AN ABSTRACT OF THE DISSERTATION OF

Hanpil Moon for the degree of Doctor of Philosophy in Applied Economics, presented on September 2, 2010.

Title: Three Essays on International Trade and Regional Productivity

Abstract approved:____

Munisamy Gopinath

A firm's productivity is composed of two parts: pure technical change and location-specific (agglomeration) externalities. Regional productivity is thus an aggregation of productivity of firms producing similar goods and located in a given region. International trade can affect both components of regional productivity. First, trade openness in a closed economy may alter its internal economic geography. Some regions which become more attractive to firms than before gain an advantage over others from integration into global markets. Second, as a competition pressure, trade liberalization forces the least productive firms to exit, resulting in the growth of aggregate productivity in the industry.

The three essays presented in this dissertation explore the relationship between international trade and regional productivity in the presence of heterogeneous firms. In the first essay, a theoretical framework is introduced in order to describe how the above two channels, through which trade affects regional productivity, shape a country's spatial distribution of productivity. Results show that industries, each having its own cost-minimizing location, can be spatially relocated within a country via heterogeneous trade liberalization across industries. Moreover, trade intensifies localization for each industry since most firms in an industry move to or gather around their industry-specific cost-

minimizing location. The consequent clustering of firms generates additional localization economies. More importantly, the intensification of localization economies can slow or delay the selection process, i.e. exit of low productivity firms, following trade liberalization. These findings suggest that trade openness induces significant industrial and spatial dynamics (entry, exit and survival) within an economy.

The second and third essays are empirical tests on the second channel through which trade openness affects regional productivity using county-level data from Korea and firm-level data from India, respectively. In addition to trade liberalization, regional infrastructure is considered to be another competition pressure for domestic firms, i.e. improved infrastructure in a region induces a similar selection process among firms. These empirical essays investigate the effect of falling trade costs and improving domestic infrastructure on the regional variation of raw productivity using a common methodology. That is, a spatial econometric procedure is applied to a production function framework to estimate total factor productivity (TFP) by region and industry, while controlling for potential external and spatial effects. The mean and alternative percentiles of the regional raw productivity distribution are then specified as functions of international and domestic competition indicators. International competition is represented by trade costs, which are estimated as frictions in a gravity-type trade model, while road density is considered to capture the level of a region's infrastructure. In both Korea and India, it is found that trade costs reduction significantly shifted to the right, particularly the 10th percentile value of, the regional productivity distribution. However, a change in the level of infrastructure appears to bring about a higher change in regional productivity relative to a change in the international competition level. Therefore, the relative contribution of trade costs and infrastructure to regional productivity should be evaluated with attention to the costs underlying these options for regional development.

Three Essays on International Trade and Regional Productivity

by

Hanpil Moon

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented on September 2, 2010
Commencement June 2011

Doctor of Philosophy dissertation of Hanpil Moon presented on September 2, 2010

APPROVED:

Major Professor, representing Applied Economics

Director of the Applied Economics Graduate Program

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University Libraries. My signature below authorizes release of my dissertation to any reader upon request.

Hanpil Moon, Author

ACKNOWLEDGEMENTS

An ancient Korean proverb states: “Praise makes even a whale dance.” Written by SeaWorld trainers, *Whale Done!* illustrates this. My advisor Dr. Munisamy Gopinath has lavished praise on me over the last four years, which has motivated me to be a better person. He always kept his office door open for discussion and showed remarkable patience with the many errors I made. This dissertation could not have been completed without the unconditional academic and personal support of Dr. Gopinath. I am deeply grateful to have Dr. Gopinath as my advisor and for the invaluable opportunities he has given me to complete my academic program successfully. Hopefully we will be able to continue our collaboration in the future.

I would also like to take this opportunity to thank my former professor Taeho Lee at Seoul National University (SNU), who initially steered me in the correct direction of learning and opened my eyes to this interesting field, agricultural economics.

I owe my immense gratitude to Professor Hanho Kim at SNU, who not only gave me the great opportunity to study at Oregon State University and to meet Dr. Gopinath but also encouraged me whenever he visited Corvallis.

Special thanks go to Dr. Yong-Taek Kim, Dr. Seong-Jae Park, Dr. Eui-Sik Hwang, and Dr. Joon-Kee Park at Korea Rural Economic Institute (KREI) whose guidance enabled me to gain a great deal of knowledge on practical subjects while I was at KREI. This knowledge undoubtedly added to my academic experience and to the completion of this dissertation.

I must also acknowledge Dr. Bruce A. Weber, Dr. Jeff Reimer, Dr. Christian Langpap, and Dr. Russell E. Ingham for serving on my committee.

I would like to thank all the staff in the department of Agricultural & Resource Economics for their support and collegial atmosphere. In particular, I am indebted to my classmates, Cyrus, Ned, Paul, Smita, Chun-Kwon, and Tae-Young for their help and companionship during my stay at OSU.

I extend my gratitude to my parents back in Korea Won-Ho Moon and Myoung-Hee Kim for the love and support they have given me throughout my life, and for making me the person I am today. I thank my beloved brother Dopil Moon for providing financial and academic support.

To my wife, Young-Joo Shin, I thank you for your relentless love and encouragement. It was because of your support that I could contribute most of my time to this work. Without your patience and devotion, I could not have got to the end of this journey. Finally, I also thank “my dream” Jangwoo and “my angel” Sophia for being well grown up and for gifting me with biggest smiles every day.

CONTRIBUTION OF AUTHORS

Dr. Munisamy Gopinath contributed to the preparation of the manuscripts, “Foreign Competition, Domestic Infrastructure and Manufacturing Productivity in Korea” (Chapter 3) and “Trade Costs and Regional Productivity in Indian Manufacturing” (Chapter 4). Dr. Hanho Kim assisted with data collection and involved with the design of Chapter 3. In particular, Dr. Gopinath contributed to the writing of the Introduction and Discussion sections in Chapter 3 and Chapter 4. His direction was instrumental to developing the methodology, interpreting results, and polishing the manuscript.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF TABLES

THREE ESSAYS ON INTERNATIONAL TRADE AND REGIONAL PRODUCTIVITY

CHAPTER ONE

Introduction and Overview

Recent literature on firm heterogeneity provides both the theoretical basis and empirical evidence that only a few firms whose productivity is highest within an industry enjoy the benefits from exposure to trade. The least productive firms are often forced to exit the industry following trade liberalization. In the spatial dimension within a country, some regions where countrywide resources newly concentrate appear to benefit more from a liberalized trade regime. Therefore, firms prefer to locate in more attractive regions, which in turn affects aggregate productivity in those regions. This circularity shapes countrywide spatial distribution of firms in an industry, and naturally productivity and income vary across regions and industries within a country as well as across countries.

This dissertation treats a firm's productivity as being composed of two components: pure technical change and location-specific externalities, i.e., agglomeration forces. Regional productivity for an industry is then defined as an aggregate productivity of firms producing similar goods and being located in the same region. Setting aside the possibility that international trade enhances individual firms' raw productivity (e.g. learning-by-exporting), this dissertation explores how trade openness affects both components of regional productivity without changing individual firms' productivity. First, trade openness in a closed economy may alter its internal economic geography. Some regions which become more attractive to firms than before gain an advantage over others from integration into global markets. Second, as a competition pressure, trade liberalization forces the least productive firms to exit and thus, reallocates the detached (floating) resources to more productive firms, resulting in growth of aggregate productivity in the industry. The above arguments are key predictions of the emerging heterogeneous-firms trade theory and its applications (Melitz 2003, Bernard et al. 2003).

This dissertation's three essays explore the relationship between international trade and regional productivity in the presence of heterogeneous firms focusing on the consequent industrial and spatial dynamics (entry, exit and survival). The first essay (Chapter 2), *Free Trade and Localization Economies with Heterogeneous Firms*, introduces a theoretical framework in order to describe how the above two channels, through which trade affects regional productivity, shape a country's spatial productivity distribution. More specifically, it studies the impacts of trade liberalization on firms' optimal location and localization economies, by linking a firm-heterogeneity trade model with a simplified core-periphery structure from new economic geography. Results show that industries, each having its own cost-minimizing location, can be spatially relocated within a country via heterogeneous trade liberalization across industries. Moreover, trade intensifies localization for each industry since most firms in an industry move to or gather around their industry-specific cost-minimizing location. The consequent clustering of firms generates additional localization economies. More importantly, the intensification of localization economies can slow or delay the selection process, i.e. exit of low productivity firms, following trade liberalization. These findings suggest that trade openness induces significant industrial and spatial dynamics (entry, exit and survival) within an economy.

The other two essays are empirical tests on the second channel through which trade affects regional productivity in the case of Korea and India. Regional infrastructure is considered to be another competition pressure for domestic firms in these empirical tests. For instance, improved infrastructure in a region brings about a selection process similar to that induced by a reduction of international trade costs. The second essay (Chapter 3), *Foreign Competition, Domestic Infrastructure and Manufacturing Productivity in Korea*, investigates the effect of foreign competition and domestic infrastructure on regional productivity variation in Korean manufacturing. For this purpose, data on manufacturing output and inputs for 231 counties during 1999-2006 are assembled. A spatial econometric procedure applied to a production function framework provides estimates of total factor productivity (TFP) by counties, while controlling for potential spatial dependence among counties and agglomeration effects. The changes in regional productivity distribution are examined by specifying its alternative percentile values as a function of changes in international competition levels and domestic infrastructure. International competition is represented by trade costs, which are estimated as frictions in a gravity-type trade model, while road density is considered to capture the level of a region's infrastructure. Consequently, the relative contribution of trade costs and infrastructure to regional productivity are evaluated. Results suggest that infrastructure improvement and trade costs reduction significantly shifted to the right, particularly the 10^{th} percentile value of, the regional productivity distribution. For all percentiles, gains in productivity associated with infrastructure improvements appear to outweigh those arising from lowering international trade costs. However, policy choices must consider the costs underlying these options for regional development.

The third essay (Chapter 4), *Trade Costs and Regional Productivity in Indian Manufacturing*, addresses the same general question posed in the second essay (the effects of falling trade costs, i.e. policy and geographic barriers, and improving domestic infrastructure on the regional variation in raw productivity), but employs a firm-level panel database for nine manufacturing industries in India. A firm-level production function using spatial econometric techniques is estimated with data on 8,462 Indian manufacturing firms during 1994-2007. Raw productivity distribution of firms in each combination of industry, region, and year is then derived. Finally, Tobit and censored least absolute deviations (CLAD) estimators are used to quantify the effects of foreign and domestic competition on the median and alternative percentiles of the regional raw productivity. The findings include an obvious synergy between improved infrastructure in surrounding regions and trade liberalization in enhancing raw productivity level. Competition-induced productivity growth is observed across industries and regions. In the context of overall manufacturing, a change in the level of infrastructure relative to that in international competition appears to bring about a higher change in regional productivity.

The three essays together improve our understanding of the relationships among trade, agglomeration and regional productivity. The results have important policy implications, especially by way of insights into how trade liberalization can be combined with regional development policies, such as improving infrastructure and encouraging export participation, to improve domestic firms' competitiveness in global markets.

CHAPTER TWO

Free Trade and Agglomeration Economies with Heterogeneous Firms

2.1. Introduction

Trade liberalization alters the nature of competition within as well as across countries (Helpman 2006). In the intra-country context, for example, urban regions may gain an advantage over others from the integration into global markets. Aside from the comparative advantage arguments, the urban lead can be inferred from the accumulation of production resources (capital and labor) in a few large cities. The question then is whether or not trade liberalization intensifies economic concentration in the core and contributes to spatially uneven economic development? The purpose of this study is to identify the effects of trade liberalization on spatial concentration of economic activity and the consequences for regional economic development.

With the remarkable success of several Asian economies, most economists have fervently advocated an open trade regime as an important ingredient to economic growth (Giles and Williams 2000). However, such a strategy may have coincided with an increase in spatial and social income inequality for some countries. The idea that trade liberalization affects the internal economic geography of a country dates back at least to Krugman (1991) and Krugman and Elizondo (1996), who argue that a low degree of trade openness in an economy creates a tendency for the spatial concentration of manufacturing activities.[1] In a more open economy, firms are less dependent on domestic markets and local sources of supply, and hence, within city agglomeration forces are weaker. Ades and Glaeser (1995) empirically find a linkage between openness and agglomeration using data from the largest cities of 85 countries. However, Ades and Glaeser (1995) cast doubt on the causality in this linkage and conclude that politics is more important in determining urban primacy than trade openness. Hanson (1996) also finds empirical evidence supporting Krugman and Elizondo's (1996) result in the case of Mexico following its trade liberalization with the United States. Alonso-Villar (2001) finds that more industrialization in a closed economy would lead to spatial concentration of activity because agglomeration forces become stronger than dispersion ones. On the other hand, at low levels of industrialization, an open economy would have a small industrial region in its center because foreign competition makes peripheral locations unprofitable (Venables 2005). Further industrial development would lead the inward oriented mono-centric economy to the export oriented multi-centric one, confirming the deconcentration effect of trade shown by Krugman and Elizondo (1996).[2]

These earlier studies, largely belonging to the 'new economic geography (NEG)' literature, have two implications.[3] First, increasing returns to scale in production implies

---

[1] Krugman (1991) shows that high tariff barriers, low share of international trade in GDP, and high costs of domestic transportation increase the degree of concentration in a single city.

[2] Some manufacturing clusters would be located in the peripheral (border) regions or trade ports.

[3] Head and Mayer (2004) provide key characteristics of a standard NEG model. First, there are increasing returns to scale in production, and these are internal to the firm. Second, the existence of increasing returns generates market power. Third, there are costs of trading over geographic space. Fourth, firms have the ability to choose their locations. Finally, there is endogenous location of demand, either through mobility of households, or via the effect on the demand for intermediate goods that occurs through the location decisions of downstream firms.

that a particular variety of good will be produced in a limited number of locations, because spreading out production over multiple places would mean that scale economies would go unexploited. Second, the concentration models embody some tension between the centripetal forces that pull population and production leading to agglomeration and the centrifugal forces that draw them apart as in dispersion (Baldwin et al. 2003, Combes et al. 2008). Centripetal forces include both pure external economies and a variety of market size effects, e.g. forward and backward linkages and transportation costs, while centrifugal forces arise from pure external diseconomies such as congestion, high land rents, and the attraction of less competitive and dispersed rural markets.

The NEG literature modeling the self-reinforcing character of spatial concentration commonly assumes identical or homogeneous firms within an industry (Fujita et al. 1999, Baldwin and Okubo 2006).[4] Recent studies show that firms within a narrowly defined industry differ in terms of capital and skill intensity, size, and productivity (Bernard and Jensen 1999, Melitz 2003, Helpman et al. 2004, Bernard et al. 2007). In an influential paper, Melitz (2003) shows that trade liberalization in a country induces not only its high-productivity firms to enter foreign markets but also its low productivity firms to exit the domestic market. The results include a shift in the industry's productivity distribution to the right, which increases its average productivity, and a countrywide selection effect, i.e. elimination of the least productive firms within a country. However, the intra-industry trade models with heterogeneous firms do not consider firms' spatial relocation decisions

[4] The central idea of NEG is cumulative causation and the consequent agglomeration or clustering of activities. Agglomeration can arise if there is labor mobility and input-output linkage; a large market creates jobs and the expenditure of these workers makes the market large, i.e. firms create the market for other firms (Krugman 1991, Venables 1996).

when exposed to trade. Moreover, firms' productivity may depend on the spatial density of economic activity, and knowledge transfer may rely on distance from the core, where investments on technology occur.

In this study, I will employ firm heterogeneity of the type in Melitz (2003) to extend a model of firms' location decision from the NEG literature. A closely related effort can be found in Baldwin and Okubo (2006), who extend the footloose capital (FC) model to allow for exogenous heterogeneity in firms' marginal costs, and show that the most efficient firms move to the big country under free trade, because efficient firms have higher sales and thus, enjoy greater savings on trade costs which outweigh the loss from the extra local competition in the big market.[5] The current study, however, differs from Baldwin and Okubo (2006) in several ways. First, I focus on firms' location decisions within a country, to analyze the effect of trade on intra-country spatial distribution of economic activities. Second, I consider a joint distribution of productivity and spatial location of firms within an industry and disentangle location-embedded productivity from raw productivity. By doing so, I can integrate NEG ideas with the Melitz (2003) framework, instead of adopting a stylized NEG model. Finally, I regard the spatial distribution of economic activity in a country as the equilibrium outcome operated through a well-known urban system.

The structure of this chapter is as follows. In the next section, I describe the core-periphery theoretical structure, followed by a model of firm location in section 2.3. In Section 2.4, I discuss the impact of trade liberalization on firm location and localization economies. Section 2.5 provides a summary and conclusions.

---

[5] Baldwin and Okubo (2006) refer to this as a "selection effect."

2.2. Theoretical Background

2.2.1. A Core-Periphery Structure

I begin by outlining a spatial structure comprising core and peripheral areas. In this structure, there exist forces shaping the spatial distribution of economic agents, people and firms. The well-functioning economic center, i.e. primacy in the core, provides most generous public services in the economy and so, attracts most of economic agents and activities. The structure supports a stylized fact that higher wages are obtained in the core rather than in the periphery. The centrifugal forces push some firms to break away and locate toward peripheral areas, but these are typically overcome by the benefits generated by increasing returns and transportation costs for being inside or near the geographic core. Here, for the sake of simplicity, I assume a small economy (country) with an economic center and composed of finite number of spatial units (regions) with the same size. Each unit can be distinguished in terms of a key geographic variable, distance from the economic center, d, and other regional characteristics labeled by a vector R.[6] The labor force (employed persons), e, which is assumed to be proportional to population and perfectly mobile across spatial units, then can be defined as a negative function of distance from the economic center:

(1) $$e_i = e(d_i\,;\ R_i, TL), \quad e^{'} < 0, \qquad \forall\, i = 1,2,\cdots,I$$

where, TL is an index of trade liberalization faced by this economy, implying that the pattern and degree of overall trade regime in a country also affects the shape of this population function. For instance, if most of trade activities take place in the economic

---

[6] Note that d_i is a time-invariant variable, but and R_i is not.

center, e_i would be steeper near the core while be flatter near peripheral areas than that in autarky since trade facilities and related services create jobs and thus attract labor from countryside.

Next, the wage in a region is represented to be increasing in the size of its labor force, reflecting the existence of urban agglomeration externalities;

$$(2) \qquad w_i = w(\, e_i), \quad w' > 0, \qquad \forall\, i = 1,2,\cdots,I.$$

This upward-sloping wage curve stands in sharp contrast with conventional wage curves that slope downwards. However, most empirical studies in urban economics show that measures of productivity per capita increase with city size (Sveikauskas 1975, Segal 1976, Moomaw 1981, Henderson 1986, Ciccone and Hall 1996). In turn, a higher productivity in larger cities can explain why a disproportionate share of economic activity takes place in a small number of places rather than spreading uniformly over space as would be predicted by traditional models.[7] Following Combes, Duranton, and Overman (2005), who diagrammatically present a simple model of a city in an urban system, the spatial distribution of labor force in a country and wage differences across regions can be regarded as an equilibrium outcome from a system of all regions. They consider the net wage, actual wage minus cost of living, as a function of city size. Upon this, an additional assumption of a flat labor supply curve, implying perfect labor mobility across regions, gives the equilibrium outcome that net wage is common in a country, while actual wages vary across regions.[8]

[7] Duranton and Puga (2004) discuss three types of mechanism: sharing, matching, and learning which bring about urban increasing returns and consequent higher wage in big cities.

[8] The equilibrium is depicted in Appendix Figure 1.

By combining equation (1) and (2), the wage in region *i* can be written as a function of distance between region *i* and economic center.

(3) $$w_i = v(\, d_i\, ;\, R_i, TL),\;\; v^{'} < 0,\;\; \forall\, i = 1,2,\cdots, I,$$

$$where\; v(\, d_i) = w[(e(\, d_i\, ;\, R_i, TL)].$$

Equation (3) means that high wage near the core reflects high labor productivity supported by externalities, particularly, benefits from urbanization economies. Hence, for some industries using high-skilled labor intensively, high wage near the core is not a centrifugal force, while it can be one of the dispersing forces for other industries in this spatial structure. Note that international trade intensifying the core, for example, shifts the wage curve upward in proportion to the increased population in each location.

Given the high concentration in the primacy, domestic transportation cost is a critical factor attracting firms to the core since firms prefer to locate close to large output markets, where transportation networks provide easy access to consumers. A firm also considers easier access to intermediate inputs and materials, and so its location decision is affected by the spatial dispersion of input sources in a country. Hence, for a specific region, the transportation costs reflecting countrywide accessibility of both output demand and materials supply can be defined in a simple way as an increasing function of the distance from the center:

(4) $$t_i^k = t^k\,(d_i\, ;\, TL^k),\quad t^{k'} > 0,\;\; t^{k''} > 0,\quad \forall\, k = 1,2,\cdots, K,\quad \forall\, i = 1,2,\cdots, I.$$

The superscript *k*, denoting the *k*-th industry, indicates that the transportation cost curve is not identical for every industry. Besides the difference in per-unit transportation cost across industries, each industry faces different spatial dispersion of demand according to the final use of its goods as well as different spatial density of intermediate goods and materials.[9] The positive second derivative implies increasing returns to scale in transportation, i.e. marginal transportation cost is lower for regions near the core. Like TL in equation (3), the industry-specific index of trade liberalization, TR^k changes the shape of transportation cost curve faced by all firms in the k-th industry. That is, in the case of trade liberalization intensifying the core, the domestic transportation curve moves upward in proportion to the distance from the core.

The core-periphery structure described above simplifies a country's internal economic geography by relating the distance from the core to regional gaps in wage and domestic transportation costs. In this framework, firms in the same location may confront different transportation costs according to their industry, while wage differences across regions are common to every firm. In the following subsection, heterogeneous firms' location decision is incorporated into the core-periphery structure.

2.2.2. Firm Heterogeneity and the Location Decision

Even in a narrowly defined industry, firms with different productivity levels coexist because they are uncertain about their productivity before making an irreversible investment to enter the industry (Melitz 2003). In this framework, I assume that a firm also makes its location decision based on geographical features such as easier market access and large pool of workers with its targeting labor productivity. That is, each firm chooses a location ensuring lower transportation costs and less expenditure on input

[9] If all firms in the k-th industry use materials at a fixed rate in order to produce a good, the per unit domestic transportation costs can be written as $t_i = t_i^{output} + s_m t_i^{material}$, where s_m indicates a constant amount of materials needed for per unit production. The transportation costs for outputs and materials (t_i^{output}, $t_i^{material}$) are weighted values reflecting distance of region i from consumers and from suppliers of intermediate goods or materials, respectively, who are unevenly spread out all over the country.

purchases, which are usually traded off each other. Thus, the location decision of a firm in the k-th industry is made based on the full knowledge of w_i and t_i^k functions. Since its productivity is still uncertain, it uses an expected productivity when finding a cost-minimizing location.[10] Note that incumbent firms may not be located at the current cost-minimizing location, because they chose different locations according to their entry times, when they faced different shapes of w_i and t_i^k.

After choosing a location and incurring the entry costs, a firm draws its productivity from a common distribution $g(\alpha)$ and chooses either to:

1) Exit if variable profit does not cover fixed production cost.

2) Serve only domestic markets if variable profit covers only fixed production cost.

3) Serve domestic and foreign markets, if variable profit covers fixed production and fixed foreign-market entry costs.

The full equilibrium will feature a break-even or cutoff level of productivity (α^*), such that firms with $\alpha < \alpha^*$ exit immediately upon learning of their draw. Given $g(\alpha)$, the equilibrium productivity distribution $h(\alpha)$ is a truncated distribution from below:[11]

(5)
$$h(\alpha) = \begin{cases} \dfrac{g(\alpha)}{1 - G(\alpha^*)} & if \;\; \alpha \geq \alpha^* \\ 0 & otherwise. \end{cases}$$

Given α^* which divides firms into 1) and 2), another cutoff productivity for export is also endogenously determined as in Melitz (2003), categorizing remaining firms into 2) and 3).

---

[10] If the expected productivity is identical to all firms in the industry, there will be an industry-specific cost-minimizing location.

[11] The equilibrium in a closed economy is represented by the cutoff productivity α^* and average profits $\bar{\pi}$ via the zero-profit and free-entry conditions. Note that the range of productivity levels of survival firms and thus, the cutoff and average productivity level (α^* and $\bar{\alpha}$, respectively) are endogenously determined (See Melitz 2003, p.1703).

With firms sorted by productivity as in equation (5), Melitz (2003) shows that a reduction in international trade costs allows more productive foreign competitors to enter the domestic market. The subsequent exit of least-productive domestic firms, whose profit become negative, increases average industry productivity, which is a key source of gains from trade. Instead of immediate exit from the industry, some of these firms may try to move their location toward the current cost-minimizing location. Some of surviving firms also have an incentive to move their location in order to improve their profitability in response to more liberalized trade. All these activities of firms in an industry together alter the internal economic geography related with the industry, and the associated agglomeration economies as will be discussed in the following section.

2.3. A Model of Heterogeneous Firms' Location

2.3.1. Cost Function

Consider a small monocentric economy composed of finite regions as described in the previous section. There is a fixed domestic supply of labor, which is mobile across regions. Workers will tend to move from the low-wage to high-wage region, so that the equilibrium that real wages are equalized among regions holds. The preferences of a representative consumer take the usual CES form over a continuum of differentiated goods in an industry:

(6) $$U = \left[\int_{a \in A} q(a)^{\rho}\, d\alpha\right]^{1/\rho},$$

where a represents a variety and A denotes the set of available goods which are assumed to be only realized through firms' different levels of technology. As in Melitz (2003), I treat a variety produced by a firm as having a one-to-one correspondence to its marginal-cost level. That is, each firm chooses to produce a distinguished variety a representing its marginal cost. These goods are close substitutes with the elasticity of substitution equal to $\sigma = 1/(1-\rho) > 1$.

Production requires labor and materials. The former is inelastically supplied at its aggregate level e_i, the equilibrium outcome of spatially distinguished labor markets. However, each location has a different nominal wage depending on the distance to the core. The price of material varies across regions and is reflected in firms' transportation costs as noted earlier (footnote 9). Production exhibits increasing returns in the usual way; firm technology is represented by a cost function that exhibits constant marginal cost with a fixed overhead cost. In addition to the production costs, firms incur different transportation costs according to their locations. Therefore, the total cost of the *j*th firm located in region *i* to produce variety a is:

$$(7) \quad C_{ij}(a) = w_i\, l_j + w_i f^k + t_i^k\, q_j$$

$$= \left(\frac{w_i}{\alpha_{ij}} + t_i^k\right) q_j + w_i f^k, \quad \alpha_{ij} > 0, \quad \forall\, i = 1,2,\cdots,I, \qquad \forall_{\mathrm{j}} = 1,2,\cdots,J,$$

where l_j and q_j are the *j*-th firm's labor use and output, respectively, f^k is a fixed overhead in industry *k*, represented by labor. Recall that w_i and t_i^k are wage and industry-specific transportation costs in region *i* as defined by equation (3) and (4), respectively. Finally, α_{ij} is the index of productivity corresponding to the *j*-th firm in region *i*.

Before making an irreversible investment to enter the industry, a firm chooses the (variable) profit maximizing location based on observed w_i, t_i^k, and the expected value of

α, i.e., $E(\alpha) = \bar{\alpha}$. For the expected productivity level, the minimized marginal cost can be defined as:

(8) $$\varphi(d_i; \bar{\alpha}) = \min_{d_i} \left[\frac{v(d_i)}{\bar{\alpha}} + t^k(d_i) \right] \quad subject\ to\ v(d_i) \text{ and } t^k(d_i).$$

For some industries, the cost-minimizing location d^* would be the core or the periphery, depending on the sum of second derivatives of $v(d_i)$ and $t^k(d_i)$. The sufficient condition for the existence of unique and interior solution to equation (8) is, therefore, $v'' + t^{k''} > 0 \;\; \forall\, d_i$. Since $t^{k''} > 0$ is already assumed in equation (4), it is enough to restrict the wage function as $v'' > -t^{k''}$ for an interior solution. Also, for each location, it is assumed that there exists a minimum level of productivity ensuring positive profit.

Note that the variable cost minimizing location derived from $\varphi(d_i)$ varies across industries in a country. For example, industries intensively using high-quality labor abundant near the core and producing consumer-use goods have steeper slopes of $t^k(d_i)$ than others using materials spread out in the peripheral areas and producing intermediate goods as outputs. With the common wage function among industries, the consumer-use industries would have their cost-minimizing locations closer to the economic center than others. Figure 2.1 illustrates the case of two industries, where the cost-minimizing locations are determined by the different shapes of transportation costs curve.[12] Panel (a) of Figure 2.1 shows a steeper $t^k(d_i)$ relative to that in panel (b), e.g. consumer goods (bulky, ready to consume) versus intermediates. As a result, d^* in panel (a) is closer to the core than that in panel (b).

---

[12] If fixed costs are considered together, then the location will further move to the right, more remote location from the economic center.

2.3.2. Productivity

After incurring the sunk-entry cost, a firm realizes its productivity level. From equation (7), higher α_{ij} implies that a firm can produce a variety with a lower labor requirement, i.e. at lower marginal cost, relative to its counterparts. Firm productivity has two parts: firm-specific and location-specific components as follows:

(9) $$\alpha_{ij} = \beta_j + \gamma_i^k, \quad \beta_j > 0, \quad \gamma_i^k \geq 0, \quad \forall\, j = 1,2,\cdots,J, \quad \forall\, i = 1,2,\cdots,I.$$

First, labor productivity of a worker can vary according to the firm where she/he is employed, because of distinct firm-specific technology and/or managerial skill. For instance, labor-saving machinery or advanced knowhow on allocating inputs to production process efficiently can improve labor productivity specific to a firm (β_j). The second component of productivity is not dependent of firm, but on the region where firm is located (γ_i^k). Co-location of firms in the same industry can generate positive externalities in the way of enhancing individual firms' productivity, i.e. localization economies (Henderson 1988, Lall et al. 2004).[13] These intra-industry benefits include access to specialized know-how, the presence of industry-specific buyer-supplier networks, and opportunities for efficient subcontracting. Workers with industry-specific skills will be attracted to such clusters giving firms access to a larger specialized labor pool. For each region, γ_i^k is an increasing function of number of firms or number of employees belonging

[13] On the contrary, so-called urbanization economies, related with inter-industry benefits are already reflected in the wage function, so it is not considered as a source of the location-specific productivity. In this framework, the denser a region is, the higher labor quality is, and hence a higher wage. Empirical studies also show that population density has a smaller effect on the productivity of manufacturing industries than other factors representing localization economies (Henderson 1988, Henderson et al. 2001, Lall et al. 2004).

to industry k. Since firms and workers gather around the cost-minimizing location, the location-specific productivity can be expressed as a function of the distance from the core (d_i) and its maximum is attained at the cost-minimizing location.

