Aalto University
School of Economics

The Impact of International Trade on Guangdong's Industrial Agglomeration -an Empirical Analysis of Manufacturing Industry in Guangdong Province, China

Economics

Master's thesis

Xiaoying Feng

2012

Department of Economics
Aalto University
School of Economics

AALTO UNIVERSITY SCHOOL OF ECONOMICS ABSTRACT

Department of Economics 02.04.2012

Master's Thesis

Xiaoying Feng

The Impact of International Trade on Guangdong's Industrial Agglomeration

— an Empirical Analysis of Manufacturing Industry in Guangdong Province, China

Guangdong, the forerunner of implementing open door policies in China, has experienced the prevailing industrial agglomeration since 1990s. Previous research attributed this concentration to many factors other than international trade which is a key characteristic of Guangdong economic development. Based on the study of the manufacturing industry in Guangdong, this paper aims to find out whether international trade pushes industrial agglomeration. This paper will observe eleven manufacturing sectors in Guangdong over the 2000-2009 period and build the models with transformed measures of industrial agglomeration and international trade as well as two other factors. Next the models will be examined in different ways such as Cross-Section Weighted Least Squares (CSWLS), Two Stage Least Squares (2SLS) and Generalized Method of Moments (GMM). The finding of this paper is that international trade does have a positive impact on agglomeration at least in some sectors of manufacturing industry in Guangdong Province, China.

Keywords: Industrial Agglomeration, International Trade, Guangdong, Manufacturing Industry

Contents

List of Tables

List of Figures

1. Introduction

1.1Background and motivation

From 1990s the global economy and technology have experienced an accelerated process of globalization. The revolution of information technology and optimized industrial structure has stimulated the allocation of resources as well as the transformation of industry and technology all over the world. Following this trend, Guangdong Province, the pioneer in the reforms and open door policies in China, has undertaken the regional industrial agglomeration on large scale, contributing to sky-rocketing regional economic growth and strong industrial competitiveness.[1] The industrial agglomeration in Guangdong is classified as foreign-investment-induced agglomeration and local-development agglomeration, both characterized by development zones (industrial parks) and professional towns. Some typical examples of spatially integration are as follows. The electronic information manufacturing park in Shenzhen city created 753.8 billion Yuan in 2009 and ranked the first in the output of mobile phones, integrated circuits, microcomputers, communication cables (and others) in China. The garment-specialized industrial area in Humen town generated 15.5 billion Yuan sales in 2009. The concentration of professional lighting industry in Guzhen town consists of more than 2500 factories and has cultivated one of the four biggest specialized lighting markets in the world.[2] Overall, the industrial

---

[1] The open door policies in China are termed as China's policies of opening up to the outside world. After Xiaoping DENG took office, the government made policies of encouraging foreign trades by abundant of benefits to firms such as low tariff and firstly implemented the promoting rules in Guangdong. Attracted by the political benefits, many firms were founded in Guangdong, doing businesses with foreign countries and subsequently the rapid economic development came up in Guangdong.

[2] Source: Guangdong Provincial Department of Foreign Trade and Economic Cooperation, 2007. Conditions of development zones in Guangdong Province, 2007.[online]Available at: <http://www.gddoftec.gov.cn/dept_detail.asp?deptid=1048&channalid=1293&contentid=10513>[Accessed 23 September 2010].

agglomeration in Guangdong Province has taken advantages of its geography, resources and policies to develop the industrial economy.

Recent studies in China have paid attention to the phenomena of regional industrial agglomeration in Guangdong and given explanations from various perspectives. Generally the industrial agglomeration in Guangdong can be attributed to internal and external reasons. Internal reasons refer to knowledge spillover, increasing scales of returns and shrinking costs of transportations, intermediate products and labors. Particularly, knowledge spillover is the main force for high-tech industrial agglomeration; increasing scales of returns exert less important influences on technology-intensive agglomeration than labor-intensive one (Li and Li, 2002). External factors include industrial transfers (Wang, 2005), foreign direct investments (Shao, 2010), government's promotions (Ren, 2005), market effects (Liu, 2003) and so forth. Luo (2002) pointed out that the industrial agglomeration in Guangdong Province is an embedded–type agglomeration. Namely the industry depends on geographical, political and low-cost benefits to attract direct investments inside; it also builds up value-added manufacturing base outside. Gradually the regional industrial clusters form. Some other authorities elaborated the incentives from locational merits (Yang and Feng, 2002; He, 2002), cultural linkages (Zheng, 2002), entrepreneurship (Li, 2000), local production system (Wang, 2001) and so on. Their points of view can be summarized as the followings. First of all, Hong Kong, Taiwan and Guangdong Province were allied in history and culture, creating some social networks. Those social networks have set up the cultural foundation for investments in Guangdong from Hong Kong and Taiwan. Second, benefiting from "East to West" gradient of reforms and open door policies in China, Guangdong got political advantages and sound institutional settings for foreign direct investments (FDI). Third, growing costs of labors and lands in East Asia strengthen Guangdong's low-cost comparative advantage, enforcing the transfer of labor-intensive industries to Guangdong and making Guangdong the leading add-valued manufacturing base in Asia even in the world. Apart from those explanations mentioned above, I lay priority to the upward degree of economic internationalization.

After China's WTO accession, the openness of industries in Guangdong has moved from the pilot stage to a completed open-up stage at a striking high speed, consistent with extraordinary growth of FDI and international trade. In 2005 the amount of international trade in Guangdong was $ 1.46 trillion, which already exceeded one trillion U.S. dollars. Guangdong's foreign trade volume increased at 20.3% annual rate, which was 10% higher than other regions' in China. Guangdong's international trade volume also took up 30% of China's and kept the first place for more than 20 years consecutively. Although in 2009 Guangdong suffered a heavy shrink in international trade due to the world-wide economic downturn, it started to recover in 2010, reaching 784.6 billion dollars at the rate of 28.4%. This growing rate was lower than Jiangsu Province's and Beijing's; however, international trades in Guangdong still made up the largest proportion of foreign trades in China. An inference can be drawn from the foregoing facts that the process of economic internationalization has kept pace with the development of industrial agglomeration in Guangdong Province.

In addition, one more interesting finding is that the export rate in industrial agglomeration zones is growing faster than that in non-industrial agglomeration zones. In 2004, 69 industrial concentrated zones created 185.83 billion Yuan, accounting for 11.59% of total GDP in Guangdong Province. And the value of import and export in those zones took up 19.13% of the Guangdong Province total and reached 68.32 billion dollars, of which 42.17 billion dollars was created by exporting. From then on, imports and exports in industrial clusters have experienced upward trends and hit the top of 135.36 billion dollars until 2007; while the GDP in those areas has fluctuated between 444 and 649 billion Yuan. The economic performance of agglomerate zones was weakened after 2008 because of global recession. However, the quantity of integrated zones reversely enhanced to 97 in 2009 and the import and export volume in those areas decreased at a lower rate than the average rate of Guangdong.[3]

Above all, the analysis of the industrial agglomeration in Guangdong should take into account an open economy perspective. There are two main indexes widely used in the areas of open economy: FDI and international trade. Many researchers and policy-makers have worked on the co-relationship of FDI and regional industrial agglomeration and found out that FDI improves the process of regional agglomeration (Shao, 2010; Liu, 2002). Whereas analyzing the relationship between international trade and regional agglomeration, specialists paid more attention to the issues about agglomeration improving international trade. What about the impact of international trade on regional industrial agglomeration? Even though some literature outside China has focused on the effect of international trade on geographic concentration (Rauch, 1991; Krugman and Venables, 1995 cited in Ottaviano and Puga, 1997; Haaparanta, 1998), little research has studied the Guangdong economic geography importance in China from an empirical point of view. Based on an empirical analysis of Guangdong Province in China, I would like to continue the discussion of the impact of international trade on regional industrial agglomeration aiming at manufacturing industry, because this industry has high degrees of agglomeration and

---

[3] Some important figures can be found in Appendix 1. Source: Guangdong Provincial Department of Foreign Trade and Economic Cooperation, 2010. Conditions of development zones in Guangdong Province, 2004-2009. [online] Available at:<http://www.gddoftec.gov.cn/dept_sub.asp?deptid=1048&channalid=1293>[Accessed 23 September 2010].

international trade. The hypothesis of this paper is that international trade may lead to industrial agglomeration in the manufacturing industry in Guangdong Province, China.

1.2 Main results

The results acquired from three estimations consistently point out that there is a significantly positive effect of international trade on agglomeration. This finding is consistent with findings in earlier literature and the hypothesis of this paper. Contrary to previous theoretical studies, the outcomes of the estimated regressions in this study show negative influences of internal economies of scale and home market effects, both of which are causes for geographical integration. The effects of these two factors are not clear due to limitations in processed data; but these findings would not disturb our studies of the impact of foreign trades on integration.

