2008). Collecting the Pieces of the FDI Knowledge Spillovers Puzzle. *The World Bank Research Observer*, 23(2):107–138.

Storper, M. (2010). Why Does a City Grow? Specialisation, Human Capital or Institutions? *Urban Studies*, 47(10):2027–2050.

Storper, M. (2013). *Keys to the City: How Economics, Institutions, Social Interaction, and Politics Shape Development*.

Syverson, C. (2004). Product Substitutability and Productivity Dispersion. *The Review of Economics and Statistics*, 86(2):534–550.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 49(2):326–65.

Thang, T. T., Pham, T. S. H., and Barnes, B. R. (2016). Spatial Spillover Effects from Foreign Direct Investment in Vietnam. *The Journal of Development Studies*, 52(10):1431–1445.

The World Bank (2011). *Special Economic Zones: Progress, Emerging Challenges, and Future Directions*. Directions in Development - General. The World Bank.

Topalova, P. and Khandelwal, A. (2010). Trade Liberalization and Firm Productivity: The Case of India. *Review of Economics and Statistics*, 93(3):995–1009.

Topalova, P. and Khandelwal, A. (2011). Trade liberalization and firm productivity: The case of India. *Review of economics and statistics*, 93(3):995–1009.

UNIDO (2011a). Business Registration Reform in Vietnam: A Situation Analysis of The Reform and of UNIDO support. Technical report, United Nation Industrial Development Organization.

UNIDO (2011b). Viet Nam Industrial Investment Report 2011. Technical report, United Nations Industrial Development Organization and Ministry of Planning and Investment, Viet Nam.

Union, E. (2011). *Greenbook*.

Vũ, T. T. A., Lê, V. T., and Võ, T. T. (2007). *Provincial Extralegal Investment Incentives in the Context of Decentralisation in Viet Nam: Mutually Beneficial or a Race to the Bottom?* United Nations Development Programme, Ha Noi, Viet Nam.

Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of economic surveys*, 26(1):98–128.

Van Biesebroeck, J. V. (2004). Robustness of Productivity Estimates. Working Paper 10303, National Bureau of Economic Research.

Voorheis, J. (2015). Mqtime: A Stata tool for calculating travel time and distance using MapQuest web services. *Stata journal*, 15(3):845–853.

Wang, J. (2013). The economic impact of Special Economic Zones: Evidence from Chinese municipalities. *Journal of Development Economics*, 101:133–147.

Wong, M. D. and Buba, J. (2017). Special economic zones : An operational review of their impacts. Technical Report 121958, The World Bank.

Wooldridge, J. M. (2009a). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114.

Wooldridge, J. M. (2009b). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114.

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Koautorenschaften sowie alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.

Bern, October 2018

Huong Quynh Nguyen