2.3.3. Equilibrium

By the assumption on demand, each firm faces a residual demand curve with constant elasticity of substitution σ, regardless of its productivity, and thus chooses the same profit maximizing markup equal to $\sigma/(1-\sigma) = 1/\rho$. Therefore, price equals the product of marginal cost and the constant markup and it is a function of the distance from the core and firm-specific productivity which are independent of each other:

$$(10) \qquad p(d_i, \beta_j) = \left(\frac{v(\,d_i)}{\beta_j + \gamma^k(\,d_i)} + \,t^k(\,d_i)\right)\frac{1}{\rho}, \qquad 0 < \rho < 1.$$

Firm profit is then:

$$(11) \qquad \pi(d_i, \beta_j) = \left(\frac{v(\,d_i)}{\beta_j + \gamma^k(\,d_i)} + \,t^k(\,d_i)\right)\frac{(1-\rho)q_j}{\rho} - \,v(\,d_i)f^k = \frac{r_{ij}}{\sigma} - \,v(\,d_i)f,$$

where r_{ij} is the firm's revenue. The optimal consumption and expenditure for an individual variety are respectively derived as;

$$(12) \qquad q(d_i, \beta_j) = Q\left[\frac{p(d,\beta)}{P}\right]^{-\sigma}, \quad r(d_i, \beta_j) = R\left[\frac{p(d,\beta)}{P}\right]^{1-\sigma},$$

where $Q, R,$ and P are aggregate output, revenue, and price, respectively.[14] From equations (10), (11) and (12) it is apparent that the closer to the cost-minimizing location a

[14] $Q \equiv U$ and $R = PQ$. A representative consumer's utility over all varieties and aggregate price according to Dixit and Stiglitz (1977) are given by:

firm is located, the bigger and the more profitable it is. Hence, in this framework, most manufacturing firms would want to move from inefficient areas into cost minimizing locations, resulting in stronger localization.

An equilibrium in an industry will be characterized by a mass of *J* firms (and hence *J* differentiated varieties) and a joint distribution of distances from the economic center and firm-specific productivity levels. Since d and β are independent, their joint density is the product of the marginal densities, that is, $f(d,\beta) = f(d)\, f(\beta)$. In such an equilibrium, the aggregate price can be written as;

$$P = J^{1/(1-\sigma)} p\left(\tilde{d}, \tilde{\beta}\right), \quad (13)$$

where $\tilde{\beta} = \left[\int_0^\infty \beta^{\sigma-1} f(\beta) d\,\beta\right]^{1/(\sigma-1)}$ and $\tilde{d} = \left[\int_1^I \varphi(\,d_i)^{1-\sigma} f(d) d\,d\right]^{1/(\sigma-1)}$ are weighted means of the firm-specific productivity level β and distance from the core *d*, respectively. Both are independent of the number of firms *J*.[15]

2.3.4. Cost-minimizing Location

Similar to Melitz (2003), the cutoff marginal cost, $c^* = w/\alpha + t$ or break-even productivity level, α^* is endogenously determined by the intersection of the zero-cutoff-profit and free-entry conditions. The above cutoff determines which firms survive to sell at least in the domestic market (equation (5)).

$$U = \left[\iint_{0\,0}^{\infty\,\mathrm{I}} q(d,\beta)\, dd\, d\beta\right]^{\frac{1}{\rho}} \text{ and } P = \left[\iint_{0\,0}^{\infty\,\mathrm{I}} p(d,\beta)^{1-\sigma} dd\, d\beta\right]^{\frac{1}{1-\sigma}}$$

[15] $\tilde{\beta}$ and $\tilde{d}$ are also regarded as aggregate productivity and distance, respectively, since they completely summarized the information of $f(d)$ and $f(\beta)$ relevant for all aggregate variables:
$R = Jr\left(\tilde{d}, \tilde{\beta}\right), Q = J^{1/\rho} q\left(\tilde{d}, \tilde{\beta}\right),$ and $\Pi = J\pi\left(\tilde{d}, \tilde{\beta}\right)$

Contrary to the case of productivity, before entering into an industry firms know which location is most profitable. So**,** in order to survive, new entrants choose their location close to the cost-minimizing location. Upon entry with a low-productivity draw, a firm may decide to immediately exit and not produce, if it cannot generate profits to cover variable cost of production. Thus, for each location, there is a possible range of productivity level under which firms can survive. For instance, a firm with a low-productivity draw can survive when located close to a cost-minimizing location. Panel (a) of Figure 2.2 describes where firms can survive given the cutoff marginal cost, c^*. Centering around the cost-minimizing location d^*, firms with $\varphi(d)$ ranging between d^{min} and d^{max} can survive in this market. On the other hand, Figure 2.2's panel (b) illustrates how the survival range of productivity levels varies for each location. Recall that $\gamma(d)$ is the location-specific productivity due to localization economies. Since $\gamma(d)$ is increasing in the number of firms, its maximum is attained at d^* (where highest concentration of firms is possible) and has positive values in the interval between d^{min} and d^{max}. With the exogenously determined location-specific productivity, the break-even level of productivity α^* and that of firm-specific productivity β^* are endogenously determined. The values β^{max} and α^{max} are the upper bound of firm- and overall (firm plus location) productivity level, respectively, which a firm in the industry can achieve. Therefore, the darkly shaded area is the allowable (survival) range of firm-specific productivity for each location. The brightly shaded area represents additionally permitted range of productivity due to location-specific productivity. It is the same as the area below the $\gamma(d)$ curve. As a result, the sum of two areas represents the possible range of productivity of firms for survival in each region. The lowest productive firm with α^* should be only located in d^*

to stay in the industry. Even a firm with higher productivity level than α^* could not survive unless it is located within the lightly shaded regions in panel (b) of Figure 2.2.

2.4. Impact of Trade Liberalization

Given the equilibrium location for each industry as in Figure 2.2, I now turn to the role of trade in this core-periphery type economy. Trade liberalization is conventionally modeled by employing additional marginal cost for export τ and/or fixed costs for entering foreign markets, f_{ex}. As shown in Melitz (2003), openness of a previously closed economy shifts up the downward slopping zero-cutoff-profit curve, so that both cutoff productivity level and average profit per firm in equilibrium increase. It induces the partitioning of firms by export status and productivity level as previously noted. Moreover, it reallocates market shares and profits to high productivity firms. Here, I provide additional insights on the changes in the spatial distribution of economic activity following trade liberalization. That is, exposure to trade alters a country's internal economic geography in two ways. It changes the cost-minimizing location changes and intensifies localization for each industry. In the following, these two cases are separately discussed in detail.

2.4.1. Impact on Cost-minimizing Location

Under the trade liberalization, a firm's cost function is rewritten as;

(14)
$$C_{ij}(a) = w_i^o\, l_j + w_i^o f^k + t_i^{o\,k}\, q_j + \tau^k\, q_j + w_i^o f_{ex}^k$$

$$= \left(\frac{w_i^o}{\alpha_{ij}} + t_i^{o\,k} + \tau^k\right) q_j + w_i^o f^k + w_i^o f_{ex}^k, \quad \tau^k > 0, \quad f_{ex}^k > 0,$$

where τ^k is the industry-specific marginal cost for export and f_{ex}^k is fixed entry costs for foreign markets expressed by labor. Compared with equation (7), note that the wage and domestic transportation cost curves changes from w_i and t_i^k to w_i^o and ${t_i^o}^k$, respectively. As defined earlier, wage is an increasing function of population and trading with foreign countries also incurs domestic transportation costs from a firm's location to the trading center, i.e. the change of trade regime is regarded as a shifter in equation (3) and (4).

Consider now two contrasting cases. The first is that openness proceeds to intensify the primacy in the current core. That is, trade facilities are mainly constructed in the existing economic center. Second, trade is mostly conducted with adjacent countries through periphery. These two cases are illustrated in Figure 2.3, whose panel (a) shows that international trade intensifying the core shifts the wage curve upward in proportion to the increased population in each location ($w \rightarrow w^o$), and also moves the domestic transportation curve upward in proportion to the distance from the core ($t \rightarrow t^o$). Compared with autarky, the minimized marginal cost curve for a firm with average productivity level moves up in this trade regime ($\varphi \rightarrow \varphi^o$). The resulting cost-minimizing location moves closer to the core than before ($d^* \rightarrow d^o$). Moreover, the cutoff marginal cost decreases due to additional trade costs for export and consequent exit of least productive domestic firms, in the line with firm heterogeneity trade model ($c^* \rightarrow c^o$). The regions where firms can survive in this trade regime are identified by the distance between two intersection points of φ^o and c^o, which is much shorter than that of autarky. The above results provide an important insight that the exposure to trade mainly through the existing core makes internal economic geography directly faced by domestic firms worse off and thus reduces their profits.

On the other hand, panel (b) depicts the case where a country mainly trades with neighboring countries via new trading centers built in peripheral areas, implying deconcentration. This trade regime shifts the wage and domestic transportation curves downward in proportion to the decreased population and distance from the core, respectively. Therefore, the minimized marginal cost curve moves down in response to trade. However, the consequent change in the cost-minimizing location depends on the relative slopes of the two curves. For a location near the core (the periphery), an advantage from relative decrease of regional wage (domestic transportation cost) is almost traded off with a disadvantage from relative increase of domestic transportation cost (regional wage). Even with the lowered cutoff marginal cost, the range of survival locations increase relative to that of autarky. Consequently, the new trading center, away from the existing economic center, can make domestic firms more profitable by altering the internal economic geography. Considering the industry dimension, there can be multiple locations, each a cost-minimizing location for a given industry. Thus, the core can likely become a cluster for the domestic market, while other clusters can specialize on exports (Venables 2005).

In sum, when trade occurs through the core, trade liberalization deepens concentration of workers and jobs in the core. This result arises because firms away from the core incur more expensive domestic transportation cost due to the growth of both output demand and imported intermediate goods or materials supply in the core (equation 4). In the second case, trade works in favor of deconcentration. Firms in peripheral areas near national borders can take advantage of cheaper domestic transportation costs in trade with neighboring countries.

2.4.2. Impact on Localization

For an industry, the pattern of openness might be different from the overall trade regime in a country. It is also possible that the wage curve does not much respond to trade (equation 3), while the industry-specific domestic transportation costs (equation 4) do. Thus, it is expected that industries, each having its own cost-minimizing location, can be spatially reallocated in a country via heterogeneous trade liberalization. Furthermore, trade plays an important role in intensifying localization. To see this, consider a case in which an industry trades through the periphery, while overall trade or other industries' trade may occur in the core. The consequences for localization for an industry trading through the periphery can be illustrated again by the upward shift of wage and downward shift of domestic transportation cost curves.

As shown in Figure 2.4, trade liberalization in this industry affects the wage and domestic transportation cost curves so that the cost-minimizing location is altered from d^* to d^{o} where firms trading with adjacent countries benefit from internal economic geography. The range of survival region shrinks and moves toward the periphery. Along with the change of optimal location for the industry, the increased cutoff productivity level from α^* to α' due to the additional trade costs τ and f_{ex}, reduces the range of firm-specific productivity levels for each location by the difference between α^* and α', which is equal to that between β^* and β^{o}. The new cutoff level of firm-specific productivity β^{o} and newly indentified survival range of locations between d_{min}^{o} and d_{max}^{o} determine the firm-specific productivity range over locations, which are represented by the darkly shaded area in Figure 2.4. Firms with productivity below α' are likely to exit regardless of their locations, while firms whose productivity is over α', but whose locations are not included in the new survival regions should choose to exit or move their location closer to d^o. That is, among firms which can survive in autarky but not in liberalized trade regime, some firms with productivity higher than α' will move their location toward d^o if their expected profits are high enough to offset the additional costs required in moving their locations. Even among the firms in the darkly shaded area, some firms which want to export and increase their market shares and profits would like to move into locations closer to the cost minimizing location d^o.

Consequently, most of firms in the industry gather near d^o and intensify localization economies, i.e. γ_i^k is an increasing function of number of firms or number of employees belonging to industry *k*. If the number of firms and that of workers in an industry does not decrease in response to trade liberalization, the location-specific productivity, γ^o for each survival region will be higher than that in autarky since the same number of firms is located in a narrower space than before. As a result, the resulting cutoff level of productivity is not α', but actually α^o which is lower than α'. This result could delay the anticipated exit of least productive firms and the consequent resource reallocation towards more efficient firms (Melitz 2003). The intensified localization may deter the proper functioning of the selection process. That is, least productive firms can continue to stay in the industry by fully taking advantage of improved localization economies.

2.5. Summary and Conclusions

In this study, I investigate the impact of trade liberalization on firms' optimal location and localization economies by linking a firm-heterogeneity trade model with a simplified core-periphery framework from new economic geography. I employ domestic transportation costs as a main source of centripetal force and adopt regional differences in labor quality and wages as the basis for a core-periphery structure. In a closed economy, internal economic geography determines the cost-minimizing location of firms in each industry. However, trade liberalization alters the optimal location of an industry depending on the region through which trade occurs. Trade through the economic center or core accelerates concentration of economic activities near the core by making firms unprofitable away from the core. In contrast, trade opportunities through the periphery can create a multi-centric export-oriented economy. Since the pattern and degree of trade liberalization may differ across industries, it is possible that industries can be spatially relocated within a country.

Moreover, trade intensifies localization for each industry since most of firms in an industry move to or gather around their industry-specific cost-minimizing location. The consequent clustering of firms generates additional localization economies, a result consistent with related empirical studies. More importantly, the intensification of localization economies can slow or delay the selection process, i.e. exit of low productivity firms, following trade liberalization. Similarly, resources may not be reallocated to high productivity firms in response to trade liberalization.

The above results provide implications of openness both for spatial distribution of economic activities and for industrial adjustment patterns in a country. Exposure to global market could lead to multi-centric economy and furthermore be a catalyst for spatially more even development, if it is accompanied by regional development policies improving infrastructure and encouraging export participation.

References

Ades, A., and E.L. Glaeser. 1995. "Trade and Circuses: Explaining Urban Giants." *Quarterly Journal of Economics* 110 (1): 195-227.

Alonso-Villar, O. 2001. "Large Metropolises in the Third World: An Explanation." *Urban Studies* 38 (8): 1359-1371.

Baldwin, R.E., and T. Okubo. 2006. "Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting." *Journal of Economic Geography* 6 (3): 323-346.

Baldwin, R.E., R. Forslid, P. Martin, G.I.P. Ottaviano, and F. Robert–Nicoud. 2003. *Economic Geography and Public Policy*. Princeton University Press, Princeton.

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

Ciccone, A., and R.E. Hall. 1996. "Productivity and the Density of Economic Activity." *American Economic Review* 86 (1): 54-70.

Combes, P-P., G. Duranton, and H.G. Overman. 2005. "Agglomeration and the Adjustment of the Spatial Economy." *Papers in Regional Science* 84 (3):311-349.

Combes, P-P., T. Mayer, and J-F. Thisse. 2008. *Economic Geography: The Integration of Regions and Nations*. Princeton University Press, Princeton.

Dixit, A.K., and J.E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67 (3): 297-308.

Duranton, G., and D. Puga. 2004. "Micro-foundations of Urban Agglomeration Economies." In: Henderson, V., Thisse, J.F. (Eds.) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland, Amsterdam, pp. 2064-2117.

Fujita, M., P.R. Krugman, and A.J. Venables. 1999. *The Spatial Economy. Cities, Regions and International Trade*. MIT Press, Cambridge MA.

Giles, J.A., and C.L. Williams. 2000. "Export-led Growth: A Survey of the Empirical Literature and Some Non-causality Results." Part I. *Journal of International Trade & Economic Development* 9 (3): 261-337.

Hanson, G.H. 1998. "Regional Adjustment to Trade Liberalization." *Regional science and urban economics* 28 (4): 419-444.

Head, K., and T. Mayer. 2004. "The Empirics of Agglomeration and Trade," in J.V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics: Cities and Geography*, vol. 4, North Holland, Amsterdam: pp. 2609-2665.

Helpman, E. 2006. "Trade, FDI, and the Organization of Firms." *Journal of Economic Literature* 44 (3): 589-630.

Helpman, E., M. Melitz, and S. Yeaple. 2004. "Export Versus FDI with Heterogeneous Firms." *American Economic Review* 94 (1): 300-316.

Henderson, J.V. 1986. "Efficiency of Resource Usage and City Size." *Journal of Urban Economics* 19 (1): 47-70.

Henderson, J.V. 1988. *Urban Development: Theory, Fact and Illusion*. New York: Oxford University Press.

Henderson, J.V., T. Lee, and Y.J. Lee. 2001. "Scale Externalities in Korea." *Journal of Urban Economics* 49 (3): 479-504.

Krugman, P.R. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99 (3): 483-499.

Krugman, P.R., and R. Livas Elizondo. 1996. "Trade Policy and the Third World Metropolis." *Journal of Development Economics* 49 (1): 137-150.

Lall, S.V., Z. Shalizi, and U. Deichmann. 2004. "Agglomeration Economies and Productivity in Indian Industry." *Journal of Development Economics* 73(2): 643-674.

Melitz, M. 2003. "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71(6): 1695-1725.

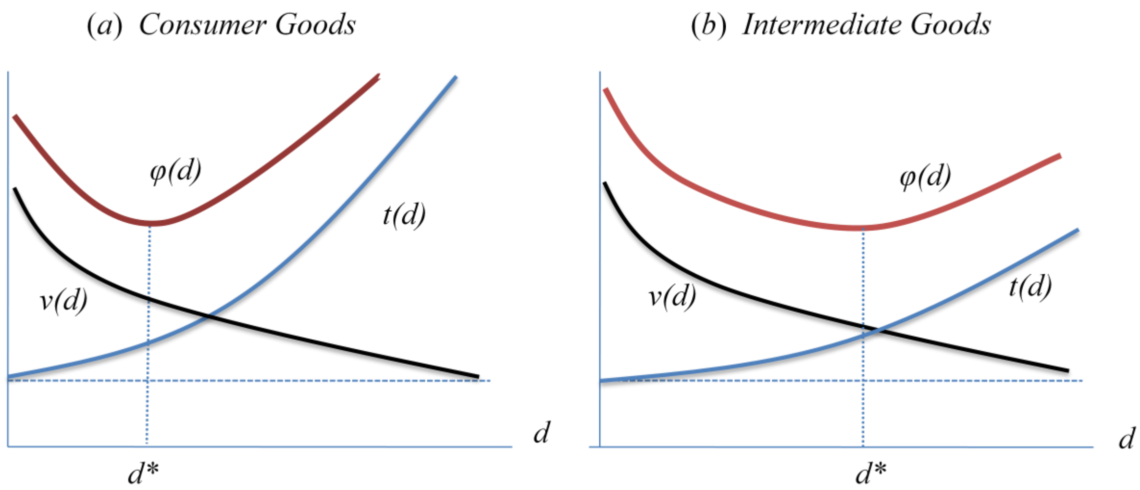
Moomaw, R.L. 1981. "Productivity and City Size: A Critique of the Evidence." *Quarterly Journal of Economics* 96 (4): 675-688.

Segal, D. 1976. "Are There Returns to Scale in City Size?" *Review of Economics and Statistics* 58 (3): 339-50.

Sveikauskas, L.A. 1975. "The Productivity of Cities." *Quarterly Journal of Economics* 89 (3): 393-413.

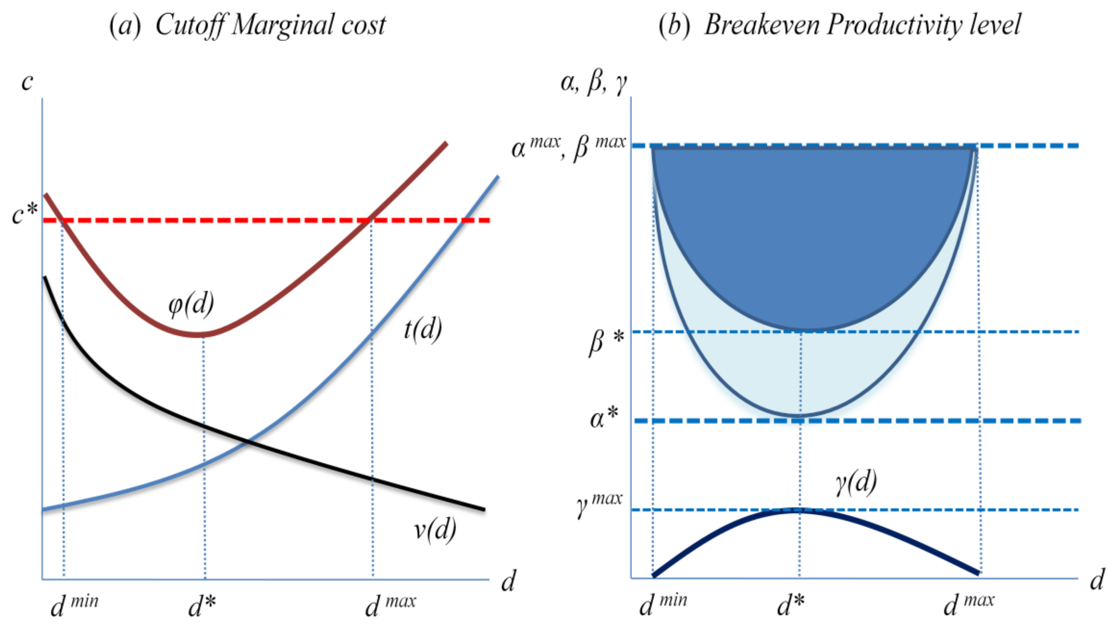
Venables, A.J. 1996. "Equilibrium Locations of Vertically Linked Industries." *International Economic Review* 37 (2): 341-360.

Venables, A.J. 2005. "Spatial Disparities in Developing Countries: Cities, Regions, and International Trade." *Journal of Economic Geography* 5 (1): 3-21.



Panel (a) **Panel (b)**

Figure 2.1. Cost-minimizing Locations in Two Industries



Panel (a) **Panel (b)**

Figure 2.2. Cutoff Marginal Cost and Survival Range of Productivity

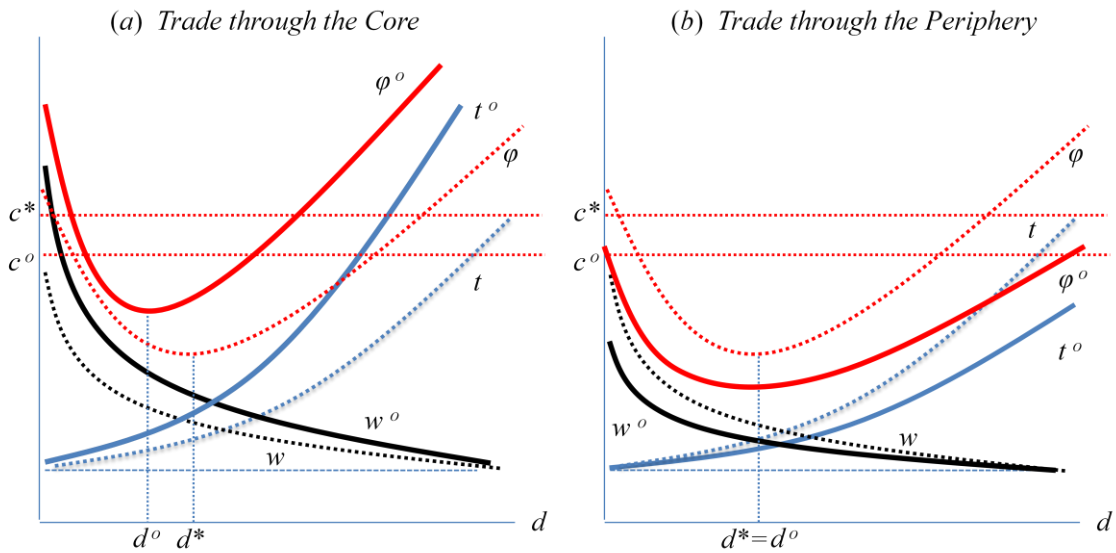


Figure 2.3. Impact of Trade on Firm's Location

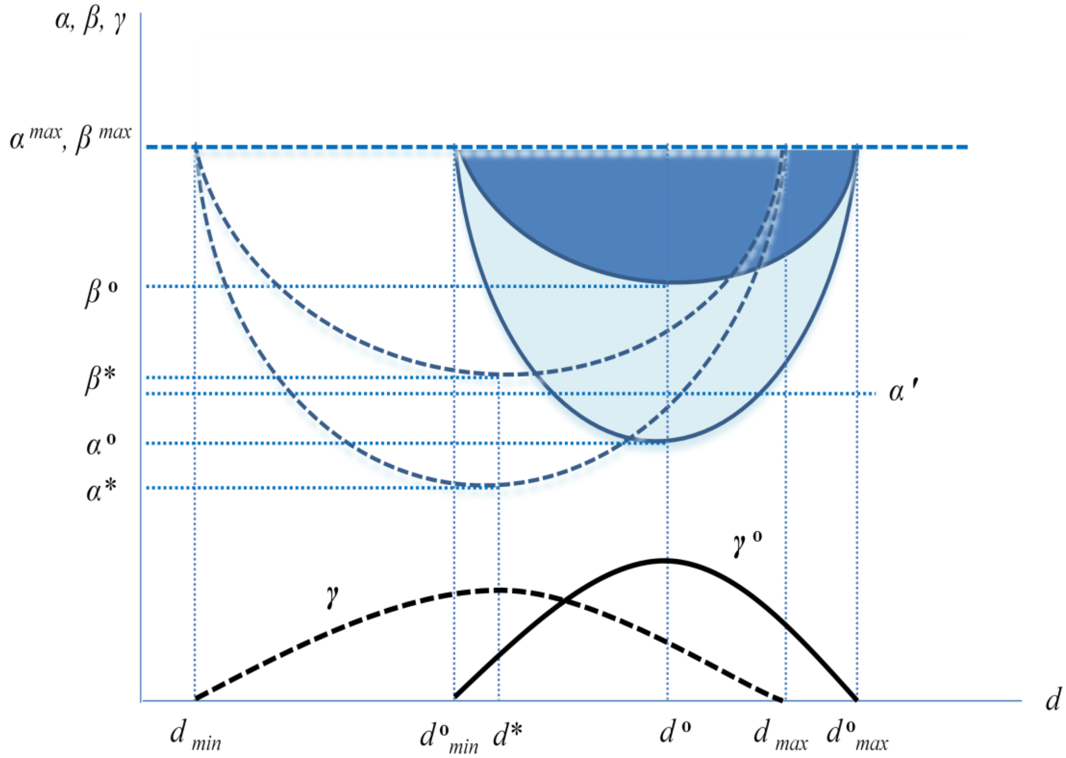


Figure 2.4. Impact of Trade on Localization

Appendix: Equilibrium Labor Force and Wage Difference across Regions

The wage and cost of living curves differ across cities because of differences in endowments, local institutions, and public goods as well as distance from the core (Combes, Duranton and Overman 2005). Therefore, differences in the net wage curve naturally lead to cities of different sizes in equilibrium. Moreover, some regional factors like amenities, for example, may shift the labor supply curve. That is, more attractive cities face a labor supply curve that is below that of less attractive ones.

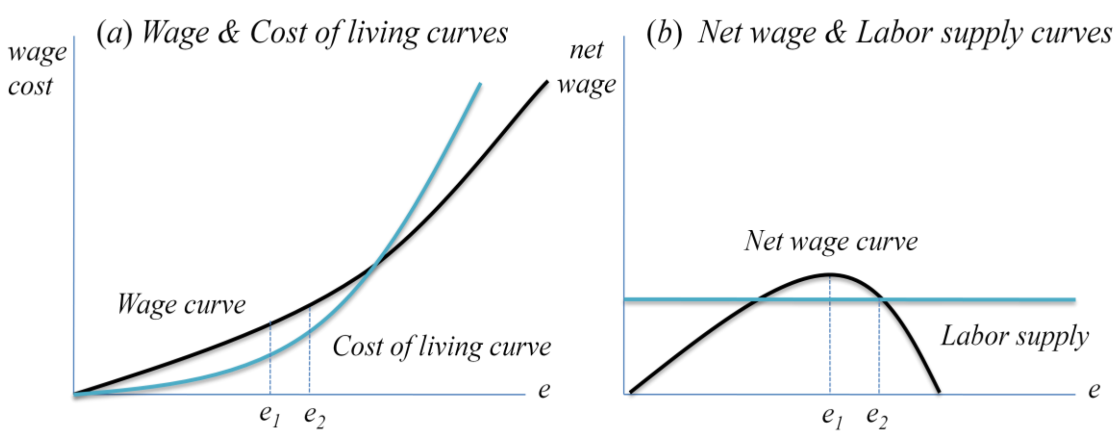


Figure 2.A.1. Upward-sloping wage curve and Equilibrium city size

CHAPTER THREE

Foreign Competition, Domestic Infrastructure and Manufacturing Productivity in Korea

3.1. Introduction

The sources and welfare consequences of spatial concentration of economic activity have been at the core of the new economic geography literature (Krugman 1991, Fujita et al. 1999, Baldwin et al. 2003, Combes et al. 2008). For instance, Henderson et al. (2001) find that highly concentrated urban regions' output, employment, and wages are persistently larger than those of their rural counterparts (Glaeser and Gottlieb 2009). In an attempt to identify the sources of regional inequalities, Rice et al. (2006) argue that the urban-rural income gap is largely due to the productivity difference between the two regions (Acemoglu and Dell 2010). Hence, in policymakers' search for solutions to regional equity in economic development, the question why we observe large inter-regional productivity differences takes on significance.

Few studies have focused on the sources of productivity variation within a country (Ciccone 2002, Rosenthal and Strange 2004, Rice et al. 2006). A primary concern of these studies has been the effect of agglomeration forces on regional productivity. Agglomeration is associated with productivity-enhancing externalities such as easier access to intermediate inputs and complementary services, availability of a large labor pool with multiple specialization, inter- and intra-industry information transfers, and co-use of general infrastructure. For example, Rice et al. (2006) find that a region's labor productivity increases by 3.5% when its working-age population is doubled. As a result, our understanding of the contribution of agglomeration forces to regional productivity has increased. However, agglomeration's importance relative to the underlying technical change, fostered by competition and innovation, has received limited research attention (Glaeser and Gottlieb 2009, Oosterhaven and Broersma 2007).