Regarding estimating approaches for models, the Two Stage Least Squares (2SLS) employed in the fixed effects model using Cross-Section Weighted Least Squares (CSWLS) seems superior among those three methods, because both the endogeneity and the heteroscedasticity are under control to some extent and the quality of chosen instrumental variable (IV) sounds good.

1.3 Structure of study

The paper is organized as follows. Section 2 reviews the theoretical backgrounds of emerging geographical concentrations and relevant literature on agglomeration from measures to factors. Main methodologies are introduced in Section 3. Then in Section 4, measures and data issues are discussed; additionally variables in the regressions are introduced. Section 5 builds up the estimated models and presents the estimation results in three different ways with comparisons. Finally Section 6 concludes.

2. Literature Review

Since the research question of this study is a subset of research in the economic geography area, let us start from the development of economic geography to capture a big picture of this paper's research background with respect to spatial agglomeration. It does help to understand our research question better and more clearly. Then we will review earlier literature about the critical object—industrial agglomeration in this study and provide theoretical grounds for its potential causes.

2.1 The development of economic geography

The inequality in population and activities across the landscape in real life triggered economists' interests in economic geography-- a study of where and why economic activity happens—long time ago. Marshall (1920 cited in Redding, 2009) brought forth three reasons behind the clustering of economic activities: knowledge spillovers, merits of pooling specialized skills and linkages associated with local markets. After that, many researches in urban and regional economics studied the existence of cities, the distribution of population in spatial terms, localized production across regions and so on. It's hypothesized that, developing from low levels, cities or countries experience regional divergence and later concentrate industrialization on a limited location (Williamson, 1965). The formation of cities is consistent with industrial agglomeration. These so-called "urban economics" and "regional science" remained the main body of economic geography and were exposed to the fourth wave of the increasing-returns revolution —"new economic geography (NEG)". In NEG, spatial interdependencies are the focus on economic agglomeration under regions' integration and new trade and new growth theories which are synthesized in terms of locations.[4]

When it comes to determinants of location, economic geography can be divided into first-nature and second-nature geography. In first-nature geography, physical locational fundamentals should be taken into account, for example, coasts, plains and other natural endowments; at the focus are exogenously given features of various locations. First nature is widely used to account for locational preference of heavy industries in the Industrial Revolution, yet it fails to give convincing reasons to many other centripetal processes of economic activities such as the formation of Silicon Valley in the USA. While in second-nature geography, the location and behaviors of economic agents related to each other in a region are under consideration; endogenous factors are the objects of investigations. With regard to these two "natures", geographic economics aims at shedding light on the economic forces by controlling first nature. It pays attention to the second nature which implies economic actors' behaviors upon the first nature.

From another point of view, economic geography starts from a static situation where locations and economic activities are homogenous across space. Then economic geography tries to find out the underlying forces that allow a small

---

4 Fujita and other economists summarized theories in geography economy in their classical book. To learn more about the development of geography economy, please read Fujita, M., Krugman, P. and Venables, A., 1999. The Spatial Economy: Cities, Regions and International Trade. Cambridge, MA: MIT Press.

asymmetric change to redistribute the unbalanced activities. Even though there are many theoretical models that have been used to address this issue, I restrict myself to a short summary of three classifications (Brülhart, 2000) under this intellectual background.

First of all, in neo-classical models, economists generally subscribe to exogenous determination of location. They assume perfect competition, homogeneous products, constant returns of scale in economic activities and completely rational agents who make geographic decisions to obtain the maximum profits. Without trade costs, demand distributed across regions affects only trade patterns but not production locations. Otherwise, the spatial dispersion of production will be adjusted by the change of demand. However, over time neo-classical models are criticized for their limitations on explaining real-life phenomena with too strict assumptions. These models only emphasize patterns of land uses and profit maximization, without considering other factors such as environment and historical causations.

Secondly, in new trade theory models, four factors are introduced: market size ("first nature") determined by immobile labor between countries, imperfect competition, differentiated products and increasing returns ("second nature"). The findings are about inter-industry and intra-industry specializations and as follows. Sectors concentrate around the places close to the core product markets. As the market is small, products are heterogeneous and returns are increasing; the inter-industry cluster becomes obvious. However, these models do not explain explicitly the sources of using some underlying assumptions: why does the division of large and small markets emerge? Why can similar countries have the disparity of production structures?

Thirdly, in NEG models, trade and location theoretical models are included. Besides, researchers pay more attention to micro-founded models, pecuniary externalities and “second nature” which dominates the whole economy. The economy is assumed that spatial sites endogenously result in geographically integrated patterns of economic activities. The initial distributions of activities and labors are unstable due to such features of “second nature” as input-output linkages and market-size externalities. Even a small shock can cause a large permanent effect across regions. This is consistent with a concept “home market effects”, stressed by Krugman (1991a; b), that contributes to explain the phenomenon of agglomeration. Meanwhile extreme agglomeration may trigger price differentials possibly followed by dispersion. Accordingly the spatial economy world is driven forward new market equilibrium by two opposing forces—agglomeration force and dispersion force. Agglomeration force promotes regional concentration of economic activities, while the other force distributes economic activities equally across locations.

Of course, there has been no lack of opposing voices about NEG. Some geographers, regional scientists argue that NEG doesn’t state clearly why some locational costs are under consideration whereas some others are not. Besides, the assumption of only two-region setting derived from trade theory limits the discussion of complex hierarchy in real economy and fails to address where the agglomeration happens. Moreover, full agglomeration and full dispersion stated in core-periphery model are too simple; as this modeling strategy just considers the homogeneity of agents. Last but not least, little literature in NEG field has done analysis of welfare.

2.2 Agglomeration

The history of human beings has seen that people's residences and activities are remarkably clustered in some spatial locations, which then usually are developed as communities, cities, regions and even countries. Even though the reasons for the formation of population aggregation have changed over time and space, the trend of concentration is more consistent.

To answer the questions about why people choose to agglomerate in the proximity of some geographical units is equivalent to finding out the causes of the development of cities. Apparently households and economic agents make these locational choices rationally, depending on compensating welfares of urban location in terms of costs decrease, output improvement and other utility optimization issues. From the events of formation of cities in the past, it's not hard to find out that transportation costs and internal economies of scale dominate the process of integration. Let's go through the following reasons why locations fostered along waterways are preferable. Shipping by sea is the cheapest way to faraway markets; water can also provide economical power to firms. Thus economic agents will build industrial denser settlements to share the benefits from internal economies of scale. And based on enhanced output, those agents will offer higher wages, attracting more workers to gather nearby in the end. In addition, accesses to technologies guarantee the sustainable development of industrial aggregation. More interestingly, technical advances do not only happen in existing clusters but also other new regions. The emergence of specialization furthermore accelerates the process of industrialization, as well as the development of the overall economy.

In respect to the incentive for agglomeration, the literature of economic geography has discussed many reasons for this manifestation:(1) centripetal forces encourage firms to locate close to each other;(2) internal economies of scale allow economic actors to concentrate their production in the same industry and give rise to intra-industry and international trade;(3) external economies of scale explain agglomeration in various industries and help to understand interplays of inter-industry(for instance, Marshallian externalities(Marshall, 1895) and pecuniary externalities(Krugman, 1991a;b)); (4) reducing trade costs such as transportation cost, tariffs and other indirect costs trigger concentrations; (5)competitions push the industrial integration to strengthen firms' comparative advantages.

2.2.1 Measures of agglomeration

Even though the phenomenon of industrial agglomeration has been seen frequently, the exact meaning of "industrial concentration" is hard to define. Does the high technology industry agglomerate in some areas? Do those firms or industries cluster in certain sites? Can we call this gathering of industries in such a scale as an aggregation? Without a standard to examine the degree of concentration, those questions are really ambiguous to answer. Therefore, economists, researchers and other experts produced various quantitative methodologies to describe the concept of "industrial agglomeration".

Duranton and Overman(2005) summarized the conditions of good measures of industrial agglomeration: (1) measures of the agglomeration in different industries or space must be comparable; (2) the overall pattern of concentrated activities is under consideration; (3) measures should take care of the structural variations in different industries or regions, namely to distinguish the sizes of firms; (4) different space scales do not change the unbias of measures' estimated values; (5) the results of estimated measures are statistically significant to reflect the actual distribution of economic activities. Based on these conditions, indicators for industrial agglomeration can be divided into four categories[5] that are described below. The first category satisfies the (1)-(2) properties, including location quotient (LQ)[6] and absolute Gini-coefficient, etc. The second category meets the (1)-(3) requirements and Hoover coefficient and locational Gini-coefficient are proxies. The third category is enhanced from the second one, considering the differences of plant levels. A typical method in this category is Ellison-Glaeser index of industrial concentration. The forth category compensates for large differences of basis units' scale in the third category's methods and satisfies all conditions above. This category is represented by intra-and inter-industry agglomeration indexes. All indexes in those four categories are more or less contributed to measuring industrial agglomeration; nevertheless, they have limitations on either theoretical foundations or empirical implications. I will briefly introduce one typical method per category, preparing for the choice of measure of industrial agglomeration in Section 4.1.1 theoretically and empirically.

2.2.1. (1) Location quotient in category one

Location quotient is the most widely used in urban economics and region economics and referred to the work of Hoover (1936) and Kim (1995). The expressions of LQ are as follows:

$$\mathrm{LQ_{ij}} = \frac{\mathrm{y_{ij}}/\sum_{\mathrm{j}} \mathrm{y_{ij}}}{\sum_{\mathrm{i}} \mathrm{y_{ij}}/\sum_{\mathrm{j}}\sum_{\mathrm{i}} \mathrm{y_{ij}}} = \frac{\mathrm{y_{ij}}/\sum_{\mathrm{i}} \mathrm{y_{ij}}}{\sum_{\mathrm{j}} \mathrm{y_{ij}}/\sum_{\mathrm{i}}\sum_{\mathrm{j}} \mathrm{y_{ij}}}, \qquad (1)$$

5 Duranton and Overman (2005) classified the existing indexes to concentration degree into three categories, losing sight of ways of absolute degree of integration and relative degree of integration matching the first two properties. Combining China's conditions, many researchers in China such as Wang and Wei (2006), He (2009), etc. used four taxonomies to analyze those indexes. As this paper is based on empirical case in China, I tend to use the latter way of classification.

6 Latter in this paper, location quotient (LQ) will be chosen as a measure of industrial agglomeration and explained more in Section 4.

where y_{ij} denotes the production of industry i in site j. The first expression $\frac{y_{ij}/\sum_j y_{ij}}{\sum_i y_{ij}/\sum_j \sum_i y_{ij}}$ measures the localization of industry i in the form of the share of site j in aggregate production of industry i, divided by the localization of total activity in site j. Analogously, the second one $\frac{y_{ij}/\sum_i y_{ij}}{\sum_j y_{ij}/\sum_i \sum_j y_{ij}}$ implies the specialization of site j in industry i, using the quotient of the share of industry i in the overall production of location j, relative to the specialization of location j in the whole production of industry i. (Overman, Redding and Venables, 2003)

The LQ reflects the distribution of localization and specialization. From the standpoint of this method, it is assumed that the deviation of observations follows the normal distribution. If the figure of a particular industry's LQ is larger than 1, it implies that this particular industry agglomerates. The higher is the figure of LQ, the more significant is agglomeration.

However, LQ is a static measure. Under some conditions with changing external and internal factors, LQ cannot present the integration trend of industries. Even though some figures of LQ of some sectors are lower than 1, those sectors still have the potentials for creating wealth by providing goods and services; this manifestation is called "emerging agglomeration". Besides, this approach relies more or less on the production or employment share in certain geographical sites, thus industries in large cities probably have higher LQs. Moreover, if the national average level is very low (the denominator is very small in equation (1)), a low level of agglomeration can be mistakenly illustrated as a high LQ value.

2.2.1. (2) Locational Gini-coefficient in category two

Krugman (1991a) brings forth locational Gini-coefficient which is related to the absolute Gini-coefficient; the formula of locational Gini-coefficient is

$LGC_i = \sum_{j=1}^{r}(X_j - S_{ij})^2$.   (2)

S_{ij} measures the employment in industry i in region j, relative to total national employment in industry i; X_j represents the proportion of employment in region j over the whole national employment. Locational Gini-coefficient describes the disparity of regional LQ and national LQ mean. If a locational Gini-coefficient is zero, then an industry is distributed across space as equally as the overall economy is. Conversely, if a locational Gini-coefficient closes to 1, an industry is strongly integrated in a given region. In shorts, locational Gini-coefficient is positively relative with agglomeration degree.

This measure only considers the relative degrees of agglomeration of industries in a region, but not the differences in the degrees of concentration of enterprises in various industries. To give an extreme example, if there is only one firm in an industry, the whole industry can be said to be concentrated in identical region and the Gini value of this industry is apparently high; nonetheless this case does not mean the exact industrial agglomeration.

2.2.1. (3) Ellison-Glaeser index in category three

Since the work of Ellison and Glaeser (1997), the issue about controlling for industrial lumpiness has been widely accepted in geography economics field. For example, an industry with larger firms is more likely to have a higher concentration level as they just have small numbers of firms; hence the Gini-coefficient fails to distinguish between random concentration and externalities-forced one. Ellison and Glaeser (1997) computed a new index under the assumptions that (1) industry i consists of N plants in a municipality, (2) this municipality is subdivided into r areas:

$$EG = \frac{G_i-(1-\sum_{j=1}^{r} X_j^2)H_i}{(1-\sum_{j=1}^{r} X_j^2)(1-H_i)}, \left(G_i = \sum_{j=1}^{r}\left(X_j - S_{ij}\right)^2, H_i = \sum_{k=1}^{N} Z_k^2\right). \qquad (3)$$

G_i and H_i represent locational Gini-coefficient addressed above and Hirschman-Herfindahl index (HHI) respectively. Z_k is the ratio of employment in plant k to the total employment in industry i, reflecting the distribution of plant scale. The higher is Z_k (up to 1), the stronger is monopolization in market. In this method, if geographical distribution of employment in plants are random, the value of EG is 0; otherwise, the positive value of EG implies regional integration in the industry.

From an empirical point of view, the Ellison-Glaeser index has its limitations. It is too sensitive to the quality of data, so that the expected values of EG in an identical industry over years fluctuate dramatically, deviating from the actual conditions. On the other hand, Ellison and Glaeser (1997) distinguished degree of integration by empirical results. They said EG<0.02 as low level of aggregation and EG>0.05 as high level. It lacks of adequate evidences to decide the boundary of agglomeration from the random one.

2.2.1. (4) Duranton-Overman index in category four

Measures in category three are based on municipal units with significant different scales. Those measures only describe the degree of industrial agglomeration in single spatial dimension and easily end up with illusions of industrial spatial patterns. Therefore, in order to clarify the pattern of economic activities, different methods appeared by means of describing industrial structures in various space scales. These methods do not describe the pattern of economic activities only in the scale of human-defined municipal units. Duranton and Overman (2005) employed the nonparametric regression model with more accurate locational data of plants and computed a more generally applied concentration index. This index (Duranton-Overman index) was named after Duranton and Overman and it is based on inter-distances:

$$i_{\mathrm{DO}}(r) = \frac{\sum_i \sum_{j \neq i} h^{-1} w(\frac{r - \| x_i - x_j \| m_i m_j}{h})}{\sum_i \sum_{j \neq i} m_i m_j)}, \qquad (4)^{7}$$

where i, j are points of locations; m is a “mark space” and $m_i m_j$ are supposed to be random variables in m; h is the bandwidth, (Silverman, 1986 cited in Duranton and Overman, 2005) ; w is a boundary correction factor.

This Duranton-Overman index eliminates some limitations of previous methods; however, the precise locational data required in this method is hard to obtain from real life.

2.2.2 The correlation between agglomeration and international trade

The correlation between agglomeration and international trade was disclosed in literature of new trade theory and NEG; the NEG literature has emphasized a non-monotonic interaction of industrial agglomeration and trade costs. Krugman (1991a) was the first economist who conducted the research of close linkages of industrial integration and international trade factors. He found that trade of products took the place of trade of factors indirectly. Regardless of the initial distribution of production factors, trade activities could integrate several productions into certain industrial areas and be followed by the formation of industrial cluster. Rauch (1991) added geographic factors to an international trade model with transportation costs for goods based on cities. He generated a positive relationship between the trade volume and home country comparative advantage. He also emphasized the volume of international trading and the

[7] For more information about the D-O index, please refer to the work of Duranton and Overman (2005).

geographic advantage which is one of the causes of industrial concentration. Haaparanta (1998) proved that free trade can lead to regional integration of economic activities within one country. His work has provided feasible theoretical foundations for this paper to discuss the regional industrial agglomeration under influences of international trade in a country.

Following the development of international trade, globalization and market integration of goods also exert crucial influences on location of economic activities. Agglomeration and international trade interact with each other.

On one hand, industrial agglomeration improves competitiveness of industries and promotes international trade. The degree of export of goods is significantly dependent on the international competitiveness of goods. Porter (1990) concluded the impact of industrial integration on competitive advantages of industries from three aspects. Firstly, industrial concentration enhances the productivity of home-based firms within integrated area by inputs of factors and complementation of technology and knowledge. Secondly, integration helps industrial innovation and speeds up firms' rate of innovation. Thirdly, cluster expands the scale of firms and pushes industrial derivations.

On the other hand, global trade boosts spatial concentration. Home market demand conditions affect an industry's ability to compete in the whole world. This competitive ability causes geographic aggregation that improves competitive advantages of industry and nations (Porter, 1990; Krugman, 1991a; b). The expansion of free trade results in the growing dependence on foreign trade, usually lowering the tariff. Low tariff cuts transportation costs, attracts foreign investments and makes it easier for firms to get access to resources internationally, promoting agglomeration across space. Besides, export of goods is the extension of home market demands which encourage spatial aggregation. Generally speaking, policies promoting exports benefit industrial concentration by providing lower trade costs and larger markets at home and abroad.