The role of international competition in enhancing productivity is highlighted in the emerging heterogeneous-firms trade theory and its applications (Melitz 2003, Bernard et al. 2003). Such models suggest that international competition forces least productive firms to exit and reallocate resources to highly productive firms, increasing average industry productivity (Syverson 2004, Bernard et al. 2007). Similarly, the effect of infrastructure on regional economic development has been well documented (Rosenthal and Strange 2004). For instance, by lowering transport costs, infrastructure improvement is often regarded as an increase in domestic competition between manufacturing firms, producing similar goods and being located at any region within a nation (Behrens 2006). Additional factors in the determination of productivity variation among regions include natural amenities, which can negatively or positively impact productivity (Deller et al. 2001, Rappaport 2009).

The objective of this study is to investigate the relative importance of agglomeration and pure technical change (or raw productivity) in regional productivity growth. In addition, we identify and quantify the determinants of variation in regional raw productivity growth. Advances in spatial econometrics allow us first to attribute productivity growth to externalities within and across spatial units (e.g. counties) and to the underlying technical change or raw productivity. Then, we examine foreign competition and domestic infrastructure as key determinants of regional raw productivity variation. Our application focuses on Korean manufacturing using county-level data during 1999-2006.[16] The sample period includes the aftermath of the 1997-98 Asian financial crisis, in which most Korean manufacturing industries experienced contraction and restructuring. More competitive and high-productive firms emerged following the crisis. Moreover, Korean manufacturing has been facing increased global competition during the sample period with China's accession to WTO (2000) and continued advances in transportation and communication technology.

The remainder of the chapter is organized as follows. In the following section we outline a spatial model to compute total factor productivity and identify the sources of its regional variation. Section 3.3 describes the data for estimating the production function and the sources of regional productivity variation. Our empirical analysis is detailed in Section 3.4. The final section summarizes and provides concluding remarks.

3.2. Agglomeration, Spatial Dependence and Regional Productivity Variation

In this section, we first describe our approach to estimating a county-level production function controlling for agglomeration and spatial spillover effects. Then, we show how to derive measures of regional (raw) productivity and examine the sources of such variation.

3.2.1. Production Function with Agglomeration and Spatial Dependence

The objective in this subsection is to quantify the productivity of each county and attribute it to agglomeration forces and the underlying pure technical change. We begin by outlining the motivation for agglomeration effects on regional manufacturing production. If average

[16] By county, we refer to Si, Gun or Gu in Korea. Each Si, Gun or Gu is represented by a government and considered equivalent to a county or district in a developed economy, e.g. the United States. For ease of reference, we maintain the county terminology throughout the manuscript.

production costs decline as the regional scale of production increases, then it is beneficial to concentrate production in particular locations, i.e. agglomeration economies (Marshall 1890, Henderson 2003).[17] In particular, positive externalities from other firms in the same industry, resulting from sharing industry-specific labor pool, know-how and business networks are referred to as localization economies. Another source of agglomeration economies points to inter-industry benefits from co-location of firms, usually termed urbanization economies. These productivity-enhancing externalities accrue from easier access to complementary services, availability of a large pool with multiple specialization, inter-industry information transfers, and co-use of general infrastructure (Rosenthal and Strange 2004, Lall et al. 2004). Localization and urbanization economies can be considered as centripetal or attractive forces leading to concentration of economic activity in specific locations.[18]

Following Henderson (2003) and others, we specify a Cobb-Douglas production function to distinguish the sources of agglomeration economies from raw TFP:

$$y_{it} = f(A_{it})B_{it}K_{it}^{\alpha}M_{it}^{\beta}L_{it}^{\gamma}, \quad (1)$$

where y_{it} is the output of the *i*-th spatial unit at time *t*, $f(A_{it})$ represents external influences on production from agglomeration economies; and B_{it} captures raw productivity, i.e. pure technology effect, since the unit-specific agglomeration effects are already controlled for by $f(A_{it})$. Hence, B_{it} is an index of raw TFP capturing Hicks-neutral technical change.

---

[17] Duranton and Puga (2004) examine three mechanisms - sharing, matching and learning- which bring about increasing returns.

[18] Centrifugal or repelling forces, e.g. high factor prices driven by competition among firms and congestion, act in the opposite direction (Baldwin et al. 2003, Lall et al. 2004).

The variables K_{it}, M_{it}, and L_{it} are *i*-th spatial unit's capital stock, intermediate goods and labor at time *t*, respectively (Moomaw 1983, Nakamura 1985, Henderson 1988).

The agglomeration economies are specified in an exponential form:

$$f(A_{it}) = e^{aAE_{it}}, \tag{2}$$

where AE_{it} is a variable or vector, and *a* the associated coefficient(s), generally used in the literature to capture agglomeration economies (Rosenthal and Strange 2004). Common measures include location quotients based on employment or output, and market size represented by population density or similar indicators. More details on variables representing agglomeration economies in our application are presented in section 3.

With agglomeration effects, a double-log production function is specified for the empirical analysis. Substituting (2) into (1), taking log, and then subtracting lnL_{it} from both sides yields:

$$ln\left(\frac{y_{it}}{L_{it}}\right) = aAE_{it} + lnB_{it} + \alpha\, ln\left(\frac{K_{it}}{L_{it}}\right) + \beta\, ln\left(\frac{M_{it}}{L_{it}}\right) + \mu lnL_{it}, \tag{3}$$

where $\mu = \alpha + \beta + \gamma - 1$ represents the returns to scale.

Since overall TFP (agglomeration effects plus raw productivity) varies across counties and over years, the estimation of (3) should include fixed effects. So, we consider a county- and time-fixed effects specification:

$$ln\left(\frac{y_{it}}{L_{it}}\right) = aAE_{it} + d_{1c} + d_{2t} + \alpha\, ln\left(\frac{K_{it}}{L_{it}}\right) + \beta\, ln\left(\frac{M_{it}}{L_{it}}\right) + \mu lnL_{it} + u_{it}, \tag{4}$$

where d_{1c} and d_{2t} are vectors of county- and time-specific intercepts, respectively, and u_{it} denotes a random disturbance term.

Equation (4) considers the production relationship for a county in isolation, i.e. the effects of production in the vicinity or neighborhood is ignored. A number of authors have argued for the role of spatial effects in production, based on recent advances in theory and empirics. On the theory side, the new economic geography models endogenize the agglomeration process, where the presence or absence of firms (output or employment) in the neighborhood critically impacts production at a particular location (Krugman 1991, Fujita et al. 1999, Baldwin et al. 2003, Combes et al. 2008). That is, intra- and inter-industry benefits noted earlier also arise from adjacent, contiguous or neighboring spatial units. On the empirical side, the recent development of spatial econometric techniques has provided flexibility in modeling agglomeration effects within a given location and across locations (Anselin 1988, Anselin et al. 2004, Vaya et al. 2004).[19] Empirical approaches include spatial lag or error models, where the former posits that the production in one spatial unit impacts that in neighboring units, but the effect decays with distance. That is, output of neighboring counties affect a given county's output, an effect well documented in growth and innovation literature (Coe and Helpman 1995, Barro and Sala-i-Martin 2003). The spatial error models hypothesize positive or negative correlation in the residuals (u_{it}) of contiguous spatial units and such correlation is also anticipated to decline with distance. Both approaches allow for testing and specification of the hypothesized spatial spillover effects.

To account for spatial effects in the production function, equation (4), we first define the spatial weights matrix *W*, which is square, symmetric and non-stochastic. Following Anselin et al. (2004), we model the spatial interaction between counties through a distance-based *W* matrix, whose diagonal elements are zero and the off-diagonal elements equal to w_{ij}, which is the relative weight of county j for county i, given by the inverse of the distance

---

[19] Agglomeration within a county is given by AE_{it}, while spatial effects arise from neighboring counties.

between the two counties. Note that spatial dependence makes usual OLS estimator biased or inefficient: bias accrues from spatial interdependence among cross-sectional units, while inefficiency results from the omission of spatially correlated regressors, which lead to spatially correlated errors. As noted earlier, the two most often-used regression models to incorporate spatial patterns are spatial lag and spatial error models (Anselin and Bera 1998, Anselin et al. 2004). However, the spatially lagged dependent variable is typically correlated with the disturbance term. Furthermore, these two spatial patterns might coexist in most of practical applications. To address these problems, Kelejian and Prucha (1998, 1999) suggest a feasible generalized spatial two-stage least squares (FGS2SLS) procedure for models that have both spatially lagged dependent variables and spatial autocorrelation in errors. Equivalent to a generalized method of moments (GMM) estimator, this nonparametric covariance estimation technique yields standard errors of the parameters that are robust to spatial dependence among the error terms. Applying Kelejian and Prucha (1998, 1999) to equation (4), while defining, in a panel setting, two $NT \times NT$ spatial weight matrices $W^* = I_T \otimes W$ and $M^* = I_T \otimes M$, we specify a spatial panel regression model as:

$$(5) \quad ln\left(\frac{y_{it}}{L_{it}}\right) = d_{1c} + d_{2t} + aAE_{it} + \boldsymbol{\lambda} \sum_{j=1}^{N} \sum_{s=1}^{T} w_{it,js}^{*} \, ln\left(\frac{y_{js}}{L_{js}}\right) + \alpha \, ln\left(\frac{K_{it}}{L_{it}}\right) + \beta \, ln\left(\frac{M_{it}}{L_{it}}\right) + \mu ln L_{it} + u_{it},$$

where $u_{it} = \boldsymbol{\rho} \sum_{j=1}^{N} \sum_{s=1}^{T} m_{it,js}^{*} u_{js} + \varepsilon_{it}$, $w_{it,js}^{*}$ and $m_{it,js}^{*}$ are *it*-th row and *js*–th column elements of W^* and M^*, respectively and ε_{it} is an element of $NT \times 1$ vector of innovations, $\varepsilon \sim iid(0, \sigma^2 I_{NT})$. Note that the two spatial weights matrices, W and M, can be identical.[20]

[20] The FGS2SLS procedure is composed of three steps. In the first step, the model is estimated by 2SLS, while using a linearly independent subset of spatially lagged exogenous variables as instruments for the spatial lag of the

Based on the equation (5), the logarithm of overall TFP of *i*-th county at year *t* is:

$$(6) \quad \widehat{lnTFP_{it}} = ln\left(\frac{y_{it}}{L_{it}}\right)^{o} - \hat{\alpha}\, ln\left(\frac{K_{it}}{L_{it}}\right)^{o} - \hat{\beta}\, ln\left(\frac{M_{it}}{L_{it}}\right)^{o} - \hat{\mu} lnL_{it}{}^{o} = \widehat{lnf(A_{it})} + \widehat{lnB_{it}},$$

where the superscript (o) indicates the transformation. The estimate of agglomeration effects in the *i*-th spatial unit at year *t* is sum of the estimated values of spatial dependence and agglomeration economies:

$$(7) \quad \widehat{lnf(A_{it})} = \hat{\lambda} \sum_{j=1}^{N} \sum_{s=1}^{T} w_{it,js}^{*}\, ln\left(\frac{y_{js}}{L_{js}}\right)^{o} + \hat{a}AE_{it}{}^{o}$$

Then, the logarithm of raw TFP estimate for the *i*-th county at year *t* is obtained by:

$$(8) \quad \widehat{lnB_{it}} = \widehat{lnTFP_{it}} - \widehat{lnf(A_{it})}$$

$$= ln\left(\frac{y_{it}}{L_{it}}\right)^{o} - \hat{\lambda} \sum_{j=1}^{N} \sum_{s=1}^{T} w_{it,js}^{*}\, ln\left(\frac{y_{js}}{L_{js}}\right)^{o} - \hat{a}AE_{it}{}^{o} - \hat{\alpha}\, ln\left(\frac{K_{it}}{L_{it}}\right)^{o} - \hat{\beta}\, ln\left(\frac{M_{it}}{L_{it}}\right)^{o} - \hat{\mu} lnL_{it}{}^{o}$$

<u>3.2.2. Regional Productivity Variation</u>

Productivity variation across countries has received much attention in the growth and development literature (Stoneman 1995, Harrigan 1999). However, few studies have focused on the variation of productivity within a country (Ciccone 2002, Rosenthal and Strange 2004, Rice et al. 2006, Glaeser and Gottlieb 2009, Acemoglu and Dell 2010). Most intra-country studies have evaluated the impact of agglomeration forces on productivity. The previous section presented an approach to distinguish between productivity that arises from

---

dependent variable. The second step uses the 2SLS residuals to consistently estimate the spatial-error autoregressive parameter (ρ) and the variance of innovations (σ^2) using Kelejian and Prucha's (1998) GMM procedure. In the final step, the estimated ρ is used to perform a Cochrane-Orcutt transformation $(1 - \hat{\rho}M^{*})$ and the transformed model is again estimated by 2SLS.

agglomeration and the rest, a pure or raw technical change. Then the interest is to analyze factors contributing to shifts or changes in the raw-productivity of regions. In particular, we seek to determine the role of international competition and domestic infrastructure in shaping the raw productivity of regions (Henderson et al. 2001, Behrens 2006). The idea that reductions in international trade costs bring about increased productivity is commonly found in the trade theory especially in the new heterogeneous-firms trade model (Melitz 2003, Bernard et al. 2003). Empirical evidence of a trade-cost or competition effect on productivity is beginning to accumulate (Kim 2000, Greenaway, Morgan and Wright 2002, Syverson 2004, Amiti and Konings 2007, Bernard et al. 2007, Novy 2008). Similarly, the effect of infrastructure on regional economic development has been well documented (Rosenthal and Strange 2004). For instance, the improvement of infrastructure is often regarded as an increase in domestic competition between manufacturing firms, producing similar goods and being located at any region within a nation (Behrens 2006). Additional factors in the determination of raw productivity variation among regions include natural amenities, which can negatively or positively impact productivity (Deller et al. 2001). Ceteris paribus, a location in the region with higher level of amenities is more attractive to firms and labor (human capital) compared to that in low-amenity region (Rappaport 2009).

With the estimated raw TFP levels of counties from equation (8), we employ a nonparametric kernel density estimator to derive the regional productivity distribution. For this purpose, we group the 231 Korean counties into four regions (see the data section for details). Such a grouping is consistent with the classification of the nine Korean provinces and also provides enough degrees of freedom to estimate kernel densities of raw productivity by region in the empirical analysis. As a nonparametric approach, kernel density estimators have no fixed structure and depend upon all the data points to derive an estimate. Specifically, a kernel function is centered at each estimating point. A spatial unit's contribution to density estimation at an estimating point depends on how far the spatial unit's productivity is apart from the given point. As a result, kernel functions yield a smooth estimation of the distribution curve (Jones 1997, Beaudry et al. 2005). Kernel estimation requires choices on kernel type and width. If we smooth too much, we throw away detail that might be informative, while if we smooth too little, we might be distracted by detail that is not informative. In the estimation, we use a Gaussian kernel at each estimating point and minimize the mean integrated squared error. Then, using the Gaussian cumulative density, we estimate alternative percentile values (10^{th} percentile, median, and 90^{th} percentile), and moments of the regional productivity distribution.

With alternative percentile values, we can partition the raw-productivity panel into three groups. Then, we specify the estimated group-wise TFP as a function of trade cost (TC_t), infrastructure (IN_{et}), and amenities (AM_{et}):

$$(9) \qquad \widehat{lnTFP}_{et}^{p} = g^{p}(TC_t, IN_{et}, AM_{et}), \qquad p = 10\%, 50\%, 90\%,$$

where the superscript p refers to three alternative percentiles, implying three different estimation functions, and the subscripts e and t refer to regions (4) and time, respectively. Data on the explanatory variables in equation (9) are described in the following section. A specification of equation (9), augmented by regional fixed effects, is:

$$(10) \qquad \widehat{lnTFP}_{et}^{p} = d_e^{p} + \delta_1^{p}\, TC_t + \delta_2^{p}\, IN_{et} + \delta_3^{p}\, AM_{et} + \varepsilon_{et}, \qquad p = 10\%, 50\%, 90\%.$$

The estimated coefficient δ_1^{p}, for example, represents the effect of trade costs change on the shift of productivity distribution for each alternative percentile group. We consider alternative specifications of equation (9) including non-linearity, e.g. interaction variables.

3.3. Data and Trade Cost Derivation

3.3.1. Production Function Data

The Si, Gun, or Gu is the basic political and administrative unit in Korea and the smallest spatial unit at which most economic data are available.[21] The base data for this study comprises of 231 spatial units (77 Si, 81 Gun, and 73 metropolitan Gu), which can be referred to as counties and the 1999-2006 period. Our sample period includes the aftermath of the 1997-98 Asian financial crisis, in which most Korean manufacturing industries experienced contraction and restructuring. More competitive and high-productive firms emerged following the crisis. Moreover, Korean manufacturing has been facing increased global competition during the sample period with China's accession to WTO (2000) and continued advances in transportation and communication technology.

Each county's manufacturing output, labor, capital and value added during 1999-2006 are obtained from *Census on Basic Characteristics of Establishment* (Statistics Korea). The *Census* for manufacturing industry is compiled on the basis of a survey which is annually administered on all manufacturing establishments with at least five employees. Output is denoted by the production value for all manufacturing establishments included in the survey. The average number of employees per month and the tangible fixed asset at the end of year are used for labor and capital, respectively. Raw or intermediate materials value -the sum of raw material costs, energy costs, repairing expenses, and outsourcing expenses- is calculated by subtracting value added from production value. Since output, capital, and materials are all measured in nominal values (million won), they are deflated by

---

[21] Gun becomes a Si when its population increases to 100,000. Gu is administratively equivalent to Si or Gun except that it is located near the 7 metropolitan cities.

manufacturing producer price index (PPI), GDP deflator, and import price index for materials/intermediates, respectively.

Consistent with prior literature, our empirical analysis employs alternative measures of agglomeration economies (Rosenthal and Strange 2004). For localization economies, we consider (i) a location quotient based on manufacturing employment, (ii) lagged industry employment and (iii) lagged industry establishments (number). The employment-based LQ is defined as $LQ_i^M = {}^{L_i^M/L_i}\Big/_{L^M/L}$, where the superscript *M* refers to manufacturing sector and the subscript *i* refers to a spatial unit. L_i and L_i^M denote *i*-th county's overall employment and that in manufacturing sector, respectively, while L^M and L denote the corresponding measures at the national level (Statistics Korea). For urbanization economies, we consider (i) population density (PD_{it}), i.e. population per square kilometer, of the spatial unit, (ii) employment density, and (iii) the (inverse) of the Herfindahl-Hirschman Index. Results of specification tests and the final specification of agglomeration economies in equation (5) are detailed in the next section.

To capture the spatial dependence among the 231 counties in county-level TFP estimation, it is necessary to define the structure of *W*, i.e. spatial weights between pairs of observations. For a single cross-section of *N* observations, a $N \times N$ spatial weighting matrix *W* is usually defined in terms of a first– or multiple-order contiguity criteria or a distance decay function such as $d_{ij}^{-\delta}$ and $exp^{-\delta d_{ij}}$, where δ is a parameter and d_{ij} is a distance between counties *i* and *j*. In this study, we employ d_{ij}^{-1} to specify spatial weights for all pairs of observations. That is, our 231×231 spatial-weighting matrix *W* has zero diagonal elements and the non-zero off-diagonal is equal to the inverse of the great-circle distance between two units. Given our panel setting (*N*=231 and *T*=8), we additionally define a $NT \times NT$ matrix $W^* = I_T \otimes W$ to obtain the spatially lagged dependent variables $W^* \, ln(y/L)$ where $ln(y/L)$ is a column vector whose element is $ln(y_{it}/L_{it})$. Following Anselin (1988), W^* is row-standardized, i.e. each row is divided by the sum of the row elements.

3.3.2. Data for Estimating Regional Productivity Variation

For estimating regional productivity distribution with county-level TFP estimates using kernel-density techniques, we group counties into 4 regions. The first region (R1) is the Seoul metropolitan area, representing national primacy and main economic center (Gyeonggi-do). The second region (R2) surrounds the first region (Gangwon-do, Chungcheongbuk-do, and Chungcheongnam-do). The third region (R3) is relatively rural and remote area (Jullabuk-do, Jeollanam-do, and Jeju-do). The fourth region (R4) is located in southeastern coastal area and is relatively specialized in traditional manufacturing industries (Gyeongsangbuk-do and Gyeongsangnam-do). The second biggest city of Korea, Busan, also plays a role as a regional economic center in the fourth region. There are 65, 51, 43, 72 counties in R1, R2, R3, and R4, respectively. Appendix Figure 3.A.1 shows the four regions and population density by county (2006).

For estimating equation (10), the computation of international trade costs follows Novy (2008), who uses international trade flows to express multilateral resistance terms as a function of observable trade and output data (Anderson and van Wincoop 2003). Thus, bilateral trade costs between home (*H*) and foreign (*F*) countries affect trade flow in both directions (x_{HF}, x_{FH}) and intra-country trade (x_{HH}, x_{FF}) can be used as sizes variables, which control for multilateral resistance. Since gross bilateral trade costs between *i* and *j* are symmetric ($t_{HF} = t_{FH}$), tariff-equivalent bilateral trade costs τ_{HF} can be written as:

$$(11) \qquad \tau_{HF} = \left(\frac{x_{HH}x_{FF}}{x_{HF}x_{FH}}\right)^{\frac{1}{2(\sigma-1)}} - 1.$$

Trade costs calculated in equation (11) capture not only traditional trade costs (transportation costs and tariffs), but also non-tariff barriers (institutional and cultural barriers). For computation of trade costs as in equation (11), we consider GDP excluding service sector as intra-country trade (x_{HH}). GDP and trade data of Korea and its major trading partners (China, US, Japan, and Taiwan) are obtained from the Bank of Korea and World Bank.[22]

Appendix Table 3.A.1 presents bilateral trade cost estimates with four major trading partners during 1999-2006. The tariff-equivalent trade costs with China and U.S., for example, considerably decreased from 0.47 to 0.18, and slightly decreased from 0.42 to 0.37, respectively. The average change in trade cost with four major partners, weighted by countries' trade volumes with Korea, declined 40% over the eight years with the annual growth rate is -7.05%. For estimation purpose, we define the freeness of trade (FT) as the inverse of the weighted average trade costs so that FT increased during the same period, as shown in the last column of Appendix Table 3.A.1.

The density of paved roads (unit: km/km^2) is used as the indicator of infrastructure level for each county and aggregate it to each extended region. The indicators representing the level of amenities for each extended region are climate variables such as clear/sunny days,

---

[22] Trade with these countries accounts for over 50 percent of total Korean trade, and the share has been fairly constant over the sample period.

precipitation (mm), mean air temperature (C^0), and natural environment variables such park-, forest-, and river- area densities (McGranahan 1999). Appendix Table 3.A.2 presents descriptive statistics on infrastructure and amenity variables in our application.

3.4. Estimation and Results

3.4.1. Production Function Estimation and Results

Prior to estimating equation (5) using spatial econometric techniques, a number of specification and error structure tests are conducted. The first of these tests is the comparison of a fixed effects model, equation (5), against a random-effects specification. A Hausman specification test rejects the hypothesis that the county-specific random effects are uncorrelated with other regressors, implying that the random-effects specification will yield biased parameter estimates. The second set of tests deals with the specification of agglomeration economies. As noted earlier, we had three alternative measures each for localization and urbanization economies. However, most of these measures are highly correlated and estimation results exhibited typical symptoms of high multicollinearity. Analysis of variance inflation factors and likelihood ratio tests lead us to choosing only population density (PD) to represent the agglomeration effects. In the above analysis, it is also found that a time trend better fits the data relative to year-specific dummies, although the former imposes a constant coefficient across all time periods.

Having chosen spillover and agglomeration variables, $W^* \, ln(y/L)$ and PD, we turn our attention to the possible endogeneity of these two regressors. Kelejian and Prucha (1998) suggest a linearly independent subset of exogenous variables and their spatial lags as instruments for the spatially-lagged dependent variable.[23] The agglomeration economies can be endogenous in regional production because of their simultaneous determination with output in many applications (Rosenthal and Strange 2004). With regard to the spatial lag, the null hypothesis of exogeneity for $W^{*}\,ln(y/L)$ is rejected when spatially-lagged materials per capita, $W^{*}\,ln(y/M)$ is chosen as an excluded instrument in 2SLS estimation. To check weak identification, we conduct the Cragg-Donald Wald F test.[24] The statistic exceeds the critical value for 10 percent maximal IV size distortion (Stock and Yogo 2005). Relevant test results are presented at the bottom of Table 3.1. With lagged population density as the instrument, the null hypothesis that *PD* can be considered exogenous in the estimation of equation (5) is not rejected.

The final estimation issue is concerned with the second step of FGS2SLS. Despite the spatial interdependence effects captured using a spatial lag in the first step, the possibility of spatial autocorrelation in errors remains.[25] That is, spatial dependency in our data arises from not only observed neighborhood effects (spatial lag) but also spatially unobserved and stochastic processes (Anselin 1998). To check for spatial autocorrelation of errors, we conduct a Lagrange multiplier (LM) test using the spatial weighting matrix M^{*} whose off-diagonal element is d_{ij}^{-2}, instead of W^{*} whose off-diagonal element is d_{ij}^{-1}. In the second step, these matrices imply different spillover processes. In the former case, a shock at one location diffuses in relatively narrower space, since the matrix gives neighboring locations

---

[23] Fingleton and Le Gallo (2008) allow more than one endogenous variables in the estimation of spatial models.

[24] Weak identification arises when the excluded instrument is weakly correlated with the endogenous regressor.

[25] Moran's I is significantly different from zero. The difference between the index ($I = 0.045$) and its expected value under the null ($E(I) = -0.001$) is larger than twice its standard error (0.004), i.e. two-sigma rule criterion.

much more weight compared to the latter case. The LM test statistic based on IV residuals from the first step is high enough to reject the null hypothesis of no spatial error autocorrelation, when employing M^* as the spatial weighting matrix in the second step.[26] This means that the autocorrelation in error term is local, while the simultaneous spatial interaction is global (nation-wide) in manufacturing production.

Table 3.1 shows the 2SLS and FGS2SLS estimates of equation (5), obtained with a heteroskedasticity-consistent (Eicker-Huber-White) covariance estimator. The 2SLS results, including the spatial lag but not the spatial error, are shown in the third column of Table 3.1. The FGS2SLS estimates in the second column of Table 3.1 indicate that the Korean manufacturing exhibits constant returns to scale, and that the elasticities of output with respect to capital, labor, and materials are 0.06, 0.20, and 0.74, respectively. These elasticities are in the range of outcomes of related literature (Oh et al. 2000).[27] The estimate of the spatial autocorrelation coefficient ρ from the second step of FGS2SLS is positive and significant (0.307).

With the estimated coefficients in the FGS2SLS, we calculate the logarithm of total TFP (equation 6), raw TFP (equation 8), and agglomeration effects for each county. Table 3.2 presents the mean and standard deviation of each extended region's raw productivity and agglomeration effect for 1999 and 2006. Note that while R1 enjoys the highest agglomeration effect, R2, which surrounds the metropolitan area (R1), exhibits the highest levels both of raw TFP and overall TFP. The latter is consistent with firms moving out of R1 and into R2 in

---

[26] The LM test statistic (66.73) exceeds the 95 percent critical value for $\chi^2_{(1)}$ (6.31).

[27] Other recent studies of Korean manufacturing include Kim (2000), Hahn (2004), and Lee et al. (2005). Some of these studies employ plant or firm-level data, but for earlier sample periods (e.g. 1990-98).

recent years. The rate of growth in TFP is highest in R3 possibly due to low initial levels (convergence). Table 3.2 suggests that raw-productivity (i.e. technical change) accounts, on average, for over 60 percent of the overall productivity. The rest is attributed to agglomeration effects of TFP. Appendix Figure 3.A.2 and 3.A.3 show the substantial spatial variation in raw TFP relative to agglomeration effects of TFP in Korean manufacturing. [28]

3.4.2. Kernel Densities and Sources of Regional Variation in Korean Productivity

Given the estimated raw TFP levels of counties, we next estimate kernel densities for each region and year and derive the first moment and the 10th and 90th percentile of regional productivity distribution. For example, Figure 3.1 shows the estimated kernel density curves of R3 for each sample year. All regions have experienced positive productivity growth during 1999 to 2006.[29] However, Table 3.2 also shows the substantial variation in the productivity growth rates of the extended regions (0.74 to 1.89%). Note again that that the relatively remote and less-developed region, R3, has considerably improved its raw productivity during the period.

To identify the sources of regional variation in productivity, we specify the estimated first moment and alternative percentile values as a function of indicators of infrastructure and natural amenities as well as a measure of trade liberalization, i.e. freeness of trade. Using the likelihood-ratio test, the base specification in equation (10) is compared to an alternative specification with interaction terms as follows:

---

[28] For the entire country, the average annual growth rate of raw productivity (0.96%) is similar to that of the agglomeration effect (0.97%).

[29] Appendix Figure 3.A.4 provides the kernel densities of all four regions for three years (1999, 2003, and 2006). It is shown that each region's density has shifted to the right as time goes.

$$(14) \qquad \widehat{lnTFP}_{et}^{p} = d_e^p + d_{99-01}^p + d_{05-06}^p + \delta_1^p\, FT_t + \delta_2^p\, IN_{et} + \delta_3^p\, AM_{et}$$
$$+ \delta_4^p\, IN_{et} * FT_t + \delta_5^p\, IN_{et} * AM_{et} + \delta_6^p\, FT_t * AM_{et} + \varepsilon_{et},$$
$$p = 10\%, 50\%, 90\%, \qquad e = 1,2,3,4, \; t = 1, \cdots, 8.$$

In equation (14), we add two year dummies to control for business cycle effects.[30] The interaction term between infrastructure and freeness of trade might capture a potential synergy or trade-off relationship between domestic infrastructure and trade liberalization in enhancing raw productivity level (Behrens 2006).

The two specifications, equation (10) and (14), are first estimated with a pooled sample including dummies for alternative percentiles. Then, they are estimated using three different samples representing the three percentiles of the kernel densities (10, 50 and 90%). For these four samples, the OLS estimates using White's heteroskedasticity consistent covariance estimator are reported in Table 3.3; the upper panel shows results from equation (10), while the lower panel shows those of equation (14). The column labeled "pooled sample" shows the parameter estimates obtained with the pooled sample. Other columns are labeled based on the corresponding sample. Since little variation in results is observed across specifications using alternative amenity variables, only those results with "clear days" are reported in Table 3.3.