2.2.3 The correlation between agglomeration and internal economies of scale

The literature of new trade theory and economic geography reaches a consensus of the relationship between agglomeration and scale economies: a low transaction cost between geographic organizations produces industrial concentration of activities. The study of Krugman (1991a) showed that economies of scale from both internal and external perspectives affect geographic concentration of economic activities, regional specialization and global industrial trades. Krugman (1991b) also explained that clustering of activities is derived from internal increasing returns to scale and transportation costs, based on predecessors' work (Henderson, 1974 cited in Krugman, 1991b; Papageorgiou and Thisse, 1985 cited in Krugman, 1991b and Fujita, 1988 cited in Krugman, 1991b). Besides, Brülhart and Torstensson (1996) found from their analysis that industrial agglomeration in central region monotonically increases the degree of scale economies. They also pointed out that plant-internal scale economies have a positive relationship with geographic aggregation and upon this they got a prediction that scale-intensive economic activities will integrate close to markets with better accesses. Moreover, an empirical model was built up to find out the relation of changes in industrial locations and economic structures by comparing EU and USA data. In this model, the coefficient of market potential and economies of scale reflect that higher economies of scale lead industries to integrate in core locations (Midelfart, Overman, Redding and Venables, 2000). In shorts, internal economies of scale create incentives for clustering firms' activities.

2.2.4 The correlation between agglomeration and home market effects

In general, studying home market effects is the first step to understand NEG models. According to Helpman and Krugman (1985 cited in Ottaviano and Puga, 1997), with transportation costs and imperfect competitions, industries tend to locate near larger markets and export to smaller ones. This is the result of home market effects that have the characteristics of "gravitational forces" which lead small changes in market size to large spatial heterogeneities. That is why "home market effects" is the core of theories and models about agglomeration. Krugman (1991a) formally brought up this concept as one of the determinants of industrial agglomeration with mobile labors. Considering the market access and final price, firms prefer to locate in areas close to demand and supply. This pattern of demand- and supply-driven specialization reflects that increasing scale economies tend to get access to good market disproportionately. In NEG models, individual's location preference is influenced by changes in expenditure; furthermore, differences of location preference can change the expenditure. Hence a growth of expenditure typically leads to a higher growth of production, and with better access to market, a region enjoys a higher factor price in NEG models. For example, if there are 8 regions, 70% of the total expenditure is equally shared by 7 regions and the rest is taken by the last one region; this larger region can supply over 30% of demands so that firms would benefit from the large market. While considering market access to intermediate production, upstream and downstream firms also have incentives to concentrate geographically to cut intermediate costs. Krugman and Venables (1995; 1996 cited in Ottaviano and Puga, 1997) pointed out that agglomeration can result from vertical linkages with immobile labors. Since even an economic agent locates its firm in a certain place, this firm can expand its upstream and downstream markets in the absence of labor mobility from other space.[8]

All in all, the "home market effects" implies a linkage between the market size and the geographic clustering of activities. However, home market effects do not give an explanation about why a small change can generate a large permanent effect on sites.

3. Methodology

In this section, I am going to specify the main econometric techniques and principles applied in this paper. This section includes methods of testing the stationarity of data and principles with regard to panel data models and how to choose models in panel data. Besides, two widely-used instrumental variables estimation ways—2SLS and GMM will be introduced.

3.1 Stationarity Testing

Granger and Newbolt (1974) showed that when running ordinary least squares (OLS) estimations with nonstationary time series, the estimated values of regression coefficients will lose the best linear unbias and the corresponding results of T-test will be useless. Li (2000) also found that nonstationary time series usually contain a mutual trend; nevertheless these series themselves may not have real relations. Consequently regression analysis for those series leads to spurious regressions, even with higher R^2 values. "Stationarity" means that the

---

[8] This is the foundation to Krugman and Venables (1995 cited in Ottaviano and Puga, 1997) where upstream and downstream sectors were simplified into one sector and to Venables (1996 cited in Ottaviano and Puga, 1997) where upstream and downstream sectors were analyzed separately.

features of random process generating data of series keep constant through time period. It means that after a time series is gotten rid of the invariable mean (namely intercept) and stochastic trend, the remaining series has the features of zero-mean and the same variance, shown as follows:

$$E(y_t) = \mu$$

$$Var\ (y_t) = \sigma^2$$

$$cov(y_t, y_{t-s}) = cov(y_{t-j,} y_{t-j-s}) = \gamma_s, \forall t, j, s \in I,$$

where μ, σ and γ_s are constants and t shows time periods. [9]

To avoid spurious regression, we will test the stationarity of panel series in the most general way —unit root testing. Levin and Lin (1993 cited in Bai, 2008) built up the early version of unit root testing for panel data on the assumption that the limited distributions of these estimators are Gaussian distributions. Levin (2002 cited in Bai, 2008) further improved this testing method and proposed the LLC-test which satisfies unit root testing for panel series in other conditions. For instance, series have different intercepts or trends, the heteroscedasticity or a high-order serial correlation. The observed samples can be middle sizes, namely the numbers of time periods and sections vary from 25 to 250 and 10 to 250 respectively. Im, Peseran and Shin (1997 cited in Bai, 2008) suggested using the IPS-test. But Breitung (2000 cited in Bai, 2008) argued that the assumption of IPS is too sensitive and put forward a new method—Breitung test—to test the unit root for panel data. Maddala and Wu (1999 cited in Bai, 2008) expanded DF-test to ADF-Fisher-test and showed one more method of testing unit roots, PP-Fisher-test. Based on the KPSS test of a single time series, Hadri (2000 cited in Bai, 2008) brought a unit root test for panel series under the

[9] Exactly they are the conditions for weak stationarity; in this paper stationarity generally refers to the weak one because the weak stationarity is sufficient.(Granger and Newbolt, 1974)

null hypothesis that individual series is stationary. Generally speaking, LLC-test, Breitung-test and Hadri-test are under the assumption of the same unit root; while IPS-test, ADF-Fisher-test and PP-Fisher-test apply to different unit roots. Particularly the null hypotheses of these tests except Hadri-test are nonstationary series. Besides, there are three patterns in the unit root testing: test equations with intercept and trend, test equations only with intercept and test equations without intercept or trend. The testing pattern of unit root testing can be inferred from the graphs of panel series that reflect the structures of variables roughly. Moreover, unit root testing starts from level series as usual; if the result shows a unit root exists, we test the first-order difference series. If the unit root still exists in first-order difference series, a higher-order differencing should be used until we get the consequence of stationary series. Accordingly if the $\{y_t\}$ series is stationary, it is called to be integrated of order zero and denoted by $\{y_t\} \sim I(0)$. The $\{y_t\}$ series becomes stationary by differencing once, then this series is said to be integrated of first order and denoted by $\{y_t\} \sim I(1)$. The $\{y_t\}$ series is made stationary by differencing d times at least and it is denoted by $\{y_t\} \sim I(d)$.

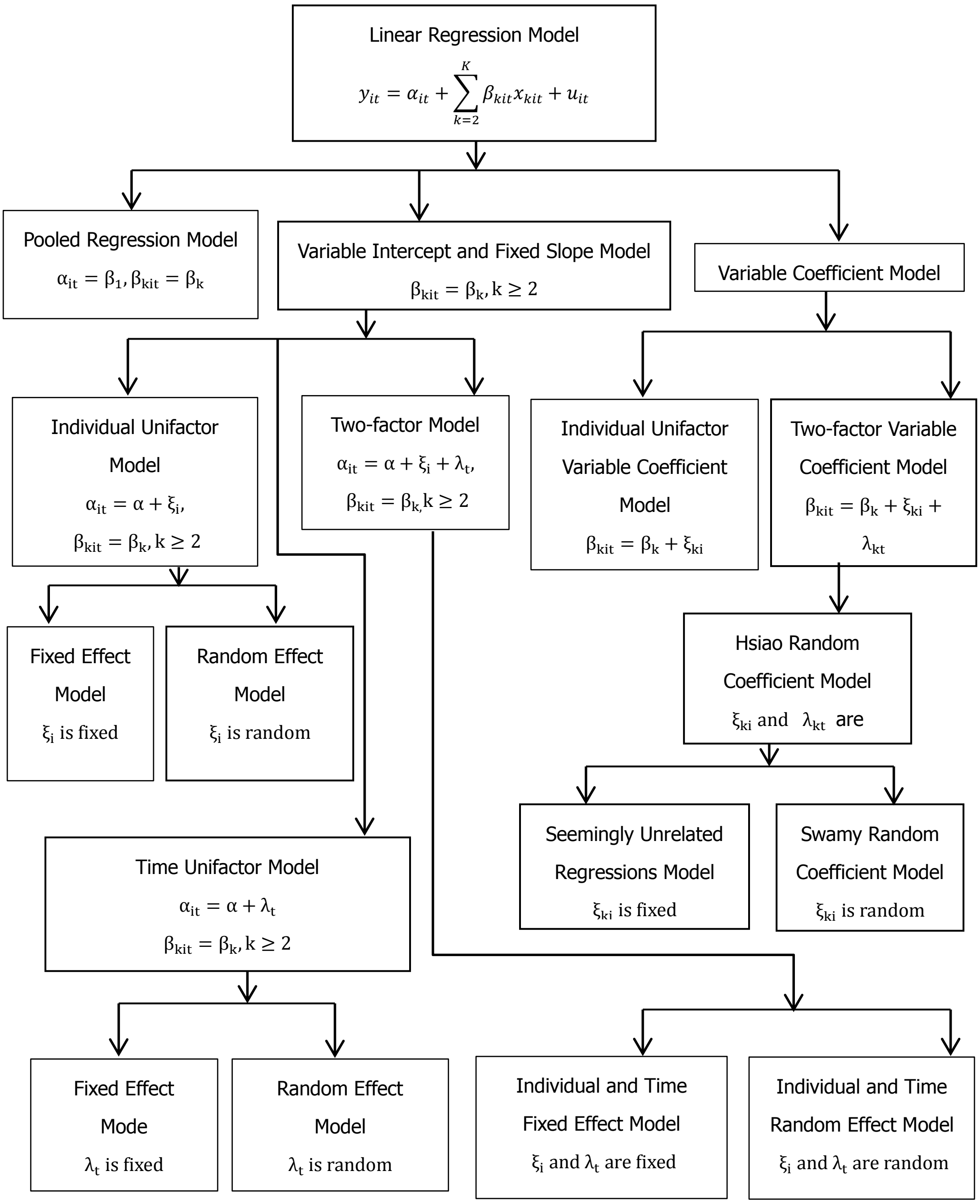
3.2 Panel data model choice

Panel data models can be divided into two categories: static regression models and dynamic regression models. In static regression models, changes of a dependent variable rely on effects of independent variables during the current period. While in dynamic regression models, changes in both current period and previous periods are under consideration.

3.2.1 Static Regression Model

Bai (2008) summarized the taxonomies and model-setting of panel data static regression models, shown as the following figure:

Figure 1: Panel Data Linear Regression Model System



Source: Bai (2008, Chapter 2, p.14).

Generally these three patterns of models in the above-mentioned system are widely used in recent studies: pooled regression model, fixed-effects regression model and random-effects regression model. The pooled regression model with OLS estimation is better when individuals do not differ from each other significantly in terms of time and section. The fixed-effects regression model is suitable for diverse sections or time series with different intercepts. The random-effects regression model is employed in the case that average effects of sectional random errors and time random errors exist in the intercepts of fixed-effects regression model and these two errors follow a normal distribution (Baltagi, 2008; Bai, 2008).

From a methodological point of view, fixed-effects-test (F-test) is used to decide between the pooled regression model and fixed-effects regression model. The null hypothesis H_0 is that all intercepts from various individual models are the same,$\alpha_{it} = \beta_1$, inferring panel data that is poolable; the alternative hypothesis H_1 is that intercepts from different individual models are diverse,$\alpha_{it} \neq \beta_1$, inferring fixed-effects regression model. In addition, Hausman-test (H-tests) (Hausman, 1978 cited in Bai, 2008) helps to choose fixed-effects or random-effects models.

3.2.2 Dynamic Regression Model

Compared with static panel data regression models, the dynamic ones introduce lagged explained variable into the right hand side of the equation of static models, reflecting dynamic lag effects. The general form of dynamic panel data regression models is $y_{it} = \gamma y_{i,t-1} + \sum_{k=2}^{K} \beta_{kit} x_{kit} + u_{it}$ (8), where i=1,2,...,N, t=1,2,…,T; $u_{it=}\mu_i + \nu_{it}$, $\mu_i \sim IID(0, \sigma_\mu^2)$ and $\nu_{it} \sim IID(0, \sigma_\nu^2)$. It is noticeable that lagged explained variable is correlated with error term even though ν_{it} is not auto correlated. In this case, least squares dummy variables method (LSDV) and feasible generalized least squares (FGLS) method are not so effective; instead, instrumental variable (IV) method and generalized method of moment (GMM) are usually used in estimations of dynamic panel data models, taking the place of OLS. Anderson and Hsiao (1981 cited in Bai, 2008) suggested an instrumental variable method with two procedures: firstly doing first difference of the equation (8) to get rid of the fixed effects:

$$y_{it} - y_{i,t-1} = \gamma(y_{i,t-1} - y_{i,t-2}) + \sum_{k=2}^{K} \beta_{kit}(x_{kit} - x_{ki,t-1}) + u_{it} - u_{i,t-1}$$

$\Delta y_{it} = \gamma \Delta y_{i,t-1} + \sum_{k=2}^{K} \beta_{kit} \Delta x_{kit} + \Delta u_{it}$.

Secondly using $y_{i,t-2}$ or $(y_{i,t-2} - y_{i,t-3})$ as an IV for $\Delta y_{i,t-1}$ to solve the endogeneity problem. These IVs are uncorrelated with the disturbance difference Δu_{it}. Then Hansen (1982 cited in Bai, 2008) pointed out another approach, GMM, which does not require so many assumptions or given distribution of errors. This approach uses an estimate of parameters' variance as weights and minimizes a chi-squared statistic when the estimator is asymptotically consistent with an estimable covariance.[10] Based on Hansen's achievement, Arellano and Bond (1991 cited in Bai, 2008) improved the GMM estimator by employing available lagged values of dependent and independent variables as instruments, namely two-step GMM estimate.[11] These one-step and two-step GMM estimators can be produced by EViews directly. Latter Arellano and Bover (1995 cited in Bai, 2008) proposed a new GMM estimator considering exogenous variables and Blundell & Bond (1998 cited in Bai, 2008) showed a consistent GMM estimator by releasing one restriction in previous GMM-proving procedures to improve the efficiency of Arellano and Bond's GMM estimators.[12]

---

10 About the detailed expression function of A-H estimator, please refer to Bai (2008) and other advanced econometrics books.

11 Interested in the completed content about what are these two GMM estimators and how to get those, readers can read more in Bai (2008).

12 More information can be referred to Bai (2008).

3.3 2SLS

In practices it is often to notice endogeneity problems caused by model specification bias (e.g. omitted variables), estimation errors and interactions between explanatory and explained variables; and such problems will yield inconsistent and biased OLS estimators. Proxy and instrumental variable approaches are applied in empirical analysis; while the latter approach is better because the proxy method causes heteroscedasticity in the case of interaction and requires stricter constraint conditions and strong awareness of omitted variables. In the following estimations, we try to get rid of the endogeneity by the instrumental variable (IV) methods that ask for instruments which are highly correlative with the endogenous variables but uncorrelated with error terms. Two Stage Least Squares (2SLS) estimation is one of the IV estimations and going to be used in this study. The model will run two regressions: first, we regress the endogenous variables on the instruments by using OLS and get fitted value of the endogenous variables. After that, we put these fitted values into the right hand side of the original model, instead of the endogenous variables and do OLS estimation on the explained variable. One issue in this process is how to evaluate the validity of IV.[13] As the IV should not correlate with other exogenous variables, one common way is to check the IV on the first stage of the 2SLS; if the IV is not significant, it is not relevant. Besides we can evaluate the IV by the R^2 and F-statistics from 2SLS to see if the estimated coefficients for other explanatory variables do not vary substantially from those in the standard OLS estimation.

3.4 GMM

[13] Instrument validity is a term used for instruments being uncorrelated with error terms. This can only be tested if there are several instruments.

GMM approach has been introduced in section 3.2.2 and it is of great help to control the endogeneity and heteroscedasticity. It should be clearly understood that the GMM estimator we are going to use in EViews 6.0 is Arellano-Bond's. The variables' transformation of first difference is applied to remove cross-section fixed effects from our dynamic panel data model; 2-step interation and White-period weighting matrix are chosen to compute estimations in the light of the properties of this empirical analysis.

4. Measures, Variables and Data

We will first address the concerned factors in our analysis and then choose one indicator to measure each concerned factor. Then, we will specify their corresponding forms introduced into the model as variables and finally give a description of data for later regression.

4.1 Measures

As can be seen in section 2, internal economies of scale and home market effects stand out above the rest of causes for the agglomeration on the following grounds: (1) they are widely accepted by theories. New trade theory stresses the importance of the impact of internal economies of scale on industrial concentration, and many articles support this point by presenting the positive effect of scale economies on industries' location preference towards centers.[14] Associated with internal economies of scale, home market effects are also generally considered as determinants of regional industrial integration by Krugman (1980 cited in Ottaviano and Puga, 1997). Moreover Krugman (1991a; b) realized that a circular causation yields from the interaction between scale

[14] For more examples and information, please refer to Section 2.2.3.

economies and home market effects. Thus, industries produce goods in concentrated proximities to large markets and then market scales are turning larger in areas with industrial aggregation.[15] (2) From a practical point of view, they play vital roles in agglomeration in the case of Guangdong. When doing empirical analysis of regional industrial clustering in Guangdong Province, researchers in China have focused on firms' scale economies as well as market potentials. In one of the most profound empirical studies about this topic, Yin and Tang(2007) threw light on their findings in Guangdong Province that the stronger the internal economies of scale, the more likely firms locate close to each other; and that micro-economic foundations aggregate spatially with convenient access to local markets.