For the case without interaction terms, results from all four samples are generally consistent. That is, most of the coefficients on infrastructure and freeness of trade are statistically significant at least at the 5% level and have positive signs consistent with

---

[30] The inclusion of these temporal dummies raises the significance of other variables of interest as well as the overall fit of estimation results. They appear to capture two recessions (2002, 2005) of the Korean economy.

economic theory. Improved infrastructure and liberalized trade should expand market access and provide cheaper intermediates to firms, but they also create additional competition from other domestic and foreign firms. The competition effect is highlighted in the literature on trade liberalization, infrastructure and TFP growth, and the emerging trade theory with heterogeneous firms (Kim 2000, Greenaway, Morgan and Wright 2002, Behrens 2006, Amiti and Konings 2007, Bernard et al. 2007). For instance, Melitz (2003) shows that cutoff productivity level for survival in an industry increases when trade costs decline, forcing low-productivity firms to exit and reallocating resources to high-productivity firms. Thus, trade liberalization leads to higher average industry productivity. Since low-productivity firms are hypothesized to exit, declining trade costs should affect the lower tail of the productivity distribution, which is observed in our results in Table 3.3. Our results show that the competition effect of trade-cost changes likely dominates the increased market access to low-productivity firms, whose exit would then increase a region's 10^{th} percentile TFP. When the density of the lower tail of a distribution is shifted to the right, the median must shift in the same direction as well. The latter is consistent with the results shown in Table 3.3, column labeled "Median." The effect of infrastructure on raw TFP also varies across the three percentiles investigated here. The 10^{th} percentile TFP effect of infrastructure is larger than that on median or 90^{th} percentile, where the latter is significant at the 10% level only. Together, the results on freeness of trade and infrastructure suggest significant effects at the left-tail of the spatial productivity distribution. That is, regions with low raw TFP are likely to face significant structural changes from domestic development and trade liberalization. Except for the 10^{th} percentile sample, coefficients on the amenity variable are not significant, implying that they do not much affect the regional distribution of manufacturing productivity in Korea. Since amenity variables are measured over extended regions, lack of variation within a region and over time might be a reason for observed results.

For the second specification, the coefficients on linear terms retain the sign and significance from the earlier estimation, but the interaction between freeness of trade and infrastructure is significant for all samples. The sign of these interaction terms is negative, implying that the marginal effect of freeness of trade on productivity is increased when infrastructure is less developed and also that the marginal effect of infrastructure on productivity increases when trade is less liberalized. These significant trade-off relationship between two variables results in the increased coefficient on the linear form of each variable, compared to the specification without interaction terms. However, in the 90th percentile sample, the coefficient on freeness of trade becomes insignificant. Meanwhile, the coefficients on other interaction terms are not significant except for 10th percentile sample. According to the LR test (last row, Table 3.3) comparing the two specifications for each sample, we cannot reject the null hypothesis that the coefficients of all interaction terms are zero for the 90th percentile sample only.

Using the estimates in Table 3.3, we quantify the effect of trade liberalization and infrastructure improvement on regional productivity and compare the magnitude of the effect among different percentile samples and regions. Since the dependent variable is the logarithm of TFP, the coefficients would be partial- or semi-elasticities in equation (10). Therefore, with interaction variables, the elasticities of manufacturing TFP with respect to infrastructure and freeness of trade are computed as:

(15) $$e_{et}^{IN} = (\widehat{\delta_2} + \widehat{\delta_4}\, FT_t + \widehat{\delta_5}\, AM_{et})\, IN_{et}, \qquad e_{et}^{FT} = (\widehat{\delta_1} + \widehat{\delta_4}\, IN_{et} + \widehat{\delta_6}\, AM_{et})\, FT_t.$$

These elasticities evaluated at the average and their counterparts for each region are given in Table 3.4. For the pooled sample, the average elasticity of productivity with respect to infrastructure is 0.455. That is, a one percent increase in the infrastructure index, on average, enhances regional manufacturing productivity by 0.455 percent, confirming the crucial role of infrastructure in regional productivity growth. The raw productivity-improvement effect of infrastructure is stronger for the low-productivity counties (0.511 for the 10th percentile sample). The results show that R1's raw-TFP appears to be more elastic to infrastructure improvement relative to other regions.

The lower panel of Table 3.4 shows that more liberalized trade represented by trade costs reduction increases regional productivity, with elasticities ranging from 0.078 to 0.115 for different samples. In particular, given the insignificant coefficient on freeness of trade with 90th percentile sample, the lower tail of the productivity distribution is more elastic to trade -costs change than its median. The cross-region difference in the freeness of trade elasticities is entirely attributed to its indirect effect on productivity through infrastructure. In the case of R1, the trade elasticities for all samples are vanishingly small, because the direct effect of international trade on productivity is likely dampened by the high level of local infrastructure. On the other hand, R2 and R3 exhibit higher trade elasticities, especially in the 10th percentile sample, suggesting that the effect of trade costs reduction on regional productivity becomes larger for the low-productivity counties located in regions with less developed infrastructure. That is, a least-efficient firm, located in a region with relatively poor infrastructure (sheltered from domestic competition), is likely more vulnerable to increased foreign competition, and responds more quickly by exiting. The exit of lowest productive firm at a county results in regional productivity enhancement in that county. In the same token, a firm located in highly urbanized area (R1) may not respond much to increased foreign competition, because it already faces a high level of domestic competition. Moreover, productivity gains from agglomeration economies may make firm exit from R1 less likely, and hence, a relatively low trade-cost elasticity is observed.

In the context of policy, Table 3.4 shows that further investments in infrastructure, especially in R1, brings about the largest growth in TFP. However, if the percentage changes, during the sample period, of infrastructure (21%) and freeness of trade (67%) are considered, gains in productivity associated with infrastructure improvements does not so much outweigh those arising from more liberalized international trade. That is, over the eight sample years, TFP growth of 9.45 % and 6.3% are respectively attributed to improved infrastructure and reduction of trade costs including a liberalized trade regime. Hence, a policy choice should factor in costs of each of these alternative measures to improve regional productivity.

In Appendix Figure 3.A.5, we match our raw TFP by counties with some evidence on firm exit and net job creation (Statistics Korea). The thick and red bordered counties have had firm exit rate over 10% or employment losses greater than 10%. Not surprisingly, most of the thick-red bordered counties have low TFP coincidental with our results on the spatial reallocation of resources following infrastructural and trade cost changes.

3.5. Summary and Conclusions

This study investigates the effect of foreign competition and domestic infrastructure on regional productivity variation in Korean manufacturing. For this purpose, we assembled data on manufacturing output and primary factors for 231 spatial units during 1999-2006. A production function, controlling for potential spatial dependence among counties as well as agglomeration effects, is used to estimate total factor productivity (TFP). We employ spatial econometric procedures to estimate TFP, which is attributed to agglomeration economies and technical change or raw TFP. Our estimates of the production function show that the Korean manufacturing exhibits constant returns to scale. Raw TFP accounted, on average, for over 60 percent of the overall TFP, while the rest is attributed to agglomeration effects of TFP. We found substantial spatial variation in raw TFP relative to agglomeration effects of TFP in Korean manufacturing.

With estimated raw TFP, the spatial units are grouped into four regions. For each region, we approximated the manufacturing (raw) productivity distribution in each year through a kernel density. Cumulative Gaussian density then allowed the estimation of the first and second moments of the distribution as well as alternative percentile values (10^{th} and 90^{th} percentile), which represented the shifts of the regional productivity distribution over time.

To identify the sources of regional variation in productivity, we specified the estimated first moment and alternative percentile values as a function of indicators of international trade costs including trade liberalization, infrastructure and natural amenities. Improved infrastructure and liberalized trade should expand market access and provide cheaper intermediates to firms, but they also create additional competition from other domestic and foreign firms. It is found that the net effect of infrastructure improvement and trade costs reduction resulted in a shift to the right, particularly the 10^{th} percentile value of, the regional productivity distribution. However, natural amenities appear to have little effects on regional productivity variation. Our results are consistent with the emerging literature on heterogeneous-firms theory that trade liberalization increases the average productivity of an industry. The latter result arises from the exit of least productive firms in an industry and the reallocation of resources to highly productive firms with the net result being higher average industry productivity.

Our study provides insights into the evolution of productivity across regions in Korea with emphasis on the roles of trade costs reduction and improved infrastructure. Substantial variation is observed in regional productivity, which is highly correlated with the level of infrastructure. Gains in productivity associated with infrastructure improvements appear to outweigh those arising from lowering international trade costs including a liberalized trade regime. However, such a comparison does not factor in costs of each of these alternative measures to improve regional productivity. Tax or debt policies, generating resources for infrastructure, have distributional effects similar to that of trade policies. Hence, costs to implement each policy should be considered. Our results help governments' evaluation of the benefits or gains from each policy option, which when compared with costs, aid in the selection of an effective policy to improve regional productivity.

References

Acemoglu, D., and M. Dell. 2010. "Productivity Differences between and Within Countries." *American Economic Journal: Macroeconomics* 2 (1): 169-188.

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

Anderson, J., and E. van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93 (1): 170-192.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht.

Anselin, L., and A. Bera. 1998. "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics." In: Giles, D.E.A., Ullah, A. (Eds.) *Handbook of Applied Economic Statistics*. Marcel Dekker, New York, pp. 237-289.

Anselin, L., R.J. Florax, and S.J Rey. 2004. "Econometrics for Spatial Models: Recent Advances." In: Anselin, L., Florax, R., Rey, S.J. (Eds.) *Advances in Spatial Econometrics. Methodology, Tools and Applications*. Springer-Verlag, Berlin, pp. 1-25.

Baldwin, R.E., R. Forslid, P. Martin, G.I.P. Ottaviano, and F. Robert–Nicoud. 2003. *Economic Geography and Public Policy*. Princeton University Press, Princeton.

Barro, R.J., and X. Sala-i-Martin. 2003. *Economic Growth*. 2nd edition. MIT Press, Cambridge, MA.

Beaudry, P., F. Collard, and D.A. Green. 2005. "Changes in the World Distribution of Output per Worker, 1960-1998: How a Standard Decomposition Tells us an Unorthodox Story." *Review of Economics and Statistics* 87 (4): 741-753.

Behrens, K. 2006. "Do Changes in Transport Costs and Tariffs Shape the Space-Economy in the Same Way?" *Papers in Regional Science* 85 (3): 379-399.

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

Bernard, A. B., J.B. Jensen, and P.K. Schott. 2003. "Plants and Productivity in International Trade." *American Economic Review* 93 (4): 1268-1290.

Ciccone, A. 2002. "Agglomeration Effects in Europe." *European Economic Review* 46 (2): 213-227.

Coe, D., and E. Helpman. 1995. "International R&D Spillovers." *European Economic Review* 39 (5): 859-887.

Combes, P-P., T. Mayer, and J-F. Thisse. 2008. *Economic Geography: The Integration of Regions and Nations*. Princeton University Press, Princeton.

Deller, S.C., T.H. Tsai, D.W. Marcouiller, and D.B.K. English. 2001. "The Role of Amenities and Quality of Life in Rural Economic Growth." *American Journal of Agricultural Economics* 83 (2): 352-365.

Duranton, G., and D. Puga. 2004. "Micro-foundations of Urban Agglomeration Economies." In: Henderson, V., Thisse, J.F. (Eds.) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland, Amsterdam, pp. 2064-2117.

Fingleton, B., and J. Le Gallo. 2008. "Estimating Spatial Models with Endogenous Variables, a Spatial Lag and Spatially Dependent Disturbances: Finite Sample Properties." *Papers in Regional Science* 87 (3): 319-339.

Fujita, M., P.R. Krugman, and A.J. Venables. 1999. *The Spatial Economy: Cities, Region and International Trade*. MIT Press, Cambridge, MA.

Glaeser, E.L., and J.D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4): 983-1028.

Greenaway, D., W. Morgan, and P. Wright. 2002. "Trade Liberalization and Growth in Developing Countries." *Journal of Development Economics* 67 (1): 229-244.

Hahn, C.H. 2004. "Exporting and Performance of Plants: Evidence from Korean Manufacturing." NBER Working Paper 10208, Cambridge, MA.

Harrigan, J. 1999. "Estimation of Cross-Country Differences in Industry Production Functions." *Journal of International Economics* 47 (2): 267-293.

Henderson, J.V. 1988. *Urban Development: Theory, Fact and Illusion*. Oxford University Press, New York.

Henderson, J.V. 2003. "Marshall's Scale Economies." *Journal of Urban Economics* 53 (1): 1-28.

Henderson, J.V., T. Lee, and Y.J. Lee. 2001. "Scale Externalities in Korea." *Journal of Urban Economics* 49 (3): 479-504.

Henderson, J.V., Z. Shalizi, and A.J. Venables. 2001. "Geography and Development." *Journal of Economic Geography* 1 (1): 81-105.

Jones, C.I. 1997. “On the Evolution of the World Income Distribution.” *Journal of Economic Perspectives* 11 (3): 19-36.

Kelejian, H.H., and I.R. Prucha. 1998. “A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances.” *Journal of Real Estate Finance and Economics* 17 (1): 99-121.

Kelejian, H.H., and I.R. Prucha. 1999. “A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model.” *International Economic Review* 40 (2): 509-533.

Kim, E., 2000. “Trade Liberalization and Productivity Growth in Korea Manufacturing Industries: Price Protection, Market Power, and Scale Efficiency.” *Journal of Development Economics* 62 (1): 55-83.

Krugman, P.R. 1991. “Increasing Returns and Economic Geography.” *Journal of Political Economics* 99 (3): 483-499.

Lall, S.V., Z. Shalizi, and U. Deichmann. 2004. “Agglomeration Economies and Productivity in Indian Industry.” *Journal of Development Economics* 73 (2): 643-674.

Lee, B.S., S. Kim, and S.H. Hong. 2005. “Sectoral Manufacturing Productivity Growth in Korean Regions.” *Urban Studies* 42 (7): 1201-1219.

Marshall, A., 1890. *Principles of Economics*, MacMillan, London (8th edition 1920).

McGranahan, D.A. 1999. *Natural Amenities Drive Rural Population Change*. Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture. Agricultural Economic Report No. 781.

Melitz, M., 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica* 71 (6): 1695-1725.

Moomaw, R.L. 1983. “Spatial Productivity Variations in Manufacturing: A Critical Survey of Cross-Sectional Analysis.” *International Regional Science Review* 8 (1): 1-22.

Nakamura, R. 1985. “Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities.” *Journal of Urban Economics* 17 (1): 108-124.

Novy, D. 2008. Gravity Redux: Measuring International Trade Costs with Panel Data. Working Paper, University of Warwick, UK.

Oh, I., J. Lee, and A. Heshmati. 2008. “Total Factor productivity in Korean Manufacturing Industries.” *Global Economic Review* 37 (1): 23-50.

Oosterhaven, J, and L. Broersma. 2007. "Sector Structure and Cluster Economies: A Decomposition of Regional Labour Productivity." *Regional Studies* 41 (5): 639-659.

Rappaport, J. 2009. "The Increasing Importance of Quality of Life." *Journal of Economic Geography* 9 (6): 779-804.

Rice, P., A.J. Venables, and E. Patacchini. 2006. "Spatial Determinants of Productivity: Analysis for the Regions of Great Britain." *Regional Science and Urban Economics* 36 (6): 727-752.

Rosenthal, S.S., and W.C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In: Henderson, V., Thisse, J.F. (Eds.) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland, Amsterdam, pp. 2118-2171.

Statistics Korea. http://kosis.kr.

Stock, J.H., and M. Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In: Stock, J.H., Andrews, D.W.K. (Eds.) *Identification and Inference for Econometric Models: a Festschrift in honor of Thomas Rothenberg*. Cambridge University Press, Cambridge, UK, pp. 80-108.

Stoneman, P. 1995 (ed.). *Handbook of the Economics of Innovation and Technological Change*. Blackwell, Oxford.

Syverson, C. 2004. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics* 86 (2): 534-550.

The Bank of Korea, Economic Statistics System. http://ecos.bok.or.kr.

Vayá, E., E. López-Bazo, M. Moreno, and J. Suriñach. 2004. "Growth and Externalities Across Economies. An Empirical Analysis Using Spatial Econometrics." In: Anselin, L., Florax, R., Rey, S.J. (Eds.) *Advances in Spatial Econometrics. Methodology, Tools and Applications*. Springer-Verlag, Berlin, pp. 433-455.

World Bank. http://www.worldbank.org/prospects/gep2004/toc.htm.

Table 3.1. Estimates of the Production Function with Si/Gun/Gu data

(Dependent variable : Log of output per worker, 1999-2006)

	FGS2SLS (Third step estimation)			2SLS Spatial Lag Model (First step estimation)		
	Coef.	Robust S.E.		Coef.	S.E.	
W^* ln(y/L)	**0.0978**	(0.0458)	**	**0.0803**	(0.0362)	**
PD	**0.0110**	(0.0067)	*	**0.0130**	(0.0068)	*
ln(K/L)	**0.0605**	(0.0115)	***	**0.0645**	(0.0114)	***
ln(M/L)	**0.7408**	(0.0159)	***	**0.7400**	(0.0164)	***
lnL	**-0.0007**	(0.0135)		**0.0013**	(0.0135)	
T	**0.0062**	(0.0024)	**	**0.0067**	(0.0017)	***
$\boldsymbol{\rho}$	**0.3066**					
Endogeneity	$\chi^2(1)$=25.1			$\chi^2(1)$=104.7		
Under identification	$\chi^2(1)$=1429.4			$\chi^2(1)$=649.3		
Weak identification	Cragg-Donald Wald F =5501.5			Cragg-Donald Wald F =8118.9		
Centered R^2	0.9873			0.9883		
Observations	1848			1848		

Notes: ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively. The endogeneity, weak identification, and under identification tests report Hausman test statistic, Anderson canonical correlations LM statistic, and Cragg-Donald Wald F statistic, respectively.

Table 3.2. Estimated TFP and Agglomeration Effects, 1999 and 2006

Region	Obs.	1999		2006		Annual growth rate (%)
		Mean	S.D.	Mean	S.D.	
Logarithm of raw TFP						
R1 (SIG)	65	0.5966	0.1617	0.6284	0.1425	0.74
R2 (GW CB DCN)	51	0.6786	0.1686	0.7119	0.1245	0.69
R3 (JB GJN JJ)	43	0.5772	0.1707	0.6581	0.1292	1.89
R4 (DGB BGN)	72	0.5961	0.1431	0.6310	0.1173	0.82
Logarithm of Agglomeration effect						
R1 (SIG)	65	0.3918	0.0919	0.4140	0.0849	0.79
R2 (GW CB DCN)	51	0.3307	0.0151	0.3566	0.0148	1.09
R3 (JB GJN JJ)	43	0.3342	0.0133	0.3620	0.0149	1.15
R4 (DGB BGN)	72	0.3510	0.0541	0.3759	0.0484	0.98

Table 3.3. Sources of Regional Variation in Raw TFP

	Pooled sample			10 percentile			Median			90 percentile		
	Coef.	S.E.		Coef.	S.E.		Coef.	S.E.		Coef.	S.E.	
Specification 1 (without interaction terms)												
Infra	0.3351	(0.1116)	***	0.3896	(0.1793)	**	0.2814	(0.1212)	**	0.3343	(0.1955)	*
FT	0.0345	(0.0084)	***	0.0395	(0.0089)	***	0.0287	(0.0086)	***	0.0353	(0.0125)	**
Clear day	-0.0003	(0.0002)		-0.0007	(0.0003)	**	-0.0004	(0.0003)		0.0001	(0.0004)	
R1	-0.3514	(0.1143)	***	-0.4321	(0.1856)	**	-0.2949	(0.1219)	**	-0.3272	(0.2031)	
R2	0.1206	(0.0191)	***	0.1073	(0.0244)	***	0.0959	(0.0188)	***	0.1585	(0.0316)	***
R3	-0.0561	(0.0223)	**	-0.0992	(0.0391)	**	-0.0754	(0.0249)	***	0.0061	(0.0438)	
D99_01	-0.0443	(0.0095)	***	-0.0398	(0.0095)	***	-0.0349	(0.0089)	***	-0.0583	(0.0177)	***
D05_06	0.2386	(0.0092)	**	0.0210	(0.0129)		0.0223	(0.0116)	*	0.0282	(0.0156)	*
cons	0.3304	(0.0794)	***	0.1766	(0.1099)		0.3973	(0.0761)	***	0.4601	(0.1357)	***
10^{th} per	-0.1594	(0.0050)	***									
90^{th} per	0.2023	(0.0061)	***									
R^2		0.9795			0.8884			0.8817			0.8211	
Obs.		96			32			32			32	
Specification 2 (with interaction terms)												
Infra	0.5529	(0.1420)	***	0.5599	(0.1871)	***	0.5525	(0.1065)	***	0.5463	(0.2613)	**
FT	0.0891	(0.0296)	***	0.1398	(0.0344)	***	0.0610	(0.0258)	**	0.0664	(0.0430)	
Clear day	0.0003	(0.0010)		0.0009	(0.0008)		-0.0003	(0.0009)		0.0004	(0.0014)	
Infra×FT	-0.0304	(0.0080)	***	-0.0282	(0.0096)	***	-0.0357	(0.0074)	***	-0.0272	(0.0122)	**
Infra×Clear	0.0003	(0.0003)		0.0008	(0.0004)	**	0.0001	(0.0004)		0.0000	(0.0006)	
FT×Clear	-0.0003	(0.0003)		-0.0008	(0.0003)	**	0.0000	(0.0003)		-0.0001	(0.0004)	
R1	-0.5186	(0.1247)	***	-0.6101	(0.1625)	***	-0.4754	(0.0989)	***	-0.4703	(0.2286)	*
R2	0.1425	(0.0225)	***	0.1278	(0.0245)	***	0.1210	(0.0159)	***	0.1788	(0.0378)	***
R3	-0.0815	(0.0234)	***	-0.1266	(0.0342)	***	-0.1028	(0.0215)	***	-0.0151	(0.0488)	
D99_01	-0.0505	(0.0105)	***	-0.0495	(0.0095)	***	-0.0395	(0.0095)	***	-0.0624	(0.0184)	***
D05_06	0.0352	(0.0097)	***	0.0308	(0.0102)	***	0.0361	(0.0088)	***	0.0388	(0.0167)	**
cons	0.0920	(0.1499)		-0.1521	(0.1650)		0.1911	(0.1108)	*	0.2800	(0.2206)	
10^{th} per	-0.1594	(0.0046)	***									
90^{th} per	0.2023	(0.0060)	***									
R^2		0.9818			0.9332			0.9257			0.8359	
Obs.		96			32			32			32	
LR test $\chi^2(3)$		11.21			16.42			14.89			2.77	

Notes: Robust standard errors are in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively.

Table 3.4. Elasticities of TFP with respect to Infrastructure and Freeness of Trade

	Pooled sample		10 percentile		Median		90 percentile	
Elasticity of TFP with respect to Infrastructure								
Average	**0.455**	**(0.0233)**	**0.511**	**(0.0463)**	**0.415**	**(0.0369)**	**0.431**	**(0.0383)**
R1	0.837	(0.0054)	0.945	(0.0128)	0.760	(0.0087)	0.789	(0.0090)
R2	0.255	(0.0028)	0.283	(0.0063)	0.233	(0.0043)	0.242	(0.0049)
R3	0.399	(0.0039)	0.439	(0.0080)	0.369	(0.0063)	0.383	(0.0074)
R4	0.330	(0.0030)	0.375	(0.0068)	0.298	(0.0045)	0.310	(0.0052)
Elasticity of TFP with respect to Freeness of trade								
Average	**0.087**	**(0.0049)**	**0.103**	**(0.0114)**	**0.078**	**(0.0085)**	**0.115**	**(0.0072)**
R1	0.014	(0.0027)	0.020	(0.0116)	0.002	(0.0029)	0.057	(0.0025)
R2	0.128	(0.0043)	0.151	(0.0110)	0.118	(0.0074)	0.146	(0.0096)
R3	0.114	(0.0036)	0.158	(0.0117)	0.088	(0.0044)	0.122	(0.0073)
R4	0.094	(0.0034)	0.083	(0.0111)	0.104	(0.0060)	0.135	(0.0086)

Note: Numbers in parentheses are standard deviations.

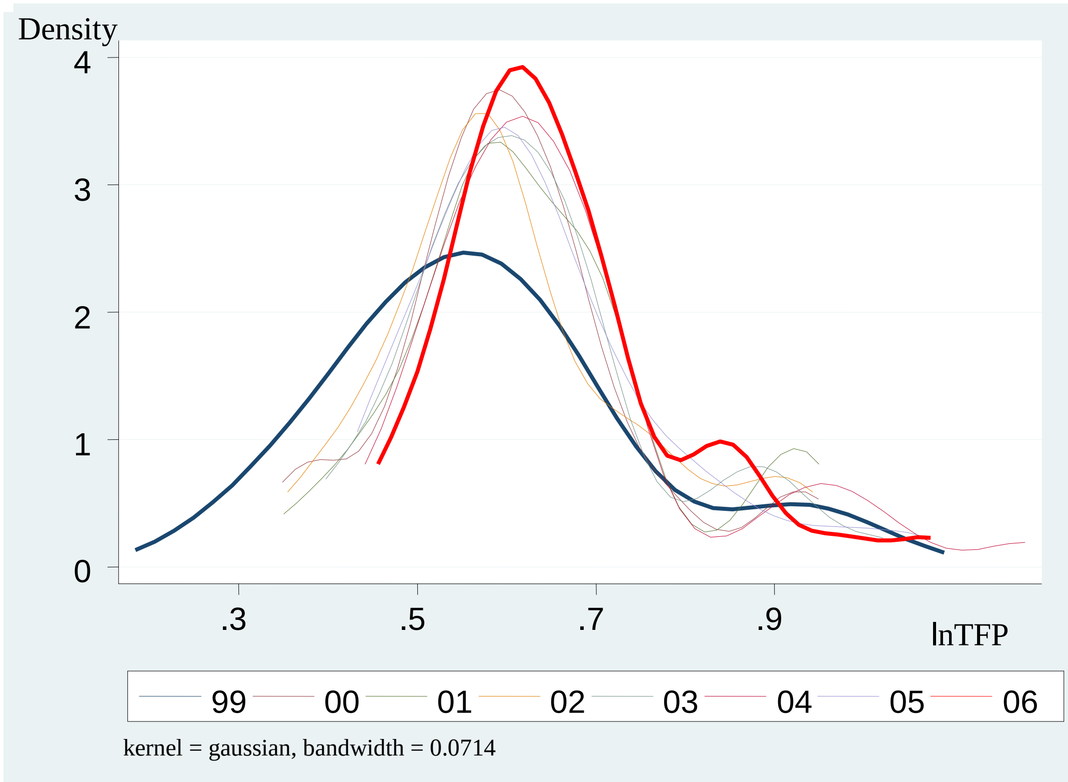


Figure 3.1. Kernel Density Estimate of Raw Productivity (*lnTFP*) for Region 3

Appendix

Table 3.A.1. Estimates of Korea's Bilateral Trade Costs during 1999 to 2006

Year	Tariff equivalent τ_{HF} (%)					**Freeness of trade**
	Korea -U.S.	Korea -Japan	Korea -China	Korea -Taiwan	**Weighted average**	
1999	0.42	0.44	0.47	0.57	**0.45**	2.25
2000	0.39	0.40	0.41	0.50	**0.40**	2.48
2001	0.43	0.42	0.42	0.54	**0.43**	2.32
2002	0.44	0.42	0.38	0.53	**0.42**	2.36
2003	0.43	0.38	0.32	0.49	**0.38**	2.61
2004	0.36	0.30	0.22	0.39	**0.30**	3.33
2005	0.39	0.31	0.22	0.40	**0.30**	3.28
2006	0.37	0.27	0.18	0.35	**0.27**	3.75
Percentage Change (%)	-11.31	-38.28	-61.57	-38.39	**-40.04**	66.78
Annual growth rate (%)	-1.70	-6.66	-12.77	-6.69	**-7.05**	7.58

Note: The estimate of bilateral trade cost between Korea and U.S. in 1999, 0.42 implies that, for instance, a Korean product (e.g. TV) sells for $1,420 in the U.S., while it sells for $1,000 in domestic market, and vice versa.

Table 3.A.2. Infrastructure and Amenities of Four Regions, 1999 to 2006

	1999	2000	2001	2002	2003	2004	2005	2006	Annual growth rate (%)
	Infrastructure (km/km^2)								
R1	1.54	1.59	1.66	1.69	1.69	1.72	1.79	1.78	2.07
R2	0.46	0.47	0.48	0.53	0.53	0.55	0.55	0.56	3.06
R3	0.73	0.74	0.77	0.82	0.83	0.86	0.89	0.89	3.02
R4	0.59	0.60	0.64	0.66	0.67	0.69	0.71	0.72	2.96
	Mean air temperature (℃)								
R1	12.8	12.6	12.6	12.6	12.5	13.0	12.0	12.9	0.04
R2	12.6	12.0	12.4	12.4	12.2	13.1	12.0	12.4	-0.25
R3	14.7	14.5	14.8	14.7	14.5	15.1	14.2	14.9	0.15
R4	14.1	14.0	14.4	14.0	13.8	14.7	13.7	14.1	0.03
	Precipitation (㎜)								
R1	1587.2	1225.0	1199.2	1217.8	1743.0	1341.2	1314.0	1448.2	-1.30
R2	1550.1	1365.1	946.9	1450.7	1750.9	1475.2	1386.2	1350.5	-1.95
R3	1918.5	1418.1	1171.2	1454.2	1925.5	1557.9	1250.5	1426.6	-4.14
R4	1859.8	1172.2	1038.3	1570.0	1975.5	1350.7	1149.2	1459.5	-3.40
	Clear days (days)								
R1	117.0	98.3	103.7	106.0	92.3	121.0	125.0	91.0	-3.53
R2	94.3	81.2	81.3	88.7	77.4	103.3	103.6	82.7	-1.85
R3	83.2	88.2	67.1	65.1	62.6	92.6	69.4	73.2	-1.81
R4	125.2	113.6	108.5	111.7	91.1	131.1	130.3	105.8	-2.37
	park/area (km^2/km^2)								
R1	0.0257	0.0261	0.0267	0.0272	0.0288	0.0319	0.0290	0.0309	2.68
R2	0.0057	0.0059	0.0060	0.0064	0.0065	0.0068	0.0074	0.0074	3.76
R3	0.0098	0.0098	0.0082	0.0078	0.0079	0.0083	0.0083	0.0085	-2.13
R4	0.0098	0.0112	0.0113	0.0114	0.0111	0.0116	0.0123	0.0118	2.62
	forestry/area (km^2/km^2)								
R1	0.5383	0.5364	0.5337	0.5315	0.5294	0.5273	0.5257	0.5231	-0.41
R2	0.7114	0.7110	0.7103	0.7095	0.7089	0.7083	0.7077	0.7071	-0.09
R3	0.5755	0.5745	0.5736	0.5717	0.5708	0.5701	0.5686	0.5675	-0.20
R4	0.6969	0.6962	0.6957	0.6952	0.6948	0.6944	0.6937	0.6931	-0.08
	river/area (km^2/km^2)								
R1	0.0390	0.0390	0.0389	0.0390	0.0388	0.0388	0.0389	0.0389	-0.06
R2	0.0254	0.0253	0.0252	0.0252	0.0251	0.0251	0.0252	0.0252	-0.07
R3	0.0236	0.0236	0.0236	0.0236	0.0236	0.0236	0.0236	0.0245	0.02
R4	0.0305	0.0305	0.0304	0.0304	0.0304	0.0304	0.0304	0.0304	-0.04

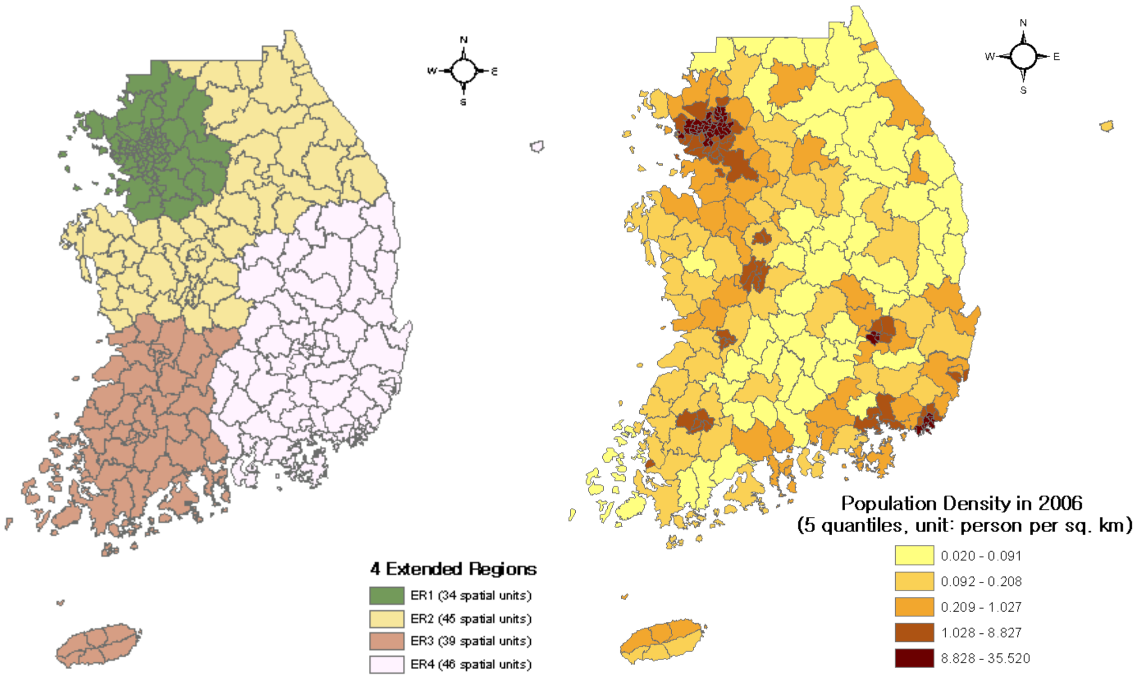


Figure 3.A.1. Four Regions and Population Density in 2006

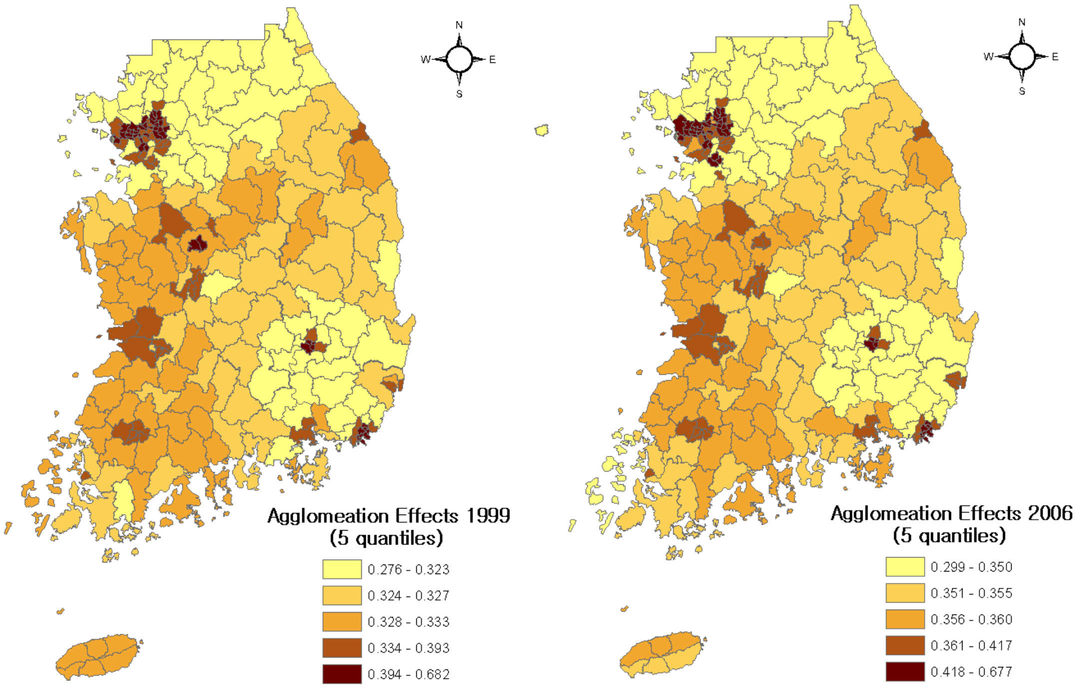


Figure 3.A.2. Estimates of Agglomeration Effects in 1999 and 2006

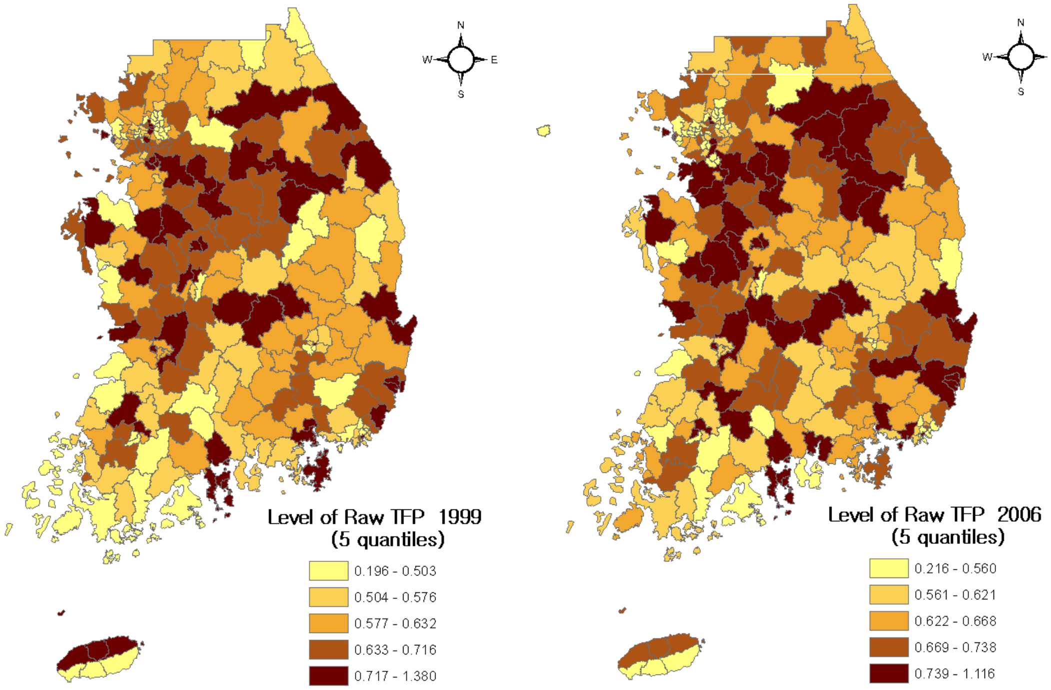


Figure 3.A.3. Estimates of Raw TFP Levels in 1999 and 2006

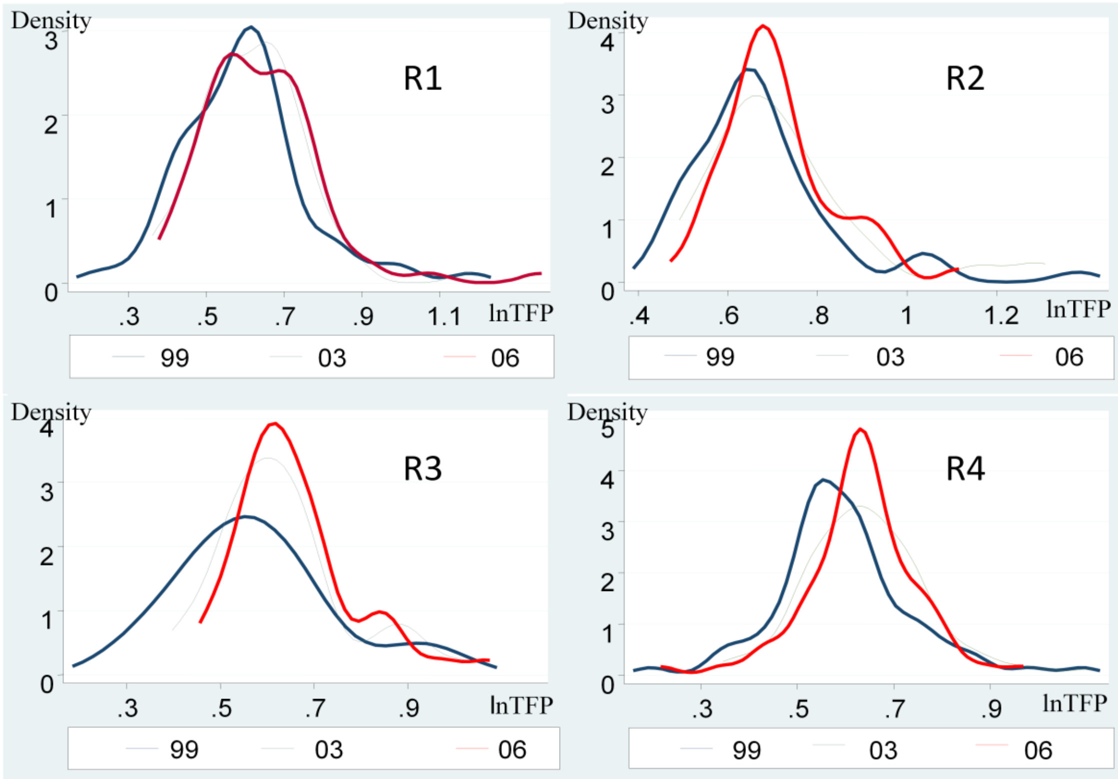


Figure 3.A.4. Kernel Density Estimate of Raw Productivity (*lnTFP*) for All Regions

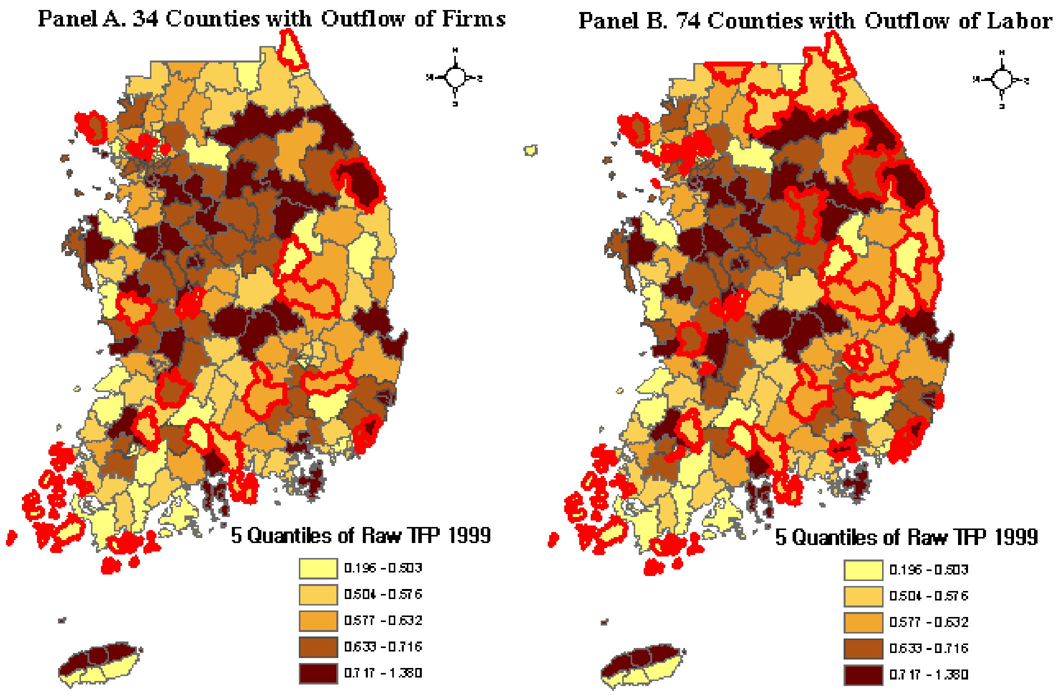


Figure 3.A.5. Counties with More Than 10% Firm Exit Rate (Panel A) or Net Employment Loss (Panel B) during 1999-2006

CHAPTER FOUR

Trade Costs and Regional Productivity in Indian Manufacturing

4.1. Introduction

Concentration of economic activities in specific regions of a country has been the core of the new economic geography literature (Krugman 1991, Fujita et al. 1999, Baldwin et al. 2003, Combes et al. 2008). Such concentration is accompanied by higher output, wages and employment relative to less concentrated regions (Henderson et al. 2001, Glaeser and Gottlieb 2009). Attempts to identify the sources of regional inequalities find that spatial variation in economic activity is largely due to the corresponding productivity difference among regions (Rice et al. 2006, Acemoglu and Dell 2010). Hence, in the search for solutions to regional equity in economic development, the question why we observe large inter-regional productivity differences takes on significance.

To date, only a few studies have focused on the sources of productivity variation within a country (Ciccone 2002, Rosenthal and Strange 2004, Rice et al. 2006, Mitra and Ural 2008). Most of these studies focus on the regional productivity effects of agglomeration, i.e. concentration of economic activities. That is, agglomeration is associated with productivity-enhancing externalities such as easier access to intermediate inputs and complementary services, availability of a large labor pool with multiple specialization, inter- and intra-industry information transfers, and co-use of general infrastructure. For example, Rice et al. (2006) find that a region's labor productivity increases by 3.5% when its working-age population is doubled. The above studies have improved our understanding of the contribution of agglomeration forces to regional productivity has increased. However, agglomeration's importance relative to the underlying technical change, fostered by competition and innovation forces, has received limited research attention (Oosterhaven and Broersma 2007, Glaeser and Gottlieb 2009).

The role of international competition in enhancing industry productivity is highlighted in the emerging heterogeneous-firms trade theory and its applications (Melitz 2003, Bernard et al. 2003, Melitz and Ottaviano 2008). Such models suggest that international competition forces least productive firms to exit, reallocates resources to highly productive firms, and increases average industry productivity (Syverson 2004, Bernard et al. 2007). Similarly, the effect of infrastructure on regional economic development has been well documented (Rosenthal and Strange 2004). By lowering transport costs, infrastructure improvement is often regarded as an increase in domestic competition between manufacturing firms, producing similar goods and being located at any region within a nation (Behrens 2006). Additional factors in the determination of productivity variation among regions include natural amenities, which can negatively or positively impact productivity (Rappaport 2009).

The objective of this study is to investigate the relative importance of agglomeration and pure technical change (or raw productivity) in regional productivity of Indian manufacturing industries. In addition, we identify and quantify the determinants of variation in regional raw productivity. Being a large and developing country, India exhibits substantial regional variation in economic development. For instance, the per capita net domestic product ranges between $217 and $1,932 among Indian states in 2006.[31] Advances in spatial econometrics allow us first to attribute firms' productivity growth to externalities within and across spatial units and to the underlying technical change or raw productivity. Then, we examine the role of foreign competition and domestic infrastructure as key determinants of regional raw productivity variation. Our application focuses on Indian manufacturing using firm-level data during 1994-2007. The sample period is the aftermath of significant (unilateral) trade reforms in Indian manufacturing (Krishna and Mitra 1998, Chari and Gupta 2008, Sivadasan 2009, Goldberg et al. 2009). Moreover, the Uruguay Round of multilateral negotiations (1994) and China's accession to WTO (2000) have added to the globally competitive environment faced by Indian manufacturing firms. While previous studies have explored the effects of trade liberalization at the industry level, few have focused on the consequences for spatial distribution of productivity in Indian manufacturing and their implications for spatially balanced economic development (Tybout 2000, Schor 2004).[32]

The rest of the chapter is organized as follows. In the next section, the theoretical basis for our application is outlined. Then, we detail the spatial model to compute firm-level total factor productivity. Firms are grouped by regions and measures of regional productivity distribution are derived. Following the description of our data in section 4.4, we discuss the results from estimating the production function and the sources of regional productivity variation. The final section summarizes and provides concluding remarks.

---

[31] The state of Goa had the highest per capita income at Rupees (Rs) 87,501 and Bihar the lowest at Rs. 9,817, both evaluated at the exchange rate of Rs. 45.28 per US $ in 2006.

[32] An exception here is Mitra and Ural (2008) who focus on state-level productivity in Indian manufacturing using industry-level data.

4.2. Theoretical Basis

Consider an industry with many firms, each indexed by its unique level of productivity and producing a differentiated variety. That is, each firm has monopolistic market power, faces a constant marginal cost inversely proportional to its productivity and a fixed entry cost. As illustrated in Melitz (2003), high productivity implies producing a variety at lower marginal cost or producing a more expensive variety at a given cost level. Then, we can define the *k*-th industry's *ex-post* productivity distribution of firms located in the *r*-th region of a country at *t*-th period as a density truncated from below as follows:

(1) $$f_{krt}(\omega) = g_{krt}\left(\omega \middle| \omega \geq \omega^*_{jrt}\right) = \frac{g_{krt}(\omega)}{Prob(\omega \geq \omega^*_{krt})},$$

where $g_{krt}(\omega)$ is the underlying *ex-ante* productivity distribution for industry *k* and region *r* at time *t* and ω^*_{jrt} is the industry-region-time specific cutoff productivity level (Melitz and Ottaviano 2008). We assume that there exists an overall and common productivity distribution for the country, $g(\omega)$.

Each firm incurs a sunk-entry cost into this industry- and region-specific market to obtain a draw from $g_{krt}(\omega)$. An entrant drawing a productivity level less than ω^*_{krt} faces negative profit and will immediately exit from the market. Note that, the zero-profit cutoff productivity, ω^*_{krt}, is determined by the short-run zero-profit condition for monopolistic firms and the long-run free entry condition (Melitz 2003). The entry and exit of firms result in the equilibrium productivity distribution, truncated from below, as in equation (1).

In this study, we specify the cutoff productivity level, ω^*_{krt}, as a function of indicators of competition in the corresponding market:

$$(2) \qquad \omega_{krt}^{*} = \omega^{*}(FC_{kt}, DC_{rt}),$$

$$\partial\omega_{krt}^{*}/\partial FC_{kt} > 0 \; and \; \partial\omega_{krt}^{*}/\partial DC_{rt} > 0,$$

where FC_{kt} and DC_{rt} are variables or vectors denoting foreign and domestic competition, respectively.[33] Changes in these exogenous factors that shift the cutoff productivity levels, e.g. tougher competition, further alter the truncation point of the distribution (Melitz 2003, Melitz and Ottaviano 2008). Consequently, the average productivity in that industry can increase without any individual firm's productivity improvement or a shift of $g(\omega)$.

The competition pressures, however, have asymmetric effects on individual firm's performance such as entry and exit, market share, and profitability according to its realized productivity. The low-productivity firms will exit and any survivors along with new entrants serve only the local market (Local Firm - LF). These low-productivity firms will lose both market share and profit. The mean-productivity firms entering into other local markets within a country (Domestic Firm - DF) will increase their market share but lose profit on account of falling prices. However, high-productivity firms additionally entering foreign markets (Exporting Firm- EF) will increase their market share but can still lose profit, and among exporting firms only the highest-productivity firms will increase both market share and profit.[34] Subsequently, productive resources, assumed to be mobile across firms and industries (labor and capital), are reallocated from LF to DF and EF.

Each cutoff productivity level dividing firms into each category as noted above, can

---

[33] For analytical simplicity, the former is assumed to be common across regions, while the latter does not vary across industries. We relax these assumptions in our empirical approach.

[34] For details of mathematical and graphical derivation, see Melitz (2003) pp. 1714.

be written as a function of ω^{*}_{krt};

(3) $$\omega^{**}_{krt} = \omega^{**}\big(\omega^{*}(FC_{kt}, DC_{rt})\big), \quad \partial\omega^{**}/\partial\omega^{*} < 0,$$

(4) $$\omega^{***}_{krt} = \omega^{***}\big(\omega^{*}(FC_{kt}, DC_{rt})\big), \quad \partial\omega^{***}/\partial\omega^{*} < 0,$$

where ω^{**}_{krt} is the cutoff level between LF and DF, while ω^{***}_{krt} is the counterpart between DF and EF. Thus, the average productivity of each category is given by:

(5) $$\overline{\omega}^{LF}_{krt} = \int_{\omega^{*}_{krt}}^{\omega^{**}_{krt}} f_{krt}(\omega)d\,\omega, \quad \overline{\omega}^{DF}_{krt} = \int_{\omega^{**}_{krt}}^{\omega^{***}_{krt}} f_{krt}(\omega)d\,\omega, \quad \overline{\omega}^{EF}_{krt} = \int_{\omega^{***}_{krt}}^{\infty} f_{krt}(\omega)d\,\omega,$$

Based on Melitz (2003), we can show that $\partial\overline{\omega}^{LF}_{krt}/\partial FC_{kt} > 0$, $\partial\overline{\omega}^{LF}_{krt}/\partial DC_{rt} \gtrsim 0$, $\partial\overline{\omega}^{DF}_{krt}/\partial FC_{kt} < 0$, $\partial\overline{\omega}^{DF}_{krt}/\partial DC_{rt} > 0$, $\partial\overline{\omega}^{EF}_{krt}/\partial FC_{kt} < 0$, and $\partial\overline{\omega}^{EF}_{krt}/\partial DC_{rt} < 0$, implying that as competition increases, average productivity of LF increases but that of EF decrease while that of DF changes ambiguously[35]. However, comparing the magnitudes of productivity improvement induced from two different sources (foreign and domestic competitions) is an empirical issue.

With the above theoretical basis, we identify the competition-productivity nexus in Indian manufacturing. Moreover, we quantify the relative contribution of the international and domestic competition pressures on aggregate productivity in Indian manufacturing sector, focusing on variations of firms' productivity distribution across regions and industries over time. We measure the shifts in the spatial/regional productivity distribution by changes in its mean, median, and alternative percentile values (e.g. 10^{th}, 20^{th}, 50^{th}, 80^{th} and 90^{th} percentiles). These alternative values are regarded as representing the average productivity of low-productivity firms (or LF), mean-

---

[35] Again note that this is the case where individual firms' productivity would not be improved by these competition sources.

productivity firms (or DF) and high-productivity firms (or EF), respectively. Regressing each alternative value of productivity, i.e. quantile regression, on two competition proxies with other controlling variables, allows us to infer on competition pressures shaping the spatial productivity distribution.

4.3. Empirical Strategy

In this section, we first describe our approach to estimating firm-level productivity. Then, we show how to derive measures of raw productivity distribution of firms belonging to the same industry and region for each time period. Finally, we present the empirical approach to examine the sources of variation in regional productivity.

4.3.1. Estimation of Firm Productivity

The first step of our empirical analysis is to estimate of firm-level productivity for nine manufacturing industries in India (Appendix Table 1). More specifically, we quantify total factor productivity (TFP) of each firm and attribute it to agglomeration forces including spatial spillovers and the underlying pure technical change or raw productivity.

Following Henderson (2003), Lall et al. (2004) and others, we begin with a Cobb-Douglas production function with external economies. In order to capture possible inter-industry benefits from co-location of firms, or urbanization economies, we use a local sum of output of all firms in other industries within a spatial unit. Positive externalities from other firms in the same industry or localization economies are also measured by employing the spatial lag model, i.e. by adding a spatially lagged dependent variable to the list of explanatory variables (Anselin, Florax and Rey 2004). Furthermore, we extend Levinsohn and Petrin's (2003) approach to control for two types of simultaneity: between conventional input levels and productivity, and between productivity and self-selection of high productivity firms into densely agglomerated locations (Saito and Gopinath 2009). The former arises because productivity is known to the profit maximizing firms when they choose their input levels. Therefore, the marginal product of an input increases with a positive productivity shock, which would result in a change in the allocation and optimal level of conventional inputs. The latter simultaneity arises when agglomeration economies enhance firm productivity and high-productivity firms also seek out more attractive locations (Baldwin and Okubo 2006).[36] Thus, firm-level raw productivity should be included as a right-hand-side variable in the estimation of production function for an industry. The above methodology allows us to obtain consistent estimates of the coefficients on conventional inputs and locational attributes, and thus, unbiased productivity measures:

(6) $$y_{it} = \beta_0 + \beta_u U_{it} + \beta_s W_{it}^{'} y + \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_e e_{it} + \omega_{it} + \varepsilon_{it} ,$$

where y_{it} is the output of the *i*-th (*i=1,2,…N*) firm at time *t* (*t=1,2,…T*) as an element of vector y, U_{it} is the variable representing external influences on production from urbanization economies, and W_{it}' is the *it*-th row vector of spatial weights matrix W_{NTxNT}. Following Anselin, Florax and Rey (2004), we model the spatial interaction between firms in the same industry through a distance-based W matrix, whose diagonal elements are zero and the off-diagonal elements equal to w_{ij}, which is the relative weight of firm *j* for firm *i*, given by the inverse of the distance between the two firms. Additionally, we assume there is no spatial interaction between firms apart from each other by above 50

[36] If high-productivity firms were disproportionately found in agglomerated areas, this would lead to an overestimation of the agglomeration effects on firm productivity.

km.[37] The variables l_{it}, k_{it}, m_{it}, and e_{it} are *i*-th firm's labor, capital stock, materials, and energy at time *t*, respectively; all lower case variables are expressed in natural logarithm. The variable ω_{it} denotes raw productivity of the firm, i.e. pure technology effect, since the firm-specific agglomeration effects are already controlled for. Hence, φ_{it} is an index of raw TFP capturing Hicks-neutral technical change. Finally, ε_{it} is an *i.i.d.* disturbance term.

Before addressing the self-selection problem, we consider the potential endogeneity resulting from $\text{cov}[U_{it},\ \varepsilon_{it}] \neq 0$ and/or $\text{cov}[W'_{it}y,\ \varepsilon_{it}] \neq 0$ (Anselin, Florax and Rey 2004). They imply the existence of another channel through which locational attributes interact with firm production other than with productivity.[38] We obtain predicted values for these two agglomeration variables through two-stage least squares (2SLS) using their own spatial lag and that of the independent variables as instruments.[39] Then, the predicted values, U^*_{it} and $W'_{it}y,^*$ are used in the estimation of productivity.

Following Levinsohn and Petrin (2003), we choose material input as a proxy to correct for both the simultaneity and self-selection biases, since the demand for materials is monotonic in the firm's productivity for all relevant levels of capital and locational attributes. Hence, by inverting the raw material demand firm's productivity can be

---

[37] Thus, off-diagonal elements for these firms are zero. This assumption is based on the empirical results of Rosenthal and Strange (2003), who show that localization economies attenuate rapidly over the first few miles and then attenuate much more slowly thereafter.

[38] For instance, a firm in a larger market is more likely to stay in the market despite a low productivity shock than a firm in a smaller market because the former could be compensated for its low productivity by higher external effects. That is, the positive correlation between market size and firm survival possibility for a given productivity shock will cause the coefficients on the locational attributes to be biased downward.

[39] Kelejian and Prucha (1999) suggest a linearly independent subset of exogenous variables and their spatial lags as instruments for the spatially-lagged dependent variable.

expressed as a function of material input, capital stock and our two variables capturing agglomeration effects:

(7) $$\omega_{it} = \omega(m_{it}, k_{it}, U_{it}^{*}, W_{it}' y^{*}).$$

By substituting equation (7) into (6) we have:

(8) $$y_{it} = \beta_l l_{it} + \beta_e e_{it} + \varphi(m_{it}, k_{it}, U_{it}^{*}, W_{it}' y^{*}) + \varepsilon_{it} ,$$

where $\varphi(\cdot) = \beta_0 + \beta_u U_{it}^{*} + \beta_s W_{it}' y^{*} + \beta_k k_{it} + \beta_m m_{it} + \omega(m_{it}, k_{it}, U_{it}^{*}, W_{it}' y^{*})$.

In the first stage, by employing a second-order polynomial series approximation for $\varphi(\cdot)$, the coefficients on the variable inputs, β_l and $\beta_e$can be consistently estimated using the ordinary least squares (OLS) procedure (Yasar and Morrison 2007, Saito and Gopinath 2009).

For consistent estimation of the remaining coefficients in the second stage, we need additional assumptions that a firm's productivity follows a Markov process and that capital stock and locational attributes adjust to a productivity shock with a time lag. Then, in the second stage, equation (6) becomes:

(9) $$\mathring{y}_{it} = \beta_0 + \beta_u U_{it}^{*} + \beta_s W_{it}' y^{*} + \beta_k k_{it} + \beta_m m_{it} + \mathrm{E}[\omega_{it} | \omega_{it-1}] + \xi_{it} + \varepsilon_{it},$$

where $\mathring{y}_{it} = y_{it} - \hat{\beta}_l l_{it} - \hat{\beta}_e e_{it}$ is defined as output net of variable inputs' contribution, and ξ_{it} is an innovation in the first-order Markov process of the productivity shock.[40] Since firms' decisions for the state variables are made prior to choosing intermediate inputs which immediately respond to innovations in their productivity, $E[\xi_{it} | U_{it}^{*}] = E[\xi_{it} | W_{it}' y^{*}] = E[\xi_{it} | k_{it}] = 0$, while $E[\xi_{it} | m_{it}] \neq 0$. Note that β_0 is not separately identified from the mean of $\mathrm{E}[\omega_{it} | \omega_{it-1}]$ without additional restrictions (Levinsohn and

---

[40] $\hat{\beta}_l$ and $\hat{\beta}_e$are the OLS estimates from the first stage.