Accordingly, to probe deeper and solve the research question, this paper considers internal economies of scale and home market effects as controlling factors and examines their impacts on locational concentration; because these two factors are salient characteristics of industrial agglomeration in Guangdong apart from international trade. In short, international trade, internal economies of scale and home market effects are used to analyze industrial agglomeration in this paper.

4.1.1 A measure of industrial agglomeration

"Agglomeration is typically used to refer to the degree to which economy activity as a whole is geographically concentrated."(Redding, 2009, p. 14) We prefer to employ location quotient (LQ) as a measure of the agglomeration degree in industry. Even though LQ has its own weakness, this approach is more useful when this paper focuses more on the agglomeration per se than its importance.

[15] The issues of home market effects as causes of industrial agglomeration can be referred to Section 2.2.4.

Besides, LQ overcomes some shortcomings of other measures of industrial concentration such as too complicated computation and limited assumptions. Most importantly, the data required in LQ method is more likely to be available even with limited accesses to province-level data in China.[16]

The LQ equation (1) shown in section 2.2.1 is adjusted to the empirical case of manufacturing industry in Guangdong Province, China as follows:

$$\mathrm{LQ_{iGuangdong}} = \frac{\mathrm{y_{iGuangdong}} / \sum_i \mathrm{y_{iGuangdong}}}{\mathrm{y_{iChina}} / \sum_i \mathrm{y_{iChina}}}, \qquad (1)^*$$

$\mathrm{LQ_{iGuangdong}}$-- The location quotient of sector i in the manufacturing industry in Guangdong Province;

$\mathrm{y_{iGuangdong}}$-- The production of sector i in the manufacturing industry in Guangdong Province;

$\mathrm{y_{iChina}}$-- the production of sector i in the manufacturing industry in China.

The numerator in expression (1)* indicates the sector i's share of the GDP in Guangdong Province, while the denominator means the sector i's share of the GDP in China. $\mathrm{LQ_{iGuangdong}}$ can reflect the disparity in the average level of production between Guangdong Province and the whole country, with the purpose of assessing the geographical structure of industrial sectors. When $\mathrm{LQ_{iGuangdong}}>1$, agglomeration exists in some sectors; whereas if a sector in Guangdong Province is not localized or specialized but distributed in accordance with productions in China, $\mathrm{LQ_{iGuangdong}}$ is equal to 1.

Data on the value or volume of aggregate manufacturing production is not available in China; instead, we insert sector i's GDP of the manufacturing industry

---

[16] A helpful discussion of methodologies of measuring agglomeration degree (including the original expression of LQ) can be found in the former part "measures of agglomeration" in section 2 in this paper.

in Guangdong Province and in China into $y_{iGuangdong}$and y_{iChina} respectively and Guangdong Province's GDP into $\sum_i y_{iGuangdong}$ as well as China's GDP into $\sum_i y_{iChina}$.

4.1.2 A measure of international trade

Referred to theoretical analysis in Rauch (1991) and Haaparanta (1998) and empirical exercises in You and Li (2010), international trade factor is tractable for evaluating international trade conditions to serve the objective of this paper. However, in China, taxonomies of trading goods are subject to criteria of Customs and different from classifications of industrial products which are based on standards of industries. Therefore, it is not probable to approach import and export data across sectors in manufacturing industry anywhere. To estimate the real annual import and export value for sector i, the equation of international trade is specified as:

$$TR_{iGuangdong} = TR_{Guangdong} * \frac{y_{iGuangdong}}{y_{Guangdong}}, \qquad (5)$$

$TR_{iGuangdong}$ —international trade factor of sector i in manufacturing industry in Guangdong Province;

$TR_{Guangdong}$—total import and export value in Guangdong Province.

That is to say, $TR_{iGuangdong}$ is denoted by the product amount of import and export in Guangdong Province and the share of Guangdong's sector i in the total GDP of Guangdong.

Consistent with the $LQ_{iGuangdong}$ method, equation (5) is further modified as

$$TR_i = \frac{TR_{iGuangdong}}{TR_{iChina}}, \qquad (5)^*$$

where $TR_{iChina} = TR_{China} * \frac{y_{iChina}}{y_{China}}$, in the same logic of equation (5).

Conclusively we will use TR_i in the equation(5)*, the ratio of trade conditions of sector i in Guangdong to those in China, to evaluate the impact of international trade on industrial concentration in manufacturing industry in Guangdong Province. The larger the TR_i, the higher free trade level the sector i.

4.1.3 A measure of internal economies of scale

"Internal economies of scale" is a measure of minimum efficient scale within micro-economic foundations. According to pervious theoretical and empirical studies, scale economies interplay with many other factors such as trade costs and market potentials. Scale economies are proved to have strong relations with concentrated industrial locations: industries with greater scale economies tend to agglomerate spatially.[17]

In terms of actual operation, firms' scale is the optimal choice to measure internal economies of scale. The study of Pratten(1988) explained that scale economies are related to the variables "products and productions runs" and "size of the establishment", providing an empirical support for this paper to choose this indicator to reflect scale economies. In the same way of evaluating agglomeration and international trade, this paper will take the value of firms' scale in sector i in Guangdong relative to that in China. The purpose is to capture the idea of how internal economies of scale affect industrial spatial integrations. This measure of internal economies of scale can be formulized:

$$\mathrm{IS_i} = \frac{y_{\mathrm{iGuangdong}}/n_{\mathrm{iGuangdong}}}{y_{\mathrm{iChina}}/n_{\mathrm{iChina}}}, \qquad (6)$$

where $n_{\mathrm{iGuangdong}}$ and n_{iChina} represent the total amount of firms in sector i in Guangdong and China respectively; the numerator suggests the average firm

[17] Studies about scale economics in geographical economics field are mentioned in the literature review part in this paper; works of Brulhart and Torstensson (1996) and Midelfart, Overman, Redding and Venables (2000) tested the importance of scale economics to industrial integration in consideration of trade costs and market potentials respectively.

scale in sector i in Guangdong, while the dominator shows the average firm scale in sector i all over China.

4.1.4 A measure of home market effects

“Home market effects” is critical in NEG theory field. It suggests a “backward linkage” of market and trade so that economic agents have incentives to locate close to larger markets for producing goods. Some scholars suspected the constancy of home market effects; for instance, Davis (1998 cited in Redding, 2009) argued that home market effects can evanish in the conditions where a sector produces homogeneous goods without fixed costs while the other sector yields differentiated products with fixed costs. However, other experts such as Krugman, Venables (1999 cited in Hanson and Xiang, 2002), Holmes and Stevens (2002 cited in Hanson and Xiang, 2002) supported the universality of home market effects-- particularly in terms of promoting regional integration-- by turning down Davis’ argument [18] and further demonstrations respectively. Apparently and reasonably, there are varying methods of identifying home market effects, e.g. interaction between supply and demand, income elasticity of exports, variables weighted by transportation costs and so on.[19]

In this paper I want to simplify the measure of home market effects by the ratio of GDP per capita in Guangdong to GDP per capita in China, according to the methodologies of estimating home market effects used by Hanson and Xiang

---

[18] Krugman and Venables(1999 cited in Hanson and Xiang, 2002) proved the existence of home market effects in the conditions that the homogeneous-goods sector has transportation costs or the differentiated-goods sector has not fixed costs, pointing against the findings in Davis(1999 cited in Hanson and Xiang, 2002).

[19] To give some examples, it can be referred to Krugman(1980 cited in Ottaviano and Puga, 1997) and Davis and Weinstein(1999 cited in Hanson and Xiang, 2002) that an interaction between supply and demand; the income elasticity of exports can be found in studies of Feenstra, Markusen and Rose(1998 cited in Hanson and Xiang, 2002) and Rauch(1999 cited in Hanson and Xiang, 2002); transportation costs were added into the gravity models to measure home market effects in Hanson and Xiang(2002).

(2002)[20] and You and Li (2010). The indicator for home market effects is showed below:

$$\mathrm{HM} = \frac{\mathrm{AGDP}_{\mathrm{Guangdong}}}{\mathrm{AGDP}_{\mathrm{China}}}, \qquad (7)$$

where $\mathrm{AGDP}_{\mathrm{Guangdong}}$ and $\mathrm{AGDP}_{\mathrm{China}}$ denote GDP per capita in Guangdong and in China respectively. When the home market effect becomes stronger, the value of HM increases, implying that a region is becoming more focused on local market.