Petrin 2003). Hence, to estimate $\beta_0 + \mathrm{E}[\omega_{it}|\omega_{it-1}]$ as a whole, we use the estimates of $\beta_0 + \omega_{it-1}$ obtained from the first stage results as:

(10) $$\beta_0 + \omega_{it-1} = \hat{\varphi}(\cdot) - \beta_u U^*_{it-1} - \beta_s W'_{it-1} y^* - \beta_k k_{it-1} - \beta_m m_{it-1}.$$

Since $E[\xi_{it}|m_{it}] \neq 0$ and $E[\xi_{it}|m_{it-1}] = 0$, consistent estimates of β_u, β_s, β_k, and β_m are obtained by a nonlinear instrument variables (NIV) estimator with a third-order approximation for $\beta_0 + \mathrm{E}[\omega_{it}|\omega_{it-1}]$ as follows (Saito and Gopinath 2009)[41]:

(11) $$\overset{\circ}{y}_{it} = \beta_u U^*_{it} + \beta_s W'_{it} y^* + \beta_k k_{it} + \beta_m m_{it} + \sum_{p=0}^{3} \delta_p (\hat{\varphi}(\cdot) - \beta_u U^*_{it-1} - \beta_s W'_{it-1} y^* - \beta_k k_{it-1} - \beta_m m_{it-1})^p + \varepsilon^*_{it},$$

where $\varepsilon^*_{it} = \xi_{it} + \varepsilon_{it}$.

Since the estimation method involves several steps, deriving the appropriate analytical standard errors is nontrivial. Levinsohn and Petrin (2003), therefore, propose a bootstrap procedure. The square root of the variance of estimated parameters across samples is used to compute the standard error. Finally, we obtain the raw productivity of firm *i* at time *t* by:

(12) $$\Omega_{it} = \exp(\omega_{it}) = \exp(y_{it} - \hat{\beta}_u U^*_{it} - \hat{\beta}_s W'_{it} y^* - \hat{\beta}_l l_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_m m_{it} - \hat{\beta}_e e_{it}).$$

Note that Ω_{it} is distinguished from external effects that arise from agglomeration economies.

After grouping firms according to regions/states (where they are located) and to time periods (when they operate), we obtain the mean, median, and alternative percentile values of firm productivity for each group (industry-region-year combination).

---

[41] The exogenous instruments for m_{it} are a subset of m_{it-1}, e_{it-1}, k_{it-1}, and $\hat{\varphi}_{it}$.

4.3.2. Sources of Regional Productivity Variation

Our interest here is to analyze factors contributing to shifts or changes in the raw-productivity across industries as well as across regions over time. In particular, we seek to determine the role of international competition and domestic infrastructure in shaping the raw productivity of regions (Henderson et al. 2001, Behrens 2006). The idea that reductions in international trade costs bring about increased productivity is commonly found in the trade theory especially in the new heterogeneous-firms trade model (Melitz 2003, Bernard et al. 2003, Melitz and Ottaviano 2008). Empirical evidence of a trade-cost or competition effect on productivity is beginning to accumulate (Greenaway, Morgan and Wright 2002, Syverson 2004, Amiti and Konings 2007, Bernard et al. 2007, Novy 2008).

Similarly, less-developed domestic infrastructure may limit intra-national trade between regions, particularly in a large and developing country such as India. In this case, most firms serve geographically limited local markets due to high transport costs across regions and face less competition from outside of their local markets. Therefore, improvement of infrastructure in a domestic region may also affect regional productivity of local firms similar to the effect of foreign competition. Because of decreased transportation costs, some high-productivity firms in neighboring regions penetrate into the given region whose market becomes easier to access, which causes least-productivity local firms to exit. In the same token, high-productivity local firms enjoy extension of their markets and some of them may have higher profit, coincidental with improved local infrastructure.[42]

[42] Both international and domestic competitions bring in cheaper intermediate inputs to local firms. The net effect of these competitive forces, i.e. increased output market competition and possibly lower intermediate prices, is again an empirical question.

To quantify how much these two competitive forces affect raw productivity variation, we specify the alternative percentile values of region- and industry-specific productivity as a function of indicators of industry-specific trade cost (TC_{kt}) and region-specific infrastructure (IN_{rt}) with other controlling variables and three-way fixed effects:

$$(13) \qquad \Omega^p_{krt} = \delta^p_0 + \delta^p_1\, TC_{kt} + \delta^p_2\, IN_{rt} + \delta^p_3\, Control_{krt} + d^p_k + d^p_r + d^p_t + \nu_{krt},$$

$$p = 10\%, 20\%, 50\%, 80\%, 90\%,$$

where the superscript *p* refers to five alternative percentiles, and the subscripts *k*, *r*, and *t* refers to industry, region and time, respectively. Hence, for example, Ω^{10}_{krt} is the 10th percentile productivity of firms which belong to industry *k*, are located in region *r*, and operate at time *t*. The three-way fixed effects capture unobservable heterogeneity in the industry, region, and time dimensions, i.e. macroeconomic business cycle, domestic industrial policy, regional characteristics including labor laws, and natural amenities.

As discussed in the theoretical framework, the coefficients δ^p_1 and δ^p_2 represent the effects of trade costs and infrastructure, respectively, on the shift of productivity distribution for each alternative percentile group. From the view of the heterogeneous firm trade model, the 10th (or 20th) percentile productivity would increase much more than others do in response to an increase in competition (Melitz 2003).

In the application of equation (13), we find that some industry-region-year combinations have zero firms. In addition, a few other regions have too small a number of firms in an industry to construct reliable measures of regional productivity measures (Ω^p_{krt}, $p = 10\%, 20\%, 50\%, 80\%, 90\%$). Hence, we exclude industry-region-year combinations with fewer than 5 firms, which results in zero values for Ω^p_{krt}. Estimating equation (13) with positive Ω^p_{krt} alone would yield biased estimates of coefficients of interest. To avoid this problem, we use a censored regression model, where the observable variable Ω_{krt}^{p} equals the latent variable (Ω_{krt}^{p*}) only when the latter takes a positive value. The left-censoring limit is then defined as a lower bound productivity (L^{p}) which is slightly smaller than the minimum productivity observed in each percentile sample. It is more appropriate than setting the threshold of zero because our firm data only cover the organized sector in Indian manufacturing, and thus the missing values for some combinations do not necessarily mean zero production or productivity in corresponding regions. In order to account for the censoring of regional productivity like in equation (14), Tobit and Powell's (1984) censored least absolute deviations (CLAD) estimator are employed.

$$(14) \qquad \Omega_{krt}^{p} = \begin{cases} \Omega_{krt}^{p*} & \text{if } \Omega_{krt}^{p*} > 0 \\ L^{p} & \text{otherwise} \end{cases},$$

$$\Omega_{krt}^{p*} = \delta_0^{p} + \delta_1^{p}\, TC_{kt} + \delta_2^{p}\, IN_{rt} + \delta_3^{p}\, Control_{krt} + d_k^{p} + d_r^{p} + d_t^{p} + \nu_{krt},$$

$$p = 10\%, 20\%, 50\%, 80\%, 90\%.$$

4.4. Data

4.4.1. Data for Productivity Estimation and Firm Characteristics

We use firm-level data from the Prowess database of the Centre for Monitoring Indian Economy (CMIE), which contains information on income statements and balance sheets of all registered firms in India from 1988 to 2007. CMIE collects annual data on all firms in the organized sector, which refers to registered companies that submit financial statements. The CMIE firms together account for more than 70% of industrial output, 75% of corporate taxes, and more than 95% of excise taxes collected by the Government of

India.[43] Focusing on manufacturing firms after early 1990's trade reform, we construct an unbalanced panel of 61,805 observations with 8,462 firms whose location information is not missing as well as whose input and output values are nonzero during our sample period 1994-2007. Based on the Indian National Industrial Classification (NIC) system (2004), we classify firms into nine manufacturing industries by regrouping two or three digit NIC sub-industries which exhibit similar production structure. For example, the food industry combines NIC 15 (food products and beverages) and NIC 16 (tobacco products). Details on our grouping are in Appendix Table 4.A.1.

Since the Prowess database provides each firm's sales revenue, capital spending and input expenditures rather than quantities of output and inputs, we need to find appropriate price deflators for each of them. The value of output, 'Sales income' is deflated by a region-specific price deflator constructed from the series on net state domestic product for registered manufacturing.[44] The item 'Compensation to employees' is converted into a measure of labor input after deflation by regional consumer price index for industrial workers.[45] Real capital stock is calculated by deflating net value of fixed assets with a deflator based on the state-wise net domestic product for industrial sector.[46]

---

[43] The Prowess database has now been used in many studies including Krishna and Mitra (1998), Khanna and Palepu (1999), Bertrand et al. (2002), Fisman and Khanna (2004), Topalova (2007), Dinc and Gupta (2007), Chari and Gupta (2007), Goldberg et al. (2008, 2009), and Alfaro and Chari (2009).

[44] The price deflator is computed as the ratio of current price 'net state domestic product at factor cost' to constant price one which are obtained from national accounts statistics.

[45] The index is from "Center-wise Consumer Price Index Numbers for Industrial Workers, base 1982=100 and 2001=100", Ministry of Commerce and Industry, Government of India. Center weights in all India are used to construct state-wise consumer price index and the indexes from two different bases are connected with adapting factor provided.

[46] Krishna and Mitra (1998) also use net fixed assets as a measure of capital, while Topalova (2007) and Goldberg et al. (2008) apply the perpetual inventory method after constructing a revaluation factor in order

Each firm's material (including intermediate goods) expenditure is sum of the four terms: raw materials, packaging, purchase of finished goods, and advertising/ marketing and distribution expenditure. Deflators for the first three terms and for the other are similarly calculated with the state-wise net domestic product for non-service sector (agriculture and industry) and service sector, respectively. Finally, the energy input comprises fuel, electricity, and water, thus the state-wise net domestic product for electricity, gas and water supply is used to obtain a quantity measure.

The variables representing agglomeration economies are also constructed using location information of the firms covered in the Prowess database, since regional data such as population density, industry-specific employment and number of establishments, are unavailable for all sample years at the district level in India. Therefore, we treat externalities as being among firms in the organized sector only. Using the six-digit postcode of each firm's headquarter, we sum the real value of output of all firms in manufacturing, services, utilities, and financial industries within each three-digit postcode area.[47] For each firm located in any three-digit postcode area, this local output sum of all industries is the proxy for potential inter-industry benefits from co-location of others firms, i.e. urbanization economies.[48] As in related literature, economic activity in neighboring three-digit postcode areas is assumed to have no effect on the area in question.

---

to convert balance-sheet data at historic cost into a measure of capital at replacement cost.

[47] The number of three-digit postal areas in India is 408, which is almost two thirds of that of districts (593) based on the 2001 Census. The average area of the postal area is about 80,000 km^2, which is 1.5 times larger than that of an average district.

[48] Our urbanization measure is specific to the industry and region. For example, two firms in different industries at the same area face different urbanization economies in size, while two firms in the same industry at the same area identically benefit from urbanization economies.

In contrast, localization economies are measured by employing the spatial lag model (Anselin, Florax and Rey 2004). That is, positive externalities from other firms in the same industry are reflected in the production function by adding a spatially lagged output variable to the set of explanatory variables. The spatial interaction among firms in industry *k* is modeled through the distance-based W^k matrix. For a single cross-section of n_t^k observations at time *t*, a $n_t^k \times n_t^k$ spatial weighting matrix W_t^k is defined in terms of the distance decay function, d_{ij}^{-1}. Here, d_{ij} is the great-circle distance between firms *i* and *j*. All distances among n_t^k firms are computed using ArcGIS, after mapping the six-digit postcodes of firm locations into the India map with six-digit postal code points. Since spatial interaction beyond 50 *km* is assumed to be asymptotically zero, the corresponding d_{ij}^{-1} becomes zero. Given our unbalanced panel setting ($N^k = \sum_{t=1}^{T} n_t^k$), the $N^k \times N^k$ matrix W^k is a block diagonal matrix, in which each diagonal element is W_t^k. Finally, each row is divided by the sum of the row elements, i.e. W^k is adjusted to be row-standardized as proposed in Anselin (1988).[49] With the spatial weight matrix, we obtain the spatially lagged dependent variable $W^k y$ for the estimation of production function for industry *k*, where y is a column vector whose element is y_{it}, the logarithm of output of firm *i* at time *t*. Descriptive statistics on the nine manufacturing industries, including measures of agglomeration economies are given in Appendix Table 4.A.2.

Our study considers 18 Indian states as regions: Andhra Pradesh, Assam, Bihar, Goa, Gujarat, Haryana, Himachal Pradesh, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, West Bengal and the

[49] We construct this matrix (W^k) for each of the nine manufacturing industries.

Union Territory of Delhi. Jammu and Kashmir and most of north-eastern states (Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura) are excluded from the sample due to lack of adequate observations. Jharkhand, Chhattisgarh, and Uttaranchal, created out respectively of Bihar, Madhya Pradesh, and Uttar Pradesh, are treated as being parts of the original states for our sample years after their creation in 2000. Except for Goa and the Union Territory of Delhi, the average area of Indian states is more than 150,000 km^2. Our regional classification is consistent with the varying environment in which firms operate, e.g. level of economic development, institutions, geography and infrastructure. For instance, difference in state labor laws may limit the effects of tough competition coming through a state or national border. The number of firms categorized both by nine industries and 18 states is provided in Appendix Table 4.A.3.[50]

4.4.2. Measurement of Trade Cost and Infrastructure

Following Novy (2008), we estimate industry-specific trade costs, which include trade and geographic barriers, in Indian manufacturing industries during 1993-2006. With observable trade and output data, Novy (2008) derives bilateral trade costs in the presence of multilateral resistance as in Anderson and van Wincoop (2003). Thus, for a specific industry (k), bilateral intra-industry trade costs between India (I) and a foreign country (c)

---

[50] Prowess also provides plants' locations and product codes for each firm. Thus, we can divide firms into four groups; single-plant firm with single-product (SS), single-plant firm with multi-products (SM), multi-plant firm with single-product (MS), and multi-plant firm with multi-products (MM). Our empirical application is conducted for each of these four samples as well as for the pooled sample to check how the role of competition pressures in enhancing firm-level productivity in response to different production structures (strategies) in terms of multi-location and multi-products. However, we only report the results from the pooled sample due to space constraints.

affect trade flow in both directions ($X_{Ic}^{k},\ X_{cI}^{k}$), while intra-country trade ($X_{II}^{k},\ X_{cc}^{k}$) can be used as a size variable controlling for multilateral resistance. A measure of industry-specific and tariff-equivalent bilateral trade costs τ_{cI}^{k}, is given by:

$$(15) \qquad \tau_{cI}^{k} = \left(\frac{t_{Ic}^{k} t_{cI}^{k}}{t_{II}^{k} t_{cc}^{k}}\right)^{\frac{1}{2}} - 1 = \left(\frac{X_{II}^{k} X_{cc}^{k}}{X_{Ic}^{k} X_{cI}^{k}}\right)^{\frac{1}{2(\sigma_k - 1)}} - 1,$$

where t_{cI}^{k} is the trade costs factor (one plus tariff equivalent, i.e. $t_{cI}^{k} = p_{cI}^{k}/p_{c}^{k}$) incurred from country *c* to India, and $\sigma_k > 1$ is the industry-specific elasticity of substitution.[51]

To compute τ_{Ic}^{k} we need an estimate of σ_k. Relying on the Armington assumption, Feenstra (1994) proposed an estimation procedure for the elasticity of substitution between domestic and foreign varieties using quantities and prices across exporters to one destination.[52] However, India's limited number of trade partners prior to 2000 in most tradable goods constrains the estimation of σ_k. Therefore, we assume that the elasticity of substitution is equal to 6.0 for all industries based on the findings of recent literature.[53]

Given the elasticity of substitution, the bilateral trade costs as in equation (15) is directly calculated with data on domestic trade ($X_{II}^{k},\ X_{cc}^{k}$) which is the difference between gross output (O_{II}^{k}, O_{cc}^{k}) and total export (X_{I}^{k}, X_{c}^{k}) in industry *k*, and bilateral trade flows ($X_{Ic}^{k},\ X_{cI}^{k}$). Countries' gross value of output data during 1993-2006 come from UNIDO

---

[51] It equals the geometric mean of the barriers in both directions, because gross bilateral trade costs factor (t_{cI}^{k}) may be asymmetric. Intuitionally, it measures bilateral trade costs relative to domestic trade costs (Novy 2008).

[52] For extensions, see Broda and Weinstein (2006), Imbs and Méjean (2009), and Chen and Novy (2009).

[53] Among others, Imbs and Méjean (2009) consider 56 industries, whose elasticities range between 3.1 and 28 with an average of 6.7, and Broda and Weinstein (2006) show that the average elasticity is around 6.6 for SITC-5 sectors in 1990s. Both studies deal with multilateral trade data for U.S. goods.

Industrial Demand-Supply Balance Database (IDSB) 2009 which is based on the 4-digit level of ISIC (Rev. 3) and other trade data for the same period is obtained from the United Nations COMTRADE database. For each industry, we first calculate about 10 bilateral trade costs between India and major trading partners and then obtain the average industry-specific trade costs, weighted by countries' trade volumes with India.[54] Trade with these countries accounts for over 50 percent of total India trade, and the share has been fairly constant over the sample period.

Table 4.1 presents average bilateral trade cost estimates for the nine industries during 1993-2006. For example, the tariff-equivalent trade costs for the food [1] industry decreased from 3.06 to 1.28. For estimation purpose, we define the freeness of trade (FT_{kt}) as the inverse of the weighted average trade costs so that the average annual growth rate of FT_{kt} ranges between 0.78 (transport vehicles and equipment [9] industry) and 4.2 (electrical machinery and electronics [8] industry) as shown in the last row of Table 4.1. The density of paved roads (unit: km/km^2) in each state is used as an indicator of its infrastructure level. As road infrastructure in a state is improved, the competition between firms within the state tends to become tougher than before since some regional markets previously separated within the state will be more integrated into bigger ones. At the same time, relatively high-productivity firms in other states, particularly in neighboring states, could access the concerned state's markets at lower transportation costs than before. Similarly, improved road infrastructure in neighboring states is likely to induce tougher competition between firms in different states. The surrounding states' infrastructure level is

---

[54] The major partners are France, Germany, Indonesia, Italy, Japan, Malaysia, Netherlands, Rep. of Korea, Singapore, USA, and United Kingdom. Other countries including China, Saudi Arabia, and UAE take nontrivial portion of Indian trade volume in our sample years, but data for these countries are unavailable.

important particularly when a state is landlocked and is not adjacent to any countries, because foreign competition can penetrate into the state mainly through the surrounding states' road infrastructure. Naturally, firms in that state can participate in foreign trade via inter-state transport network. Thus, we expect that better infrastructure in surrounding regions as well as in a region make its market more competitive. In this sense, road infrastructure in adjacent regions is considered as an additional regressor in equation (14).

4.5. Estimation and Results

4.5.1. Production Function Estimation and Results

To obtain firm-level raw productivity, we estimate the production function in equation (6) for each manufacturing industry using a three-way fixed effects specification: three-digit NIC sub-industry, three-digit postal code area, and year dummy variables. To avoid potential distortion from extreme outliers, we remove firms whose input-output ratios ($l\,n\,(y/L), l\,n\,(y/K), l\,n\,(y/M),$ and $ln\,(y/E)$) are beyond a three-standard-deviation cut from above and below. As a result, 3,309 outliers are excluded from the original data with 61,805 observations. The estimation results are reported in Table 4.2.

Industry-wise heterogeneity in production technology including agglomeration effects is clearly visible in the results. First, all industries significantly benefit from localization (spatial spillovers) economies. In particular, the food [1] , textiles [2], chemicals [4] are strongly affected by spatial spillovers, while machinery [7] and the wood, paper and printing [3] industries gain the most from urbanization economies. Second, each industry has distinct input intensities in production even though intermediate input accounts for the highest cost share without exception: textiles [2] and transport vehicles and equipment [9] are relatively capital intensive, while food [1] and wood, paper and printing [3] use labor more intensively than do others. Finally, returns to scale, which equals the sum of coefficients on conventional inputs, varies across industries. The null hypothesis of constant returns to scale technology is not rejected at the 5% level only for wood, paper and printing [3] industry. The estimate of the returns to scale in fuels and mineral [5] and food [9] industries are larger than that of other manufacturing industries.

Based on the estimates in Table 4.2, the firm-specific raw TFP measure is calculated as the difference between actual and predicted output as in equation (12). Table 4.3 presents the mean and standard deviation of each industry's raw productivity and agglomeration effects as well as overall TFP for the sample period, 1994-2007. Raw productivity (technical change) accounts for, on average, about 65 percent of the overall productivity. The rest is attributed to the agglomeration component including spatial spillovers. Note that the wood, paper and printing [3] and textiles [2] industry respectively record the highest and lowest average raw TFP. Also, during the sample period, textiles [2] industry has experienced the fastest improvement in raw TFP relative to other industries. Surprisingly, we find negative TFP growth rates in chemicals [4] and machinery [7]. These arise from a sharp drop in productivity during one year of the sample (1995), which completely offsets positive productivity growth in most other years.

Table 4.4 shows the substantial spatial variation in raw productivity across industries. For each industry, bold figures indicate that corresponding states rank top-three in average productivity among 18 states. During the sample period Karnataka, Goa, and Kerala record the highest regional productivity for five industries, and Andhra Pradesh and Tamil Nadu do so for three industries. Appendix Figure 4.A.2 visually presents the spatial distribution of raw productivity and agglomeration effects for a low-tech (textiles) and a hi-tech (electrical machinery and electronics) industry.

4.5.2. Competition and Regional Productivity Variation

With nine industries and 14 years of data, we obtain 126 annual industry-specific raw productivity distributions. Grouping firms by regions, i.e. 18 states, we have 2,268 regional productivity distributions distinguished in three dimensions; region, industry, and year. With these distributions, the censored equation in (14) is rewritten as follows:

(16) $$\Omega_{krt}^{p} = \begin{cases} \Omega_{krt}^{p*} & \text{if } \Omega_{krt}^{p*} > 0 \\ L^{p} & \text{otherwise} \end{cases},$$

$$\Omega_{krt}^{p*} = \delta_0^p + \delta_1^p\, FT_{kt-1} + \delta_2^p\, IN_{rt-1} + \delta_3^p SIN_{rt-1} +$$

$$\delta_4^p FT_{kt-1} \times IN_{rt-1} + \delta_5^p FT_{kt-1} \times SIN_{rt-1} + \delta_6^p IN_{rt-1} \times SIN_{rt-1} +$$

$$\delta_7^p\, HHI_{kt-1} + \delta_8^p PS_{rt-1} + \delta_9^p N_{krt}^p + \delta_{10}^p PL_{krt}^p + \delta_{11}^p PR_{krt}^p + d_r^p + d_k^p + \nu_{krt},$$

$$p = 10\%, 20\%, 50\%, 80\%, 90\%,$$

The dependent variable in equation (16), Ω_{krt}^{p*} is representative productivity of firms near the *p*-th percentile of a specific combination of region (*r*), industry (*k*), and year (*t*), instead of a single firm's productivity corresponding to the *p*-th percentile. For example, Ω_{krt}^{10} (Ω_{krt}^{90}) is average productivity of firms whose productivity indices belong in the interval between zero and 10th percentiles (90th and 100th percentiles) of the *krt*-th distribution. The above definition is consistent with theory (equation 5) and also useful in exploring the resource reallocation following changes in trade costs.

For ease of interpretation, freeness of trade (*FT*), the inverse of trade costs, is used in equation (16). Moreover, a variable capturing surrounding regions' infrastructure (*SIN*) is included in the right hand side of equation (16). In order to capture a potential synergy or trade-off relationship between domestic infrastructure and trade liberalization in enhancing raw productivity level, three interaction terms among own local infrastructure, surrounding regions' infrastructure and freeness of trade are added in equation (16). Additional explanatory variables include: the Herfindal-Hirschman Index (HHI_{kt}) based on output for each industry to reflect the industry's degree of competition;[55] and share of each state in India's population (PS_{rt}) represents regional market size. Firms grouped within a percentile interval (e.g. 10th to 20th) may have certain features explaining their productivity level, e.g. the number of firms (*N*), the average number of plants each operates (PL), and the average number of products each produces (PR). We also control for these firm-level characteristics in equation (16). Finally, two regional dummies each representing coastal states and north-eastern states, and dummies for nine industries are included in the estimation to capture unobserved heterogeneity across regions and industries. All regressors are lagged by one year to avoid possible bias due to potential endogeneity.

For median and each representative percentile value, equation (16) is first estimated by a classical Tobit model under the assumption of normal and homoskedastic disturbance term. Estimated parameters are presented in Table 4.5. Marginal effects are reported in Table 4.6 only for variables of interest, i.e. competiton, in order to provide economic interpretation of our findings. Since we are particularly interested in the effect of competition pressures on

---

[55] $HHI_{krt} = \sum_i S_{krti}^2$, where$S_{krti}^2$ is the squared output share of firm *i* from industry *k*, in state *r*, and at year *t*. If HHI_{krt} is close to one, it means that the corresponding market is near monopoly, and if it is close to zero, the market is under perfect competition.

productivity variation in the entire geographic system of a country, we focus on the marginal effects for the unconditional expected value of the *p*-th percentile productivity, $\mathrm{E}(\Omega_{krt}^{p}|X_{krt})$, instead of the conditional marginal effects, $\mathrm{E}(\Omega_{krt}^{p}|X_{krt}, \Omega_{krt}^{p} > 0)$. Following Ai and Norton (2003), the overall marginal effects of interacted variables $(x_1,\ x_2)$ and interaction term $(x_1 x_2)$ at the means of independent variables are obtained from:

(17)
$$\frac{\partial E(\Omega|X)}{\partial x_1} = \delta_1 + \delta_{12}\bar{x}_2\Phi\left(\frac{X\delta}{\sigma}\right),$$

$$\frac{\partial^2 E(\Omega|X)}{\partial x_1 \partial x_2} = \delta_{12}\Phi\left(\frac{X\delta}{\sigma}\right) + \frac{1}{\sigma}(\delta_1 + \delta_{12}\bar{x}_2)(\delta_2 + \delta_{12}\bar{x}_1)\Phi'\left(\frac{X\delta}{\sigma}\right),$$

where Φ is the standard normal cumulative distribution, σ is the standard error and $X\delta$ is the predicted value using the mean of regressors.

As shown in Table 4.5, the estimates of the coefficient on FT and SIN are significantly negative while those on IN are significantly positive in all percentile subsamples. However, the interaction between freeness of trade and infrastructure of surrounding states is significantly positive supporting the expected synergy between domestic infrastructure and trade liberalization in enhancing raw productivity level. On the contrary, the interaction term between FT and IN appears to be insignificant and that between IN and SIN is significantly negative. All coefficients on other control variables have the expected signs. Even though HHI is significant only in the 10th percentile subsample, its negative sign is consistent across subsamples. It means that higher competition within an industry (lower value of HHI) leads to productivity growth of, particularly, least productive firms (Syverson 2004). The coefficient on population share is positive and significant suggesting self-selection of heterogeneous firms to locate in specific markets (Saito and Gopinath 2009). The locational advantage of firms in coastal states and the disadvantage in remote and less developed north-eastern states are confirmed by respective coefficients. The better performance of multi-plant and/or multi-productive firms is also supported by the parameter estimates.

Our key hypotheses are clearly validated by the marginal effects and their statistical significance presented in Table 4.6. The marginal effects calculated following (17) are positive for SIN in all subsamples and for FT in 10th and 90th percentile subsamples, unlike the coefficient estimates in Table 4.5. The elasticities based on the marginal effects and evaluated at the means of data are presented in the last three rows of Table 4.6. Free trade and own infrastructure appear to shift the tails of the productivity distribution only, leaving the median unchanged (Saito and Gopinath 2009). The estimated elasticity of surrounding states' infrastructure is significant in all subsamples and increases in magnitude for the right-tail of the productivity distribution. These results indicate that lowering trade costs including trade liberalization and improved domestic infrastructure has not only forced the least productivity firms to exit, but also has shifted the productivity distribution to the right and this effect might be stronger for high productivity firms in a region. For instance, foreign competition (domestic competition) may force domestic (local) firms to raise firm size and scale efficiency, and also provide incentives for innovation to compete with foreign (other region's) competitors. In the context of trade, learning-by-exporting and easier access to foreign inputs and technology may also increase firm-level productivity. Local infrastructure is also regarded as accumulated physical capital for firms beyond the industrial capital stock directly used in production, bringing about increasing returns in transportation and thus, yielding larger input and output markets. The significant and positive interaction effect between FT and SIN implies that the marginal effect of freeness of trade on regional productivity is increased when infrastructure in surrounding regions is well developed and also that the marginal effect of neighboring regions' infrastructure on regional productivity increases when trade is more liberalized.

The maximum likelihood estimates of the Tobit model are not consistent when disturbances do not follow a homoskedastic and normal distribution.[56] To correct this problem, we employ Powell's (1984) censored least absolute deviation (CLAD) estimator. The CLAD estimator requires the median of disturbances to be zero, and thus, it is consistent when disturbances do not follow a homoskedastic normal distribution. However, obtained parameters cannot be directly converted into marginal effects as in Tobit models. In the CLAD model, a parameter estimate measures median response of the p^{th} percentile of productivity distribution to one unit change in the explanatory variable, provided the distribution has positive p^{th} percentile values. The CLAD estimates are presented in Table 4.7 along with the percentage change in the conditional median of productivity when an independent variable increases by one percent. The latter measures, referred to as elasticities, are computed at the mean and found to be similar to those of the Tobit model with single key exception: the elasticities of three competition sources are lower in the 90^{th} percentile subsample compared with the Tobit results. That is, trade liberalization and improvement of domestic infrastructure induces a greater shift of the left-tail of the raw productivity distribution than the median or right tail.

---

[56] Pegan and Vella's (1989) conditional moment tests are carried out to check the validity of assumption of homoskedastic and normal distribution. The null is rejected at 1% level for all subsamples but 20^{th} percentile one as shown at the bottom of Table 5. However, the conditional moment test rejection does not necessarily imply serious misspecification of the Tobit estimator. In the sense Hausman (1979) suggested, the choice of model between Tobit and CLAD depends on whether non-normality and heteroskedasticity are serious enough to distort the Tobit estimator from the CLAD estimator (Wilhelm 2008).

Table 4.8 provides a better understanding of how much the competition pressures increase the aggregate productivity. During 1994 to 2007, trade costs changed by about 37.5 percent, while domestic road infrastructure improved by as much as 28 percent. The cumulative effects of variables on raw productivity of least productive firms (10th percentile sample) are obtained by multiplying their growth rate by their elasticities derived from the CLAD estimates. Compared with the aggregate growth rate of raw TFP for all manufacturing firms, the magnitude of each cumulative effect is not trivial at all. Trade liberalization has raised the aggregate productivity of least productive firms by 1.69 percent, accounting for a large share of the 3.85 percent growth in the aggregate productivity of all manufacturing firms in India during the sample period. The two infrastructure variables together increase the average productivity of firms belonging to the 10th percentile group by as much as trade-cost reduction does.

The finding that a change in the international competition level brings about a higher change in regional productivity relative to own infrastructure (or surrounding states' infrastructure) has policy implications. While it is tempting to support free trade, we need information on the cost of each strategy (trade versus infrastructure) to identify efficient strategies for regional development in India. The higher elasticity of surrounding infrastructure than that of own infrastructure indicates that constructing more efficient domestic transportation network would be more helpful for regional productivity growth in less-developed or remote regions than focusing on development of some targeted regions.

Our empirical findings supporting competition effects of trade liberalization and improved road infrastructure on productivity should be considered in a broader context. The Prowess database likely contains relatively large firms and it is likely that firms in the organized sector are more productive than those in the unorganized sector. Then, the increase in the real cutoff level of productivity may not be fully captured in our empirical application. Second, as discussed by Topalova (2007) and others, the competition effects of trade liberalization and infrastructure improvement might not be realized unless complementary domestic policies (free entry and exit) are in place. Finally, within our sample period, India has experienced a severe recession during 2001-2004, which introduces additional complexity in measuring the relationship between competition and firm-level productivity.

4.6. Summary and Conclusions

The objective of this study is to identify regional productivity arising from pure technical change (raw productivity) and agglomeration effects in Indian manufacturing industries. Furthermore, we examine the effects of falling trade costs, i.e. policy and geographic barriers, and improving domestic infrastructure on the regional variation in raw productivity. For this purpose, we augment a firm-level production function framework to include external economies: urbanization and localization economies. Spatial econometric techniques are employed to derive estimates of firm-level productivity by industry, region and time, while correcting for the commonly observed simultaneity between conventional inputs and productivity. Then, we derive the mean, median, and alternative percentiles of the firm-level productivity distribution by region and industry for each time period. Using quantile regression techniques, we estimate the relationship between measures of firm-level (raw) productivity distribution and international and domestic competition indicators.

We employ firm-level data with location information from the Prowess database of the Centre for Monitoring Indian Economy during 1994-2007. Variables required for the estimation of the firm-level production function are derived in constant Rupees: output, capital, labor and energy. In addition, urbanization and localization economies are represented by the sum of output of all industries in a three-digit postal code area and the spatial lag of the dependent variable in the neighborhood (50 km), respectively. For the quantile regression, international competition is represented by industry-specific trade cost, which includes all trade and geographic barriers. Observed trade and output data are used to derive industry-specific trade costs following a gravity model. Domestic competition is represented by the level of regional road infrastructure.

Production function estimates reveal differences in conventional input intensities and returns to scale among Indian manufacturing industries. On average, raw productivity accounts for about 65 percent of the overall productivity, while the rest is attributed to agglomeration effects. Substantial industrial and regional variation in raw productivity is observed during our sample years, and international and domestic competition-induced productivity growth is also observed across industries and regions. Quantile regression of measures of firm-level productivity distribution by region and industry shows that falling trade costs boost average productivity. Improvements in domestic infrastructure also bring about productivity growth, but the effect on the distribution tails is different from that observed for trade costs. The effect of domestic infrastructure on the right tail of the distribution is larger than that on the left tail. There is an obvious synergy relationship between improved infrastructure in surrounding regions and trade liberalization in enhancing raw productivity level. In the context of overall Indian manufacturing, a change in the level of infrastructure appears to bring about a higher change in regional productivity relative to a change in the international competition level. However, information on the cost of each of these options is required to identify effective and efficient strategies for regional development in India.

References

Acemoglu, D., and M. Dell. 2010. "Productivity Differences between and Within Countries." *American Economic Journal: Macroeconomics* 2 (1): 169-188.

Ai, C., and E.C. Norton. 2003. "Interaction Terms in Logit and Probit Models." *Economic Letters* 80 (1): 123-129.

Alfaro, L., and A. Chari. 2009. "India Transformed? Insights 1988-2005." Working Paper, UNC-Chapel Hill and Harvard Business School.

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

Anderson, J., and E. van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93 (1): 170-192.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht.

Anselin, L., R.J. Florax, and S.J. Rey. 2004. "Econometrics for Spatial Models: Recent Advances." In: Anselin, L., Florax, R., Rey, S.J. (Eds.) *Advances in Spatial Econometrics. Methodology, Tools and Applications*. Springer-Verlag, Berlin, pp. 1-25.

Baldwin, R.E., and T. Okubo. 2006. "Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting." *Journal of Economic Geography* 6 (3): 323-346.

Baldwin, R.E., R. Forslid, P. Martin, G.I.P. Ottaviano, and F. Robert–Nicoud. 2003. *Economic Geography and Public Policy*. Princeton University Press, Princeton.

Behrens, K. 2006. "Do Changes in Transport Costs and Tariffs Shape the Space-Economy in the Same Way?" *Papers in Regional Science* 85 (3): 379-399.

Bernard, A. B., J.B. Jensen, and P.K. Schott. 2003. "Plants and Productivity in International Trade." *American Economic Review* 93 (4): 1268-1290.

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

Bertrand, M., P. Mehta, and S. Mullainathan, 2002. "Ferreting Out Tunneling: An Application to Indian Business Groups." *Quarterly Journal of Economics* 117 (1): 121-148.

Broda, C., and D. Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121 (2): 541-585.

Chari, A., and N. Gupta. 2008. "Incumbents and Protectionism: The Political Economy of Foreign Entry Liberalization." *Journal of Financial Economics* 88 (3): 633-656.

Chen, N., and D. Novy. 2009. "International Trade Integration: A Disaggregated Approach." CESIFO WORKING PAPER NO. 2595, University of Warwick, UK.

Ciccone, A. 2002. "Agglomeration Effects in Europe." *European Economic Review* 46 (2): 213-227.

Combes, P-P., T. Mayer, and J-F. Thisse. 2008. *Economic Geography: The Integration of Regions and Nations*. Princeton University Press, Princeton.

Dinc, I. S., and N. Gupta, 2007. "The Decision to Privatize: Finance, Politics and Patronage." Mimeo, MIT, Cambridge, MA.

Feenstra R.C. 1994. "New Product Varieties and the Measurement of International Prices." *American Economic Review* 84 (1): 157-177.

Fisman, R., and T. Khanna, 2004. "Facilitating Development: The Role of Business Groups". *World Development* 32 (4): 609-628.

Fujita, M., P.R. Krugman, and A.J. Venables. 1999. *The Spatial Economy: Cities, Region and International Trade*. MIT Press, Cambridge, MA.

Glaeser, E.L., and J.D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4): 983-1028.

Goldberg, P., A. Khandelwal, N. Pavcnik, and P. Topalova. 2008. "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India." NBER Working Paper 14416, Cambridge, MA.

Goldberg, P., A. Khandelwal, N. Pavcnik, and P. Topalova. 2009. "Trade Liberalization and New Imported Inputs." *American Economic Review* 99 (2): 494-500.

Greenaway, D., W. Morgan, and P. Wright. 2002. "Trade Liberalization and Growth in Developing Countries." *Journal of Development Economics* 67 (1): 229-244.

Hausman, J.A. 1978. "Specification Tests in Econometrics." *Econometrica* 46 (6): 1251-1271.

Henderson, J.V. 1988. *Urban Development: Theory, Fact and Illusion*. Oxford University Press, New York.

Henderson, J.V. 2003. "Marshall's Scale Economies." *Journal of Urban Economics* 53 (1): 1-28.

Henderson, J.V., T. Lee, and Y.J. Lee. 2001. "Scale Externalities in Korea." *Journal of Urban Economics* 49 (3): 479-504.

Imbs J., and I. Méjean. 2009. "Elasticity of Optimism." CEPR Discussion Paper 7177.

Kelejian, H.H., and I.R. Prucha. 1999. "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model." *International Economic Review* 40 (2): 509-533.

Krishna, P., and D. Mitra. 1998. "Trade Liberalization, Market Discipline and Productivity Growth: New Evidence from India." *Journal of Development Economics* 56 (2): 447-62.

Krugman, P.R. 1991. "Increasing Returns and Economic Geography." *Journal of Political Economics* 99 (3): 483-499.

Lall, S.V., Z. Shalizi, and U. Deichmann. 2004. "Agglomeration Economies and Productivity in Indian Industry." *Journal of Development Economics* 73 (2): 643-674.

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

Melitz, M. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695-1725.

Melitz, M., and G.I.P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies* 75 (1): 295-316.

Mitra, D., and B. Ural. 2008. "Indian Manufacturing: A Slow Sector in A Rapidly Growing Economy." *Journal of International Trade and Economic Development* 17 (4): 525-559.

Novy, D. 2008. "Gravity Redux: Measuring International Trade Costs with Panel Data." Working Paper, University of Warwick, UK.

Oosterhaven, J., and L. Broersma. 2007. "Sector Structure and Cluster Economies: A Decomposition of Regional Labour Productivity." *Regional Studies* 41 (5): 639-659.

Pagan, A., and F. Vella. 1989. "Diagnostic Tests for Models Based on Individual Data: A Survey." *Journal of Applied Econometrics* 4: S29-S59.

Rappaport, J. 2009. "The Increasing Importance of Quality of Life." *Journal of Economic Geography* 9 (6): 779-804.

Rice, P., A.J. Venables, and E. Patacchini. 2006. "Spatial Determinants of Productivity: Analysis for the Regions of Great Britain." *Regional Science and Urban Economics* 36 (6): 727-752.

Rosenthal, S.S., and W.C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In: Henderson, V., Thisse, J.F. (Eds.) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland, Amsterdam, pp. 2118-2171.

Saito, H., and M. Gopinath. 2009. "Plants' Self-Selection, Agglomeration Economies and Regional Productivity in Chile." *Journal of Economic Geography* 9 (4): 539-558.

Schor, A. 2004. "Heterogeneous Productivity Responses to Tariff Reduction: Evidence from Brazilian Manufacturing Firms." *Journal of Development Economics* 75 (2): 373-396.

Sivadasan, J. 2009. "Barriers to Competition and Productivity: Evidence from India." *B.E. Journal of Economic Analysis and Policy* 9 (1) (Advances).

Syverson, C. 2004. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics* 86 (2): 534-550.

Topalova, P. 2007. "Trade Liberalization and Firm Productivity: The Case of India." IMF Working Paper, WP/04/28, Washington DC.

Tybout, J. 2000. "Manufacturing Firms in Developing Countries: How Well They Do and Why?" *Journal of Economic Literature* 38 (1): 11-44.

Wilhelm, M.O. 2008. "Practical Considerations for Choosing Between Tobit and SCLS or CLAD Estimators for Censored Regression Models with an Application to Charitable Giving." *Oxford Bulletin of Economics and Statistics* 70 (4): 559-582.

Yasar, M., and C. J. Morrison. 2007. "International Linkages and Productivity at the Plant Level: Foreign Direct Investment, Exports, Imports and Licensing." *Journal of International Economics* 71 (2): 373-388.

Table 4.1. Industry-Specific Trade Costs

	Industry								
Year	1	2	3	4	5	6	7	8	9
1993	3.057	1.730	2.946	2.176	1.953	2.474	2.258	2.401	2.370
1994	2.764	1.585	2.856	2.028	1.918	2.351	2.060	2.366	2.536
1995	2.572	1.621	2.682	2.050	1.918	2.345	2.033	2.128	2.883
1996	2.404	1.595	2.609	2.019	1.920	2.248	1.977	2.061	2.654
1997	2.594	1.664	2.498	1.978	1.860	2.304	1.829	2.045	2.835
1998	2.270	1.676	2.608	2.002	1.719	2.290	1.835	2.191	2.800
1999	2.178	1.673	2.673	1.993	1.626	2.252	1.958	2.197	3.028
2000	2.351	1.575	2.593	2.013	1.758	2.159	1.871	2.035	3.010
2001	2.376	1.494	2.547	1.926	1.867	2.140	1.793	1.982	2.586
2002	2.127	1.454	2.395	1.881	1.784	2.192	1.723	1.883	2.707
2003	2.031	1.376	2.308	1.837	1.635	2.203	1.631	1.843	2.404
2004	2.088	1.475	2.274	1.793	1.717	2.113	1.593	1.839	2.353
2005	2.159	1.243	2.156	1.606	1.539	2.098	1.457	1.483	2.272
2006	2.183	1.240	2.221	1.549	1.380	2.015	1.359	1.407	2.142
AG(TC)	**-2.56**	**-2.53**	**-2.15**	**-2.58**	**-2.63**	**-1.57**	**-3.83**	**-4.03**	**-0.78**
AG(FT)	**2.62**	**2.59**	**2.20**	**2.65**	**2.71**	**1.59**	**3.98**	**4.20**	**0.78**

Notes: AG (TC) and AG (FT) are average annual growth rate of trade costs and freeness of trade, respectively. The elasticity of substitution is assumed to be equal to 6 for all industries.

Industry Definitions: 1 Food; 2 Textiles & Apparel; 3 Wood, Paper & Printing; 4 Chemicals & Rubber; 5 Fuels & Mineral; 6 Metals; 7 Machinery; 8 Electricals & Electronics; 9 Transport Vehicles & Equipment.

Table 4.2. Firm-Level Production Function Estimation Results

(Dependent Variable: Log of Output, *ln*(*y*))

Industry	*Wln(y)*	*ue*	*ln(k)*	*ln(m)*	*ln(l)*	*ln(e)*	RTS	Obs.
1	0.077	0.0003[(a)]	0.157	0.687	0.195	0.065	1.104	8,180
	(0.020)	(0.003)	(0.041)	(0.008)	(0.005)	(0.004)	(0.013)	
2	0.060	0.002	0.138	0.717	0.122	0.068	1.045	8,559
	(0.008)	(0.001)	(0.009)	(0.010)	(0.004)	(0.004)	(0.009)	
3	0.037	0.003	0.061	0.728	0.169	0.073	1.031[(c)]	2,997
	(0.013)	(0.001)	(0.015)	(0.026)	(0.009)	(0.006)	(0.022)	
4	0.056	0.001[(a)]	0.129	0.703	0.168	0.077	1.077	16,672
	(0.018)	(0.001)	(0.020)	(0.009)	(0.003)	(0.003)	(0.010)	
5	0.034	0.002[(a)]	0.122	0.711	0.151	0.120	1.105	3,165
	(0.010)	(0.003)	(0.026)	(0.019)	(0.009)	(0.008)	(0.019)	
6	0.045	0.002[(b)]	0.082	0.759	0.124	0.086	1.051	7,799
	(0.008)	(0.001)	(0.008)	(0.007)	(0.007)	(0.005)	(0.009)	
7	0.021[(b)]	0.004	0.103	0.725	0.173	0.050	1.051	4,320
	(0.009)	(0.001)	(0.012)	(0.009)	(0.015)	(0.006)	(0.019)	
8	0.036	0.001[(a)]	0.100	0.779	0.127	0.055	1.062	6,403
	(0.020)	(0.001)	(0.021)	(0.016)	(0.009)	(0.006)	(0.016)	
9	0.051	0.0004[(a)]	0.153	0.699	0.152	0.074	1.078	3,858
	(0.010)	(0.001)	(0.027)	(0.014)	(0.008)	(0.005)	(0.015)	

Note: Value in parenthesis is the bootstrapped standard error based on 200 iterations. All estimates are statistically significant at 1% level except (a) and (b) indicating statistically insignificance at 10% level, and significance at 5% level, respectively.
In the RTS (returns to scale) column, (c) implies the corresponding industry exhibits constant returns to scale, statistically at 5% level.

Industry Definitions: 1 Food; 2 Textiles & Apparel; 3 Wood, Paper & Printing; 4 Chemicals & Rubber; 5 Fuels & Mineral; 6 Metals; 7 Machinery; 8 Electricals & Electronics; 9 Transport Vehicles & Equipment

Table 4.3. Estimated Raw TFP and Agglomeration Effects (1994-2007 average)

	Raw TFP		AE		Overall TFP		Annual Growth Rate		
Industry	Mean	S.D.	Mean	S.D.	Mean	S.D.	RTFP	AE	OTFP
1	2.198	1.069	1.270	0.122	3.469	1.041	0.58	0.10	0.40
2	1.980	0.863	1.214	0.075	3.194	0.852	1.07	0.02	0.68
3	2.415	0.766	1.149	0.102	3.564	0.748	0.10	0.26	0.15
4	2.127	0.866	1.212	0.075	3.338	0.852	-0.08	0.04	-0.04
5	2.257	0.925	1.129	0.081	3.386	0.918	0.36	0.23	0.32
6	2.120	0.763	1.181	0.070	3.301	0.751	0.47	0.13	0.35
7	2.279	0.712	1.124	0.090	3.403	0.698	-0.18	0.23	-0.05
8	2.136	0.836	1.136	0.055	3.273	0.828	0.09	0.16	0.12
9	2.186	0.777	1.213	0.073	3.399	0.769	0.22	0.24	0.23
Total	2.151	0.867	1.194	0.094	3.345	0.852	0.29	0.11	0.23

Note: Overall TFP is the sum of Raw TFP and Agglomeration Effects (AE), which is the sum of the estimated values of spatial dependence and urbanization economies.

Industry Definitions: 1 Food; 2 Textiles & Apparel; 3 Wood, Paper & Printing; 4 Chemicals & Rubber; 5 Fuels & Mineral; 6 Metals; 7 Machinery; 8 Electricals & Electronics; 9 Transport Vehicles & Equipment

Table 4.4. Estimated Raw TFP (1994-2007 State-Industry Average)

	Industry								
State	1	2	3	4	5	6	7	8	9
1	2.29	1.74		2.15	2.16			2.22	
	(0.740)	(0.707)		(0.888)	(0.205)			(0.876)	
2	1.89	1.81	**2.68**	2.02	2.36	2.35	2.43	2.05	1.99
	(0.601)	(0.446)	(0.515)	(0.494)	(0.296)	(0.931)	(0.505)	(0.984)	(0.445)
3	1.78	2.10	2.34	2.28	**2.75**	2.18	2.23	2.12	2.04
	(0.496)	(0.675)	(0.574)	(0.899)	(0.896)	(0.597)	(0.467)	(0.676)	(0.388)
4	2.00	1.95	2.27	2.08	2.22	2.10	2.12	2.03	2.15
	(0.709)	(0.677)	(0.524)	(0.873)	(0.743)	(0.525)	(0.647)	(0.523)	(0.644)
5	2.20	**2.15**		1.98	2.18	2.20	2.35	2.06	2.18
	(0.759)	(1.204)		(0.648)	(0.563)	(0.796)	(0.678)	(0.555)	(0.512)
6	1.80	2.04	2.52	2.05	1.96	2.15	2.43	2.13	2.00
	(0.766)	(1.104)	(0.490)	(0.658)	(0.557)	(0.554)	(1.092)	(0.952)	(0.472)
7				**2.45**	2.47	2.28	2.21		1.92
				(0.629)	(0.869)	(0.646)	(0.393)		(0.423)
8	2.48		2.08	2.25	2.16	2.30			
	(1.107)		(0.560)	(0.811)	(0.470)	(0.533)			
9	2.33	1.74	2.10	1.93	2.03	1.95	2.26	1.90	2.09
	(1.004)	(0.678)	(0.565)	(0.750)	(0.809)	(0.860)	(1.024)	(0.894)	(1.081)
10	2.53		2.49	2.18	1.94	2.24	2.26	**2.44**	
	(1.699)		(0.439)	(0.649)	(0.445)	(0.718)	(0.301)	(1.019)	
11	2.14	1.74	2.87	2.31	2.36	2.32	**2.47**	2.21	2.06
	(0.756)	(0.363)	(1.493)	(0.795)	(0.914)	(0.726)	(0.574)	(0.744)	(0.314)
12	2.28	2.00	2.51	2.18	2.46	2.18	2.40	2.18	2.36
	(1.173)	(0.747)	(0.481)	(0.769)	(0.752)	(0.634)	(0.675)	(0.738)	(0.569)
13	2.01	1.93	2.23	1.99	2.13	1.92	2.10	2.06	2.10
	(0.927)	(0.817)	(0.640)	(0.761)	(1.115)	(0.558)	(0.623)	(0.839)	(0.781)
14	2.11	1.96	**2.59**	2.20	**2.54**	2.17	2.42	2.29	**2.83**
	(0.989)	(0.935)	(0.667)	(0.793)	(0.798)	(0.518)	(0.656)	(0.863)	(0.766)
15	**2.70**	**2.21**	2.46	**2.42**	**2.54**	**2.51**	2.35	2.10	2.22
	(1.677)	(1.061)	(1.471)	(1.282)	(0.975)	(2.025)	(0.591)	(0.807)	(0.618)
16	**2.57**	**2.99**		2.19	2.13	**2.38**	**2.53**	2.03	**2.32**
	(1.022)	(1.219)		(0.528)	(0.339)	(0.676)	(0.440)	(0.394)	(0.969)
17	**2.63**	2.10	2.46	**2.94**	**2.74**	**2.36**	2.08	**2.36**	**2.64**
	(1.437)	(1.154)	(0.734)	(1.725)	(1.954)	(0.678)	(0.413)	(0.840)	(1.176)
18	2.33	2.07	**2.87**	2.33	2.13	2.26	**2.56**	**2.43**	2.29
	(1.271)	(0.934)	(1.207)	(1.022)	(0.702)	(0.779)	(0.721)	(1.068)	(0.982)

Notes: Numbers in parentheses are standard deviation. Bold figures indicate that corresponding states recode top-three average productivity among states for each industry.

Industry: 1 Food; 2 Textiles & Apparel; 3 Wood, Paper & Printing; 4 Chemicals & Rubber; 5 Fuels & Mineral; 6 Metals; 7 Machinery; 8 Electricals & Electronics; 9 Transport Vehicles & Equipment
State: 1 Himachal Pradesh; 2 Punjab; 3 Haryana; 4 Delhi; 5 Rajasthan; 6 Uttar Pradesh; 7 Bihar; 8 Assam; 9 West Bengal; 10 Orissa; 11 Madhya Pradesh; 12 Gujarat; 13 Maharashtra; 14 Andhra Pradesh; 15 Karnataka; 16 Goa; 17 Kerala; 18 Tamil Nadu

Table 4.5. Estimation Results of Tobit, Coefficients

(Dependent Variable: Average Productivity in Each Percentile Group)

	Percentiles				
	10th	20th	50th	80th	90th
FT	-0.658 **	-1.293 ***	-1.126 ***	-1.362 **	-1.950 *
	(0.282)	(0.362)	(0.415)	(0.544)	(1.013)
IN	0.074 ***	0.072 ***	0.007	0.086 ***	0.197 ***
	(0.024)	(0.022)	(0.025)	(0.033)	(0.061)
SIN	-0.520	-0.692 **	-0.634 *	-0.623	-1.777 **
	(0.343)	(0.322)	(0.369)	(0.482)	(0.901)
FT×IN	0.005	-0.035	-0.013	0.014	-0.048
	(0.035)	(0.033)	(0.037)	(0.049)	(0.092)
FT×SIN	1.674 ***	2.235 ***	1.938 ***	2.775 ***	6.356 ***
	(0.648)	(0.609)	(0.698)	(0.912)	(1.701)
IN×SIN	-0.095 ***	-0.067 **	0.019	-0.103 **	-0.202 **
	(0.036)	(0.034)	(0.038)	(0.050)	(0.095)
HHI	-2.675 **	-1.763	-0.390	-2.572	-5.840 *
	(1.314)	(1.230)	(1.430)	(1.842)	(3.461)
PS	0.706 **	0.804 ***	0.512	1.678 ***	3.275 ***
	(0.311)	(0.294)	(0.338)	(0.441)	(0.821)
Coast	0.107 ***	0.080 ***	0.140 ***	0.274 ***	0.664 ***
	(0.028)	(0.027)	(0.031)	(0.040)	(0.075)
NE	-0.407 ***	-0.429 ***	-0.412 ***	-0.405 ***	-0.622 ***
	(0.067)	(0.067)	(0.078)	(0.102)	(0.185)
N	0.0009 ***	0.0012 ***	0.0020 ***	0.0017 ***	0.0039 ***
	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0008)
Plant	0.048 ***	0.069 ***	0.117 ***	0.207 ***	0.225 ***
	(0.006)	(0.007)	(0.010)	(0.015)	(0.028)
Product	0.101 ***	0.078 ***	0.153 ***	0.197 ***	0.298 ***
	(0.008)	(0.006)	(0.009)	(0.013)	(0.026)
Sigma	0.527	0.490	0.550	0.726	1.365
	(0.010)	(0.010)	(0.011)	(0.014)	(0.025)
Pseudo R^2	0.221	0.250	0.298	0.249	0.136
Log-Likelihood	-1758	-1609	-1695	-2144	-3177
Obs.	2268	2268	2268	2268	2268
Censored Obs.	652	727	826	727	652
L.B.	0.518	0.874	1.11	1.43	1.93
F(6,2247)	8.91	11.98	8.55	18.65	19.43
F(8,2247)	16.47	14.57	9.24	14.65	11.47
CM-normal	28.628 (13.8)	<u>14.08</u> (12.1)	31.277 (14.0)	32.029 (15.2)	84.489 (13.5)
CM-Hetero	150.32	181.07	168.61	201.35	220.22

Notes: Numbers in parentheses are standard deviation. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively. Coast and NW are dummies indicating coastal states and north-western states, respectively. Sigma is the standard error. L.B. indicates the lower bound productivity which is the left-censoring limit. Model also includes industry dummies and constant term; effects not shown
F(6,2247) and F(8,2247) are F-test statistics for joint significance of interacted variables and industry dummies, respectively. CM-normal and CM-Hetero are conditional moment test statistics. The critical values (1%) are provided in parentheses.

Table 4.6. Estimation Results of Tobit, Marginal Effects and Elasticities

(Dependent Variable: Average Productivity in Each Percentile Group)

	Percentiles				
	10th	20th	50th	80th	90th
$\partial E(\Omega\|X)/\partial FT$	0.065 *	-0.187 *	-0.163	-0.061	0.480 *
	(0.054)	(0.113)	(0.140)	(0.181)	(0.350)
$\partial E(\Omega\|X)/\partial IN$	0.023 ***	0.013 ***	0.005	0.027 ***	0.049 ***
	(0.006)	(0.004)	(0.005)	(0.007)	(0.013)
$\partial E(\Omega\|X)/\partial SIN$	0.123 **	0.176 ***	0.210 ***	0.353 ***	0.657 ***
	(0.057)	(0.041)	(0.051)	(0.066)	(0.128)
$\partial^2 E(\Omega\|X)/\partial FT\partial IN$	0.005	-0.022	-0.009	0.006	-0.019
	(0.025)	(0.017)	(0.021)	(0.028)	(0.056)
$\partial^2 E(\Omega\|X)/\partial FT\partial SIN$	1.168 ***	1.144 ***	1.068 ***	1.572 ***	3.901 ***
	(0.445)	(0.323)	(0.399)	(0.516)	(0.995)
$\partial^2 E(\Omega\|X)/\partial IN\partial SIN$	-0.063 **	-0.032 *	0.012	-0.050 *	-0.107 **
	(0.025)	(0.017)	(0.021)	(0.027)	(0.054)
Elasticity of	0.040 *	-0.086 *	-0.051	-0.009	0.116 *
FT	(0.033)	(0.063)	(0.082)	(0.142)	(0.084)
Elasticity of	0.024 ***	0.015 ***	0.009	0.021 **	0.028 ***
IN	(0.006)	(0.004)	(0.011)	(0.009)	(0.008)
Elasticity of	0.029 **	0.070 ***	0.073 ***	0.081 ***	0.116 ***
SIN	(0.014)	(0.023)	(0.022)	(0.023)	(0.034)

Notes: Numbers in parentheses are standard deviation. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively.
The marginal effects and elasticities are calculated using the nlcom command of STATA version 10.1. To save space we only report the overall marginal effects of our main competition pressures at the mean here.

Table 4.7. Estimation Results of CLAD, Coefficients and Elasticities

(Dependent Variable: Average Productivity in Each Percentile Group)

	Percentiles				
	10th	20th	50th	80th	90th
FT	-0.883 **	-1.247 ***	-0.478 *	-1.486 ***	-0.140
	(0.291)	(0.293)	(0.278)	(0.406)	(0.704)
IN	0.049 **	0.050 **	-0.017	0.034 *	0.091 *
	(0.019)	(0.022)	(0.022)	(0.029)	(0.049)
SIN	-0.949 **	-0.816 **	-0.500 **	-1.236 **	-0.365
	(0.346)	(0.348)	(0.255)	(0.422)	(0.726)
FT×*IN	-0.043 *	-0.051 **	-0.006 **	-0.040 *	0.001
	(0.030)	(0.025)	(0.031)	(0.034)	(0.050)
FT×SIN	2.364 ***	2.321 ***	1.021 **	3.466 ***	2.021 *
	(0.620)	(0.602)	(0.553)	(0.750)	(1.487)
IN×SIN	-0.028	-0.032 *	0.042 *	0.003	-0.113 *
	(0.029)	(0.028)	(0.030)	(0.042)	(0.057)
HHI	0.252	0.042	0.531	-1.901 *	0.069
	(0.902)	(1.014)	(0.820)	(1.411)	(2.181)
PS	-0.013	-0.044	0.015	0.270	1.038 **
	(0.274)	(0.283)	(0.211)	(0.375)	(0.510)
Coast	0.026	0.030 *	0.013	0.111 **	0.223 ***
	(0.029)	(0.027)	(0.024)	(0.042)	(0.070)
NE	-0.490 **	0.059	-0.455 **	-0.003	0.649 **
	(0.155)	(0.206)	(0.214)	(0.152)	(0.261)
N	0.00024 ***	0.0008 **	0.0021 ***	0.0008 *	0.0063 ***
	(0.0001)	(0.0004)	(0.0004)	(0.0006)	(0.0015)
Plant	0.081 ***	0.063 ***	0.128 ***	0.263 ***	0.235 ***
	(0.010)	(0.010)	(0.016)	(0.029)	(0.069)
Product	0.096 ***	0.096 ***	0.149 ***	0.191 ***	0.180 ***
	(0.012)	(0.012)	(0.015)	(0.025)	(0.029)
Elasticity_FT	0.0456	-0.0812	-0.0098	-0.0008	0.0206
Elasticity_IN	0.0230	0.0106	-0.0014	0.0134	0.0104
Elasticity_SIN	0.0377	0.0542	0.0079	0.0846	0.0170
Pseudo R^2	0.15	0.14	0.14	0.15	0.12
Obs.	2268	2268	2268	2268	2268
Censored Obs.	652	727	826	727	652
L.B.	0.518	0.874	1.11	1.43	1.93

Notes: Numbers in parentheses are standard deviation. ***, **, and * indicate significance at 1%, 5%, and 10% level, respectively.
Coast and NW are dummies indicating coastal states and north-western states, respectively. L.B. indicates the lower bound productivity which is the left-censoring limit. Model also includes industry dummies and constant term; effects not shown.
To save space we only report the elasticities of our main competition pressures at the mean here.