4.2 Variables Specification

The aim of this paper is to find out whether international trade is a factor creating industrial agglomeration in Guangdong Province of China. However, when the above-mentioned measures of industrial agglomeration and international trade are directly taken into the analysis as variables, a problem of endogeneity or simultaneity may come up. The more activities are located in Guangdong Province, the more likely is international trade to become larger; namely both agglomeration and trade have causal effects on each other. In order to solve this problem, transformations of both measures of agglomeration and international trade are used as variables in the models. The principle of these transformations is to use the variations of data between the observed year and the initial year (Year 1985) when the open door policy was encouraged by Chinese government. Therefore the variables in analysis are as equations (8) and (9)[21], instead of measures of industrial agglomeration ($\mathrm{LQ}_{\mathrm{iGuangdong}}$) and international trade (TR_{i}) respectively:

---

[20] In "The Home Market Effect and Bilateral Trade Patterns" Hanson and Xiang (2002) used "relative GDP to capture relative market size" in their model; I am going to use the similar way of identifying home market effects in the model of this paper.

[21] The computations of the variables follow the equations (1)* and (5)* in section 4.1.1 and 4.1.2 and the processed data of variables can be found in the appendix.

$\mathrm{LQCHANGE}_{i,t} = \mathrm{LQ}_{\mathrm{iGuangdong,t}} - \mathrm{LQ}_{\mathrm{iGuangdong,1985}}$ (8);

$\mathrm{TRCHANGE}_{it} = \mathrm{TR}_{\mathrm{iGuangdong,t}} - \mathrm{TR}_{\mathrm{iGuangdong,1985}}$(9).

As is introduced earlier in section 1.1, Guangdong Province pioneered the regional economic development with the benefits of open door policies in China. These policies encouraged establishing export processing zones or special economic zones and many of them are located in Guangdong Province. Firms in those zones were requested to export most of their productions and allowed to import resources without paying tariffs. That is to say the expansion of international trade in Guangdong Province where export processing zones and special economic zones were built up was initially induced by exogenous policies and followed by regional economic clusters. Hence we can consider that the changes in agglomeration after 1985 are caused by the changes in foreign trades; the causality between agglomeration and trade is not simultaneous.[22] The variables $\mathrm{LQCHANGE}_{i,t}$ and $\mathrm{TRCHANGE}_{i,t}$ can eliminate endogeneity or simultaneity problems.

In line with $\mathrm{LQCHANGE}_{i,t}$ and $\mathrm{TRCHANGE}_{i,t}$, the measures IS_i and HM are transformed into the logarithmic forms of $\ln(\mathrm{IS}_{it})$ and $\ln(\mathrm{HM}_t)$ respectively. Because the logarithmic transformation can reflect the changes of data, keep the linear relationship of dependent and independent variables and remove heteroscedasticity from original data.

In conclusion, under the theoretical assumptions, the analyzed factors, their corresponding variables and estimated controlling variables' influences on agglomeration are shown in the table 1 below:

[22] It means that the industrial agglomeration is caused by international trade while international trade is not resulted from agglomeration after 1985.

Table 1: Variables Specification

Factors	Variables	Direction of Influence (estimated)
Industrial Agglomeration	Change in Location Quotient ($LQCHANGE_{i,t}$)	NA
International Trade	Changes in Ratio of international Trade ($TRCHANGE_{i,t}$)	+
Internal Economies of Scale	$LnIS_{i,t}$	+
Home Market Effects	$LnHM_{i,t}$	+

Source: Author's predictions, based on previous literature.

4.3 Data source and testing

4.3.1 Data Source

This paper is going to use the panel data in Guangdong Province in China from 2001 to 2009 to study the impact of international trade on regional agglomeration in manufacturing industry. From the perspective of industrial agglomeration, this paper aims at manufacturing industry but not mining industry, supply of energy (electric and heat) or supplies of gas and water because of the reasons below. Firstly, mining industry is strictly limited by the distribution of natural resources; secondly, supplies of energy, gas or water highly rely on home market demand and most of their products are not tradable. Contrasting with these industries, manufacturing industry provides more feasible locational options for plants and happens on varying patterns of concentration; this diversity of agglomeration in manufacturing industry is worth further studying.

Out of all 33 manufacturing sectors, such sectors that have strong performances in trading are chosen. From the statistics in appendix 2, the following sectors have relatively high degrees of internationalization and should be analyzed in this paper: (1) Manufacture of Textile Garments, Footwear and Headgear; (2) Manufacture of Leather, Fur, Feather, Down and Related Products;(3) Manufacture of Timber Processing, Bamboo, Cane, Palm Fiber and Straw Products;(4) Manufacture of Furniture;(5) Manufacture of Printing and Record Medium Reproduction; (6) Manufacture of Cultural, Educational and Sports Articles;(7) Manufacture of Plastic Products;(8) Manufacture of Nonmetal Mineral Products;(9) Manufacture of Metal Products;(10) Manufacture of Electrical Machinery and Equipment;(11) Manufacture of Communication Equipment, Computers and Other Electronic Equipment.

The annual data for the estimation mainly comes from China Statistical Yearbook and Guangdong Statistical Yearbook during the ten-year-period of 2001-2010. The data covers GDP in certain sectors, the import and export values, the amount of plants in a particular sector and GDP per capita at national and provincial levels. Since 2004, the classification of industries in Guangdong Statistical Yearbook has been updated, following the rules of a new system (GB/T4754-2002) instead of the old system (GB/T4754-94). Thereby, this paper chooses sectors that are not affected by the modification in classification criterion. Regarding the general information of industrial agglomeration and international trade in Guangdong Province in section 1, that data is from the official website of Guangdong Provincial Department of Foreign Trade and Economic Cooperation.

4.3.2 Data testing

In section 3 we have discussed the gains of testing the data stationarity before running regressions of the model. Unit root tests, therefore, are applied to the selected panel data. Based on the EViews 6.0 testing results in the tables below, the data for variables used in this study can be said to be stationary.[23]

Table 2: Unit Root Test for Variable LQCHANGE

Panel unit root test: Summary

Series: LQCHANGE

Date: 04/12/12 Time: 09:34

Sample: 2000 2009

Exogenous variables: Individual effects

Automatic selection of maximum lags

Automatic selection of lags based on SIC: 0 to 1

Newey-West bandwidth selection using Bartlett kernel

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-4.94069	0.0000	11	97
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	0.00050	0.5002	11	97
ADF - Fisher Chi-square	18.2840	0.6890	11	97
PP - Fisher Chi-square	14.0115	0.9011	11	99

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

[23] The level variables with individual intercepts in the test equations are tested by EViews 6.0. In general, if the probabilities for all testing methods are smaller than 5%, the null hypothesis of unit root is rejected and the data is I (0).

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Table 3: Unit Root Test for Variable TRCHANGE

Panel unit root test: Summary

Series: TRCHANGE

Date: 04/12/12 Time: 09:35

Sample: 2000 2009

Exogenous variables: Individual effects

Automatic selection of maximum lags

Automatic selection of lags based on SIC: 0 to 1

Newey-West bandwidth selection using Bartlett kernel

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-3.29900	0.0005	11	97
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	0.57374	0.7169	11	97
ADF - Fisher Chi-square	13.2860	0.9249	11	97
PP - Fisher Chi-square	14.7293	0.8736	11	99

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Table 4: Unit Root Test for Variable LnIS

Panel unit root test: Summary

Series: LNIS

Date: 04/12/12 Time: 09:36

Sample: 2000 2009

Exogenous variables: Individual effects

Automatic selection of maximum lags

Automatic selection of lags based on SIC: 0

Newey-West bandwidth selection using Bartlett kernel

Balanced observations for each test

Cross-

Method	Statistic	Prob.**	sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-9.35708	0.0000	11	99
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-4.22139	0.0000	11	99
ADF - Fisher Chi-square	58.1971	0.0000	11	99
PP - Fisher Chi-square	64.9663	0.0000	11	99

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Table 5: Unit Root Test for Variable LnHM

Panel unit root test: Summary

Series: LNHM

Date: 04/12/12 Time: 09:36

Sample: 2000 2009

Exogenous variables: Individual effects

Automatic selection of maximum lags

Automatic selection of lags based on SIC: 1

Newey-West bandwidth selection using Bartlett kernel

Balanced observations for each test

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-8.90368	0.0000	11	88
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-4.13739	0.0000	11	88
ADF - Fisher Chi-square	62.2455	0.0000	11	88
PP - Fisher Chi-square	9.42640	0.9908	11	99

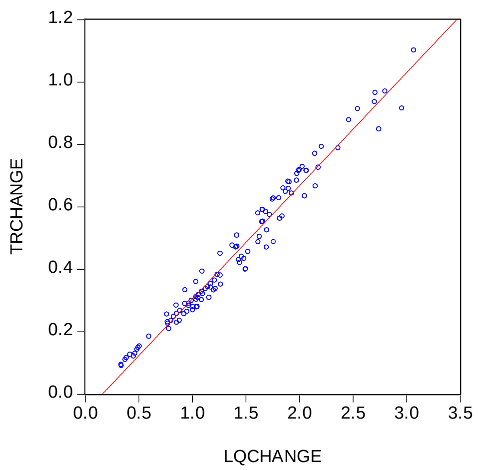
** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

In table 2 and 3, only the “Levin, Lin & Chu t” results point out the stationarity of variables' data but other tests fail. In this particular study, the sample is small, only covering 11 sections over 10 years; it is reasonable for us to make an assumption of common unit root in panel data and trust the testing results of "Levin, Lin & Chu t" method. LnIS and LnHM pass the stationarity testing similarly.