Table 4.8. Cumulative Effect of Competition Sources (based on CLAD Model)

	(A) 1994 Average	(B) 2007 Average	(C) Growth rate % = (B/A)×100 -100	(D) Elasticity from CLAD for 10th percentile	(E) Cumulative Effects = C×D
Free Trade	0.453	0.623	37.53	0.045	1.69
Infra	1.300	1.670	28.46	0.023	0.65
Surrounding states' Infra	0.360	0.460	27.78	0.038	1.06
Raw TFP	2.160	2.243	3.85		

Appendix

Table 4.A.1. Industry Classification

Industry	Details (NIC codes)
1 Food	NIC 15 Manufacture of food products and beverages NIC 16 Manufacture of tobacco products
2 Textiles & Apparel	NIC 17 Manufacture of textiles NIC 18 Manufacture of wearing apparel; dressing and dyeing of fur NIC 19 Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear
3 Wood, paper, & Printing	NIC 20 Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials NIC 21 Manufacture of paper and paper products NIC 22 Publishing, printing and reproduction of recorded media NIC 361 Manufacture of furniture
4 Chemicals & Rubber	NIC 24 Manufacture of chemicals and chemical products NIC 25 Manufacture of rubber and plastics products
5 Fuels & Mineral	NIC 23 Manufacture of coke, refined petroleum products and nuclear fuel NIC 26 Manufacture of other non-metallic mineral products
6 Metals	NIC 27 Manufacture of basic metals NIC 28 Manufacture of fabricated metal products, except machinery and equipment
7 Machinery	NIC 29 Manufacture of machinery and equipment n.e.c.
8 Electricals Machinery & Electronics	NIC 30 Manufacture of office, accounting and computing machinery NIC 31 Manufacture of electrical machinery and apparatus n.e.c. NIC 32 Manufacture of radio, television and communication equipment and apparatus NIC 33 Manufacture of medical, precision and optical instruments, watches and clocks NIC 369 Manufacturing n.e.c.
9 Transport Vehicles & Equipment	NIC 34 Manufacture of motor vehicles, trailers and semi-trailers NIC 35 Manufacture of other transport equipment

Table 4.A.2. Descriptive Statistics on Indian Manufacturing Industries

	Industry	1	2	3	4	5	6	7	8	9
Output	Mean	80.57	56.51	43.98	88.49	649.55	138.79	95.52	137.72	218.05
	S.D.	3.26	1.35	1.87	2.39	73.59	8.15	5.96	6.32	15.09
	AGR	7.42	7.85	4.14	6.46	7.57	7.43	6.53	13.34	9.98
Localization (Wlny)	Mean	2.60	2.64	2.08	2.67	2.64	2.91	2.71	2.81	3.39
	S.D.	0.01	0.01	0.03	0.01	0.03	0.01	0.02	0.02	0.02
	AGR	1.32	1.72	1.23	1.17	1.68	1.84	1.03	2.45	2.96
Urbanization (UE)	Mean	5.72	6.73	7.93	8.79	6.44	7.34	9.10	10.42	7.41
	S.D.	0.10	0.12	0.21	0.09	0.18	0.12	0.18	0.16	0.16
	AGR	8.42	9.02	8.78	9.65	10.33	9.03	7.80	9.59	10.30
Labor	Mean	4.07	4.10	4.24	5.10	16.36	9.89	10.05	5.72	13.24
	S.D.	0.15	0.12	0.23	0.13	1.37	1.14	0.88	0.25	0.76
	AGR	4.18	3.46	6.20	5.33	8.62	5.38	3.25	4.94	5.38
Capital Stock	Mean	43.22	48.17	60.90	90.35	556.98	205.64	55.18	46.93	119.51
	S.D.	1.78	1.36	3.13	2.72	59.38	16.38	2.91	2.20	7.51
	AGR	7.72	7.36	7.02	6.15	12.96	5.16	4.75	7.34	9.43
Materials	Mean	46.45	25.67	20.80	47.02	620.16	69.49	50.41	58.06	110.62
	S.D.	1.60	0.56	0.79	1.31	88.01	3.40	2.77	2.41	7.82
	AGR	7.00	3.87	5.75	6.40	14.32	9.35	6.83	10.31	9.26
Energy	Mean	1.55	2.49	3.92	4.44	18.67	8.71	1.22	0.84	2.64
	S.D.	0.06	0.07	0.28	0.16	1.01	0.58	0.09	0.03	0.13
	AGR	0.47	0.57	-3.95	0.88	2.26	3.23	-2.40	2.57	2.02

Note: All (mean) values are constant Rs. Crore (1993=100). AGR indicates annual growth rate during 1994-2007.

Table 4.A.3. Number of firms and Observation by Industries and State

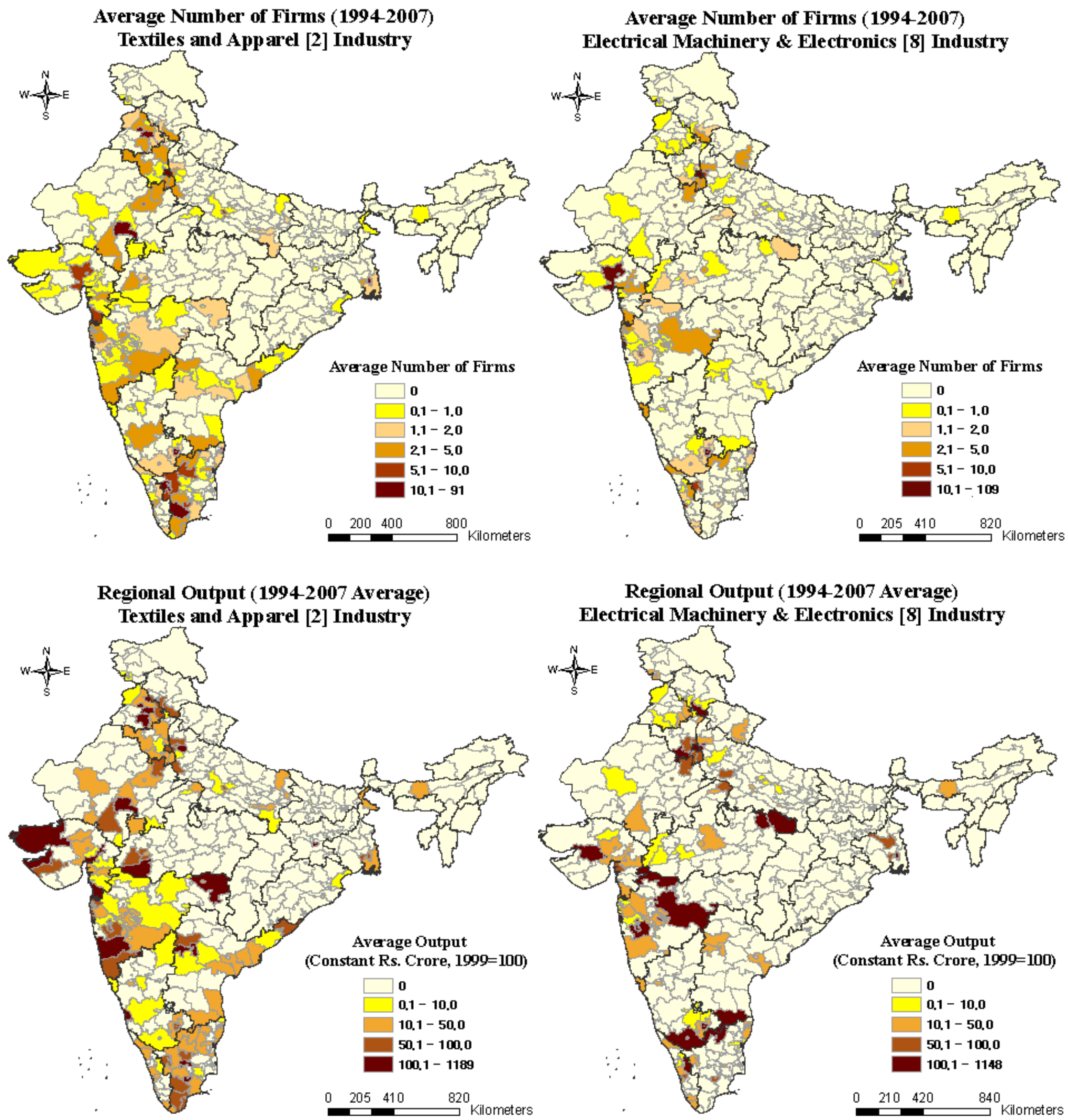
Number of firms (1994-2007)

	Industry									
State	1	2	3	4	5	6	7	8	9	Total
1	7	7	1	10	1	8		7	2	43
2	46	63	18	48	1	46	6	17	21	266
3	22	28	5	35	4	28	9	25	35	191
4	113	98	54	137	34	132	56	114	102	840
5	24	60	1	52	19	24	10	18	5	213
6	55	43	33	65	10	36	8	35	13	298
7	9	2	5	10	8	22	2		3	61
8	26	1	7	9	14	11	2	2		72
9	198	105	20	117	43	137	44	56	24	744
10	12	3	5	9	12	25	1	6		73
11	63	15	9	54	17	44	6	23	4	235
12	72	102	39	328	56	109	73	72	21	872
13	183	237	123	785	90	285	178	265	113	2,259
14	111	59	27	181	53	73	23	55	16	598
15	61	62	24	62	12	40	40	62	21	384
16	6	2	1	21	5	10	2	8	8	63
17	48	23	19	59	10	14	3	29	2	207
18	126	295	39	178	35	95	78	95	102	1,043
Total	1,182	1,205	430	2,160	424	1,139	541	889	492	8,462

Observations (1994-2007)

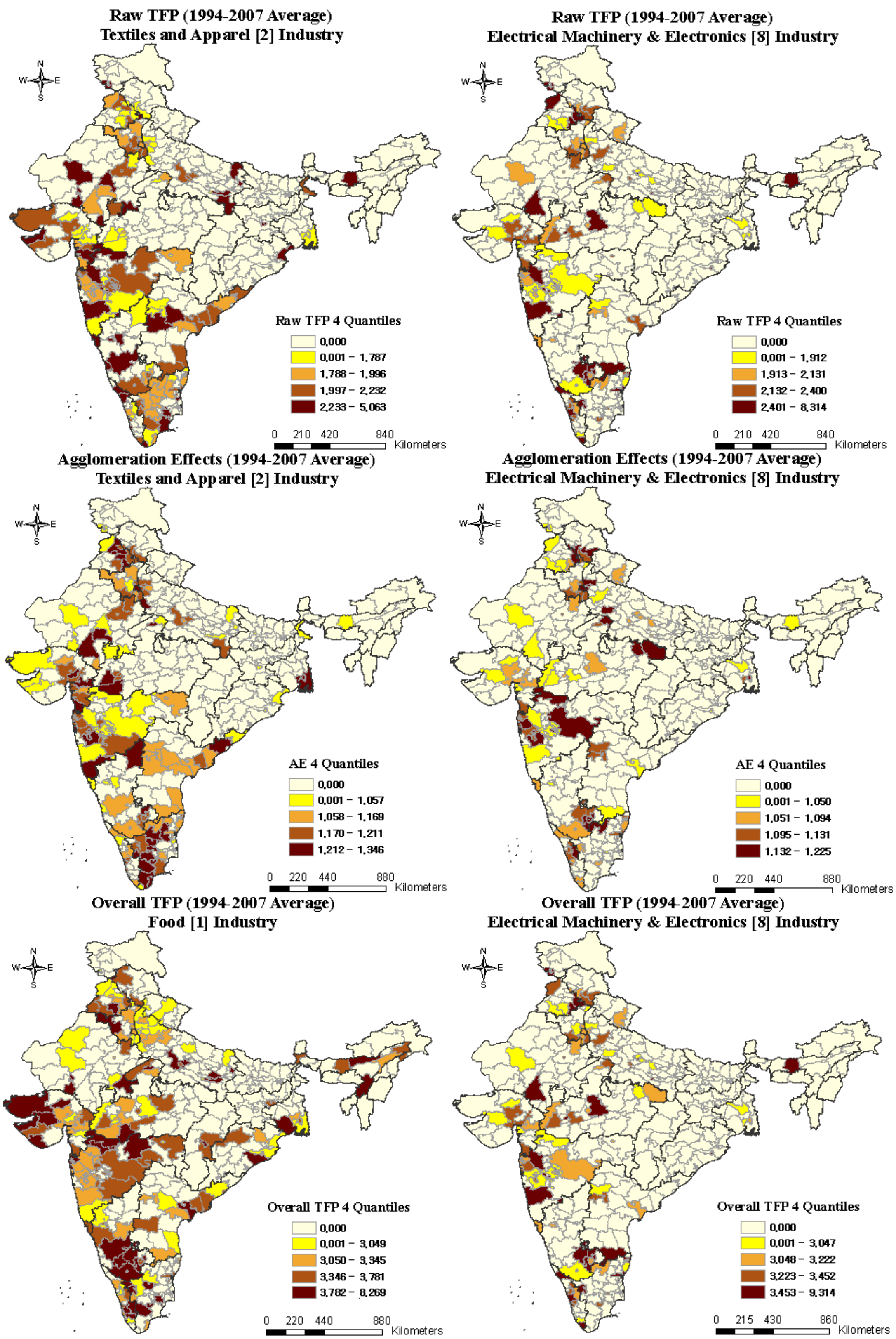
	Industry									
State	1	2	3	4	5	6	7	8	9	Total
1	73	40	6	91	14	50		70	13	357
2	329	457	135	366	5	324	55	116	177	1,964
3	146	269	42	297	19	206	76	166	321	1,542
4	693	647	349	928	215	812	350	795	716	5,505
5	161	462	7	453	125	198	64	151	43	1,664
6	437	336	218	495	95	224	87	241	129	2,262
7	55	6	27	65	49	153	20		29	404
8	164	2	58	51	89	53	11	15		443
9	1,473	851	161	934	337	965	361	408	197	5,687
10	58	14	50	74	107	203	10	26		542
11	422	139	55	380	107	285	61	177	35	1,661
12	442	753	273	2,679	342	789	595	528	165	6,566
13	1,283	1,679	902	5,989	703	1,999	1,500	1,950	838	16,843
14	655	440	181	1,321	426	473	169	348	141	4,154
15	425	303	120	461	108	241	316	418	178	2,570
16	35	13	4	154	31	86	7	57	49	436
17	402	146	132	466	98	89	19	215	18	1,585
18	907	1,974	268	1,428	294	639	608	696	806	7,620
Total	8,160	8,531	2,988	16,632	3,164	7,789	4,309	6,377	3,855	61,805

Note: For reference to industries and states, see Table 4.4 (Notes).



Appendix Figure 4.A.1. Spatial Variation of Number of Firms and Regional Output

Notes: Spatial units are 3-digit postal areas. Thick lines indicate state borders.



Appendix Figure 4.A.2. Spatial Variation of Raw TFP, AE, and Overall TFP

BIBLIOGRAPHY

Acemoglu, D., and M. Dell. 2010. “Productivity Differences between and Within Countries.” *American Economic Journal: Macroeconomics* 2 (1): 169-188.

Ades, A., and E.L. Glaeser. 1995. “Trade and Circuses: Explaining Urban Giants.” *Quarterly Journal of Economics* 110 (1): 195-227.

Ai, C., and E.C. Norton. 2003. “Interaction Terms in Logit and Probit Models.” *Economic Letters* 80 (1): 123-129.

Alfaro, L., and A. Chari. 2009. “India Transformed? Insights 1988-2005.” Working Paper, UNC-Chapel Hill and Harvard Business School.

Alonso-Villar, O. 2001. “Large Metropolises in the Third World: An Explanation.” *Urban Studies* 38 (8): 1359-1371.

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

Anderson, J., and E. van Wincoop. 2003. “Gravity with Gravitas: A Solution to the Border Puzzle.” *American Economic Review* 93 (1): 170-192.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht.

Anselin, L., and A. Bera. 1998. “Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics.” In: Giles, D.E.A., Ullah, A. (Eds.) *Handbook of Applied Economic Statistics*. Marcel Dekker, New York, pp. 237-289.

Anselin, L., R.J. Florax, and S.J Rey. 2004. “Econometrics for Spatial Models: Recent Advances.” In: Anselin, L., Florax, R., Rey, S.J. (Eds.) *Advances in Spatial Econometrics. Methodology, Tools and Applications*. Springer-Verlag, Berlin, pp. 1-25.

Baldwin, R.E., and T. Okubo. 2006. “Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting.” *Journal of Economic Geography* 6 (3): 323-346.

Baldwin, R.E., R. Forslid, P. Martin, G.I.P. Ottaviano, and F. Robert–Nicoud. 2003. *Economic Geography and Public Policy*. Princeton University Press, Princeton.

Barro, R.J., and X. Sala-i-Martin. 2003. *Economic Growth*. 2nd edition. MIT Press, Cambridge, MA.

Beaudry, P., F. Collard, and D.A. Green. 2005. "Changes in the World Distribution of Output per Worker, 1960-1998: How a Standard Decomposition Tells us an Unorthodox Story." *Review of Economics and Statistics* 87 (4): 741-753.

Behrens, K. 2006. "Do Changes in Transport Costs and Tariffs Shape the Space-Economy in the Same Way?" *Papers in Regional Science* 85 (3): 379-399.

Bernard, A. B., J.B. Jensen, and P.K. Schott. 2003. "Plants and Productivity in International Trade." *American Economic Review* 93 (4): 1268-1290.

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

Bertrand, M., P. Mehta, and S. Mullainathan, 2002. "Ferreting Out Tunneling: An Application to Indian Business Groups." *Quarterly Journal of Economics* 117 (1): 121-148.

Broda, C., and D. Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121 (2): 541-585.

Chari, A., and N. Gupta. 2008. "Incumbents and Protectionism: The Political Economy of Foreign Entry Liberalization." *Journal of Financial Economics* 88 (3): 633-656.

Chen, N., and D. Novy. 2009. "International Trade Integration: A Disaggregated Approach." CESIFO Working Paper No. 2595, University of Warwick, UK.

Ciccone, A. 2002. "Agglomeration Effects in Europe." *European Economic Review* 46 (2): 213-227.

Ciccone, A., and R.E. Hall. 1996. "Productivity and the Density of Economic Activity." *American Economic Review* 86 (1): 54-70.

Coe, D., and E. Helpman. 1995. "International R&D Spillovers." *European Economic Review* 39 (5): 859-887.

Combes, P-P., G. Duranton, and H.G. Overman. 2005. "Agglomeration and the Adjustment of the Spatial Economy." *Papers in Regional Science* 84 (3): 311-349.

Combes, P-P., T. Mayer, and J-F. Thisse. 2008. *Economic Geography: The Integration of Regions and Nations*. Princeton University Press, Princeton.

Deller, S.C., T.H. Tsai, D.W. Marcouiller, and D.B.K. English. 2001. "The Role of Amenities and Quality of Life in Rural Economic Growth." *American Journal of Agricultural Economics* 83 (2): 352-365.

Dinc, I. S., and N. Gupta, 2007. "The Decision to Privatize: Finance, Politics and Patronage." Mimeo, MIT, Cambridge, MA.

Dixit, A.K., and J.E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67 (3): 297-308.

Duranton, G., and D. Puga. 2004. "Micro-foundations of Urban Agglomeration Economies." In: Henderson, V., Thisse, J.F. (Eds.) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland, Amsterdam, pp. 2064-2117.

Feenstra R.C. 1994. "New Product Varieties and the Measurement of International Prices." *American Economic Review* 84 (1): 157-177.

Fingleton, B., and J. Le Gallo. 2008. "Estimating Spatial Models with Endogenous Variables, a Spatial Lag and Spatially Dependent Disturbances: Finite Sample Properties." *Papers in Regional Science* 87 (3): 319-339.

Fisman, R., and T. Khanna, 2004. "Facilitating Development: The Role of Business Groups". *World Development* 32 (4): 609-628.

Fujita, M., P.R. Krugman, and A.J. Venables. 1999. *The Spatial Economy. Cities, Regions and International Trade*. MIT Press, Cambridge MA.

Giles, J.A., and C.L. Williams. 2000. "Export-led Growth: A Survey of the Empirical Literature and Some Non-causality Results." Part I. *Journal of International Trade & Economic Development* 9 (3): 261-337.

Giles, J.A., and C.L. Williams. 2000b. "Export-led Growth: A Survey of the Empirical Literature and Some Non-causality Results." Part II. *Journal of International Trade & Economic Development* 9 (4): 445-470.

Glaeser, E.L., and J.D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4): 983-1028.

Goldberg, P., A. Khandelwal, N. Pavcnik, and P. Topalova. 2008. "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India." NBER Working Paper 14416, Cambridge, MA.

Goldberg, P., A. Khandelwal, N. Pavcnik, and P. Topalova. 2009. "Trade Liberalization and New Imported Inputs." *American Economic Review* 99 (2): 494-500.

Gopinath, M., and H. Kim (Editors). 2009. *Globalization and the Rural-Urban Divide*. Seoul, Korea: Seoul National University Press.

Greenaway, D., W. Morgan, and P. Wright. 2002. "Trade Liberalization and Growth in Developing Countries." *Journal of Development Economics* 67 (1): 229-244.

Hahn, C.H. 2004. “Exporting and Performance of Plants: Evidence from Korean Manufacturing.” NBER Working Paper 10208, Cambridge, MA.

Hanson, G.H. 1998. “Regional Adjustment to Trade Liberalization.” *Regional science and urban economics* 28 (4): 419-444.

Harrigan, J. 1999. “Estimation of Cross-Country Differences in Industry Production Functions.” *Journal of International Economics* 47 (2): 267-293.

Hausman, J.A. 1978. “Specification Tests in Econometrics.” *Econometrica* 46 (6): 1251-1271.

Head, K., and T. Mayer. 2004. “The Empirics of Agglomeration and Trade,” in J.V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics: Cities and Geography*, vol. 4, North Holland, Amsterdam: pp. 2609-2665.

Helpman, E. 2006. “Trade, FDI, and the Organization of Firms.” *Journal of Economic Literature* 44 (3): 589-630.

Helpman, E., M. Melitz, and S. Yeaple. 2004. “Export Versus FDI with Heterogeneous Firms.” *American Economic Review* 94 (1): 300-316.

Henderson, J.V. 1986. “Efficiency of Resource Usage and City Size.” *Journal of Urban Economics* 19 (1): 47-70.

Henderson, J.V. 1988. *Urban Development: Theory, Fact and Illusion*. Oxford University Press, New York.

Henderson, J.V. 2003. “Marshall’s Scale Economies.” *Journal of Urban Economics* 53 (1): 1-28.

Henderson, J.V., T. Lee, and Y.J. Lee. 2001. “Scale Externalities in Korea.” *Journal of Urban Economics* 49 (3): 479-504.

Henderson, J.V., Z. Shalizi, and A.J. Venables. 2001. “Geography and Development.” *Journal of Economic Geography* 1 (1): 81-105.

Imbs J., and I. Méjean. 2009. “Elasticity of Optimism.” CEPR Discussion Paper 7177.

Jones, C.I. 1997. “On the Evolution of the World Income Distribution.” *Journal of Economic Perspectives* 11 (3): 19-36.

Kelejian, H.H., and I.R. Prucha. 1998. “A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances.” *Journal of Real Estate Finance and Economics* 17 (1): 99-121.

Kelejian, H.H., and I.R. Prucha. 1999. “A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model.” *International Economic Review* 40 (2): 509-533.

Kim, E., 2000. “Trade Liberalization and Productivity Growth in Korea Manufacturing Industries: Price Protection, Market Power, and Scale Efficiency.” *Journal of Development Economics* 62 (1): 55-83.

Krishna, P., and D. Mitra. 1998. “Trade Liberalization, Market Discipline and Productivity Growth: New Evidence from India.” *Journal of Development Economics* 56 (2): 447-62.

Krugman, P.R. 1991. “Increasing Returns and Economic Geography.” *Journal of Political Economy* 99 (3): 483-499.

Krugman, P.R., and R. Livas Elizondo. 1996. “Trade Policy and the Third World Metropolis.” *Journal of Development Economics* 49 (1): 137-150.

Lall, S.V., Z. Shalizi, and U. Deichmann. 2004. “Agglomeration Economies and Productivity in Indian Industry.” *Journal of Development Economics* 73 (2): 643-674.

Lee, B.S., S. Kim, and S.H. Hong. 2005. “Sectoral Manufacturing Productivity Growth in Korean Regions.” *Urban Studies* 42 (7): 1201-1219.

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

Marshall, A., 1890. *Principles of Economics,* MacMillan, London (8th edition 1920).

McGranahan, D.A. 1999. *Natural Amenities Drive Rural Population Change*. Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture. Agricultural Economic Report No. 781.

Melitz, M. 2003. “The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity.” *Econometrica* 71 (6): 1695-1725.

Melitz, M., and G.I.P. Ottaviano. 2008. “Market Size, Trade, and Productivity.” *Review of Economic Studies* 75 (1): 295-316.

Mitra, D., and B. Ural. 2008. “Indian Manufacturing: A Slow Sector in A Rapidly Growing Economy.” *Journal of International Trade and Economic Development* 17 (4): 525-559.

Moomaw, R.L. 1981. “Productivity and City Size: A Critique of the Evidence.” *Quarterly Journal of Economics* 96 (4): 675-688.

Moomaw, R.L. 1983. "Spatial Productivity Variations in Manufacturing: A Critical Survey of Cross-Sectional Analysis." *International Regional Science Review* 8 (1): 1-22.

Nakamura, R. 1985. "Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities." *Journal of Urban Economics* 17 (1): 108-124.

Novy, D. 2008. "Gravity Redux: Measuring International Trade Costs with Panel Data." Working Paper, University of Warwick, UK.

Oh, I., J. Lee, and A. Heshmati. 2008. "Total Factor productivity in Korean Manufacturing Industries." *Global Economic Review* 37 (1): 23-50.

Oosterhaven, J, and L. Broersma. 2007. "Sector Structure and Cluster Economies: A Decomposition of Regional Labour Productivity." *Regional Studies* 41 (5): 639-659.

Pagan, A., and F. Vella. 1989. "Diagnostic Tests for Models Based on Individual Data: A Survey." *Journal of Applied Econometrics* 4: S29-S59.

Rappaport, J. 2009. "The Increasing Importance of Quality of Life." *Journal of Economic Geography* 9 (6): 779-804.

Rice, P., A.J. Venables, and E. Patacchini. 2006. "Spatial Determinants of Productivity: Analysis for the Regions of Great Britain." *Regional Science and Urban Economics* 36 (6): 727-752.

Rosenthal, S.S., and W.C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In: Henderson, V., Thisse, J.F. (Eds.) *Handbook of Regional and Urban Economics*, vol. 4. North-Holland, Amsterdam, pp. 2118-2171.

Ruan, J., and M. Gopinath. 2008. "Global Productivity Distribution and Trade Liberalization: Evidence from Processed Food Industries." *European Review of Agricultural Economics* 35 (4): 439-460.

Saito, H., and M. Gopinath. 2009. "Plants' Self-Selection, Agglomeration Economies and Regional Productivity in Chile." *Journal of Economic Geography* 9 (4): 539-558.

Schor, A. 2004. "Heterogeneous Productivity Responses to Tariff Reduction: Evidence from Brazilian Manufacturing Firms." *Journal of Development Economics* 75 (2): 373-396.

Segal, D. 1976. "Are There Returns to Scale in City Size?" *Review of Economics and Statistics* 58(3): 339-50.

Sivadasan, J. 2009. "Barriers to Competition and Productivity: Evidence from India." *B.E. Journal of Economic Analysis and Policy* 9 (1) (Advances).

Stock, J.H., and M. Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In: Stock, J.H., Andrews, D.W.K. (Eds.) *Identification and Inference for Econometric Models: a Festschrift in honor of Thomas Rothenberg*. Cambridge University Press, Cambridge, UK, pp. 80-108.

Stoneman, P. 1995 (ed.). *Handbook of the Economics of Innovation and Technological Change*. Blackwell, Oxford.

Sveikauskas, L. A. 1975. "The Productivity of Cities." *Quarterly Journal of Economics* 89 (3): 393-413.

Syverson, C. 2004. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics* 86 (2): 534-550.

Topalova, P. 2007. "Trade Liberalization and Firm Productivity: The Case of India." IMF Working Paper, WP/04/28, Washington DC.

Tybout, J. 2000. "Manufacturing Firms in Developing Countries: How Well They Do and Why?" *Journal of Economic Literature* 38 (1): 11-44.

Vayá, E., E. López-Bazo, M. Moreno, and J. Suriñach. 2004. "Growth and Externalities Across Economies. An Empirical Analysis Using Spatial Econometrics." In: Anselin, L., Florax, R., Rey, S.J. (Eds.) *Advances in Spatial Econometrics. Methodology, Tools and Applications*. Springer-Verlag, Berlin, pp. 433-455.

Venables, A.J. 1996. "Equilibrium Locations of Vertically Linked Industries." *International Economic Review* 37(2): 341-360.

Venables, A.J. 2005. "Spatial Disparities in Developing Countries: Cities, Regions, and International Trade." *Journal of Economic Geography* 5(1): 3-21.

Wilhelm, M.O. 2008. "Practical Considerations for Choosing Between Tobit and SCLS or CLAD Estimators for Censored Regression Models with an Application to Charitable Giving." *Oxford Bulletin of Economics and Statistics* 70 (4): 559-582.

Wu, J., and M. Gopinath. 2008. "What Causes Spatial Variations in Economic Development in the United States?" *American Journal of Agricultural Economics* 90 (2): 392-408.

Yasar, M., and C. J. Morrison. 2007. "International Linkages and Productivity at the Plant Level: Foreign Direct Investment, Exports, Imports and Licensing." *Journal of International Economics* 71 (2): 373-388.