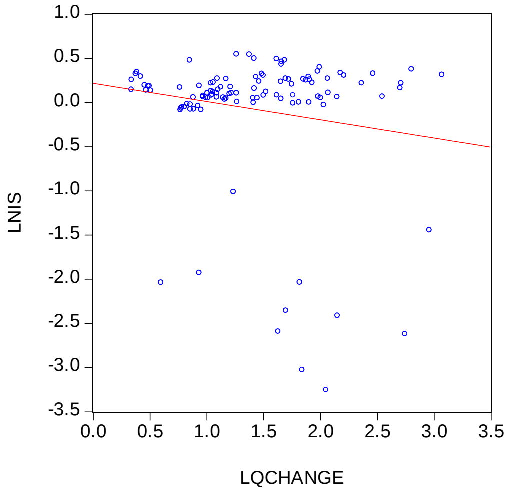
Next step, the scatter graphs of the explained variable LQCHANGE and explanatory variables TRCHANGE, LnIS and LnHM are made respectively to decide their relationships and the suitable setting of the analysis model. The scatter graphs are shown as follows:

Figure 2: Panel Data Scatter Graph between LQCHANGE and TRCHANGE



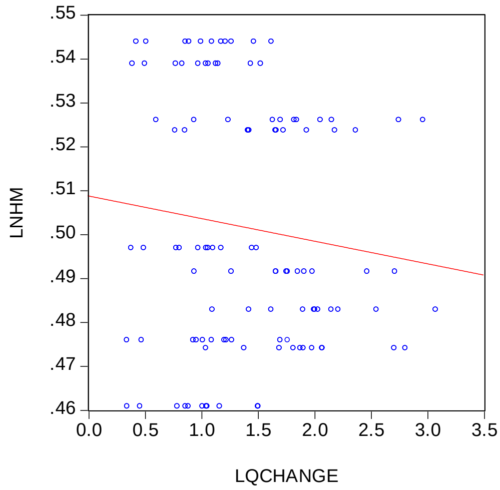
Source: graphs from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Figure 3: Panel Data Scatter Graph between LQCHANGE and LnIS



Source: graphs from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Figure 4: Panel Data Scatter Graph between LQCHANGE and lnHM



Source: graphs from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

From figure 2, LQCHANGE and TRCHANGE can be found to have a strong linear relationship. In figure 3, heteroscedasticity exists in the panel data while it seems that LQCHANGE and LnIS have a weak linear relationship.[24] From figure 4, LQCHANGE and LnHM have horizontal linear relationships over years, meaning

[24] The scatter graph of LQCHANGE and Ln(IS)^2 is tried to draw and the linear relationship is more obvious and stronger; however, the significance of Ln(IS)^2 in the model is too low to take it as a factor for agglomeration. Finally LnIS is chosen to put into the analysis model after comparing with other forms of measures of internal economies of scale.

that cross-section data in the same year is fixed across various sectors when period data changes.

5. the Estimated Model and Results

This empirical analysis part will set up the model and give an explanation of the model choice; later we will see the results of estimations with and without further controlling the endogeneity. In the third step we will estimate the panel data in another dynamic model and then control the endogeneity on this foundation.

5.1 The model

Building on the above-analyzed graphs and empirical facts that manufacturing industry varies from sector to sector over time; a fixed-effects model[25] should be used in this study to restrict the potential heterogeneity among different sectors in manufacturing industry. The analysis model can be set as

$\mathrm{LQCHANGE}_{\mathrm{it}} = \alpha + \beta_1 TRCHANGE_{\mathrm{it}} + \beta_2 LnIS_{it} + \beta_3 LnHM_{it} + \xi_i + \mathrm{u}_{\mathrm{it}}$ (9),

where α is a constant term; ξ_{i} denotes unobserved individual effects, changing across sectors (i=1, 2… N); u_{it} represents a random disturbance and specifications of variables have been explained in section 4.

Before going through the estimation of the model, we should figure out the suitability of the model setting. At first glance, it is obvious that the fixed-effects model is more appropriate than a random-effects one because we did not draw

[25] The fixed effects model is assumed that variations in the constant term can reflect the unobserved variations between individuals. It allows us to estimate the effects of differences within individual levels.

samples randomly from a wide range of data; instead, we collected typical targeted samples that experienced industrial concentration and foreign trades. Besides, we need to support our model choice by means of methodological tools. Firstly we use F-tests to decide to build up whether a fixed-effects model or a pooled model. EViews 6.0 can do F-tests automatically in the estimation of fixed-effects model. The result of F-tests (table 6) shows that a fixed-effects model is better than a pooled model; as the probability of F-tests is zero, smaller than the given significance level of 5%, meaning that the null hypothesis is rejected.[26]

Table 6: Results of Fixed Effects Tests

Redundant Fixed Effects Tests
Equation: Untitled
Test cross-section fixed effects

Effects Test	Statistic	d.f.	Prob.
Cross-section F	30.933097	(10,96)	0.0000
Cross-section Chi-square	158.439132	10	0.0000

Cross-section fixed effects test equation:
Dependent Variable: LQCHANGE
Method: Panel Least Squares
Date: 04/03/12 Time: 18:35
Sample: 2000 2009
Periods included: 10
Cross-sections included: 11
Total panel (balanced) observations: 110

	Coefficient	Std. Error	t-Statistic	Prob.
TRCHANGE	2.654373	0.041557	63.87242	0.0000
LNIS	-0.058709	0.012731	-4.611621	0.0000
LNHM	-0.261472	0.349333	-0.748491	0.4558
C	0.332082	0.179159	1.853563	0.0666

26 The principle of F-test (Fixed effects tests) is introduced in section 3.2.

R-squared	0.976082	Mean dependent var	1.415768
Adjusted R-squared	0.975405	S.D. dependent var	0.618483
S.E. of regression	0.096995	Akaike info criterion	-1.792634
Sum squared resid	0.997246	Schwarz criterion	-1.694434
Log likelihood	102.5949	Hannan-Quinn criter.	-1.752804
F-statistic	1441.954	Durbin-Watson stat	0.372116
Prob(F-statistic)	0.000000		

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Next, H-tests help to make a choice between a fixed-effects model and a random-effects model. The tests' results can be got by EViews 6.0 from the random-effects model estimation. The results from table 7 agree to set the fixed-effects model because the probability of random effects denoted by H-tests is zero, rejecting the null hypothesis of setting up the random effects models.[27]

Table 7: Results of Random Effects Tests

Correlated Random Effects - Hausman Test

Equation: Untitled

Test cross-section random effects

Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.
Cross-section random	307.825607	3	0.0000

** Warning: estimated cross-section random effects variance is zero.

Cross-section random effects test comparisons:

Variable	Fixed	Random	Var(Diff.)	Prob.
TRCHANGE	2.375782	2.654373	0.000254	0.0000
LNIS	-0.071465	-0.058709	0.000002	0.0000
LNHM	-0.642063	-0.261472	0.000550	0.0000

Cross-section random effects test equation:

27 The principle of H-tests (Random Effects-Hausman tests) is also introduced in section 3.2.

Dependent Variable: LQCHANGE
Method: Panel Least Squares
Date: 04/03/12 Time: 18:43
Sample: 2000 2009
Periods included: 10
Cross-sections included: 11
Total panel (balanced) observations: 110

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.649040	0.093456	6.944855	0.0000
TRCHANGE	2.375782	0.026571	89.41343	0.0000
LNIS	-0.071465	0.006689	-10.68365	0.0000
LNHM	-0.642063	0.180177	-3.563506	0.0006

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.994335	Mean dependent var	1.415768
Adjusted R-squared	0.993568	S.D. dependent var	0.618483
S.E. of regression	0.049602	Akaike info criterion	-3.051171
Sum squared resid	0.236191	Schwarz criterion	-2.707474
Log likelihood	181.8144	Hannan-Quinn criter.	-2.911766
F-statistic	1296.222	Durbin-Watson stat	1.752473
Prob(F-statistic)	0.000000		

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

In brief the setting of the fixed effects model (9) is well justified by not only chosen data but also methodological tests.

5.2 Results of estimations

5.2.1 Fixed Effects Estimation

The data collected from various manufacturing sectors is generally assumed to differ from each unit, because the heteroskedasticity is more likely to happen in sectional data at a given time than in time series data. This fact can be further seen from the scatter graphs of variables in Section 4.3.2. Consequently I would like to present the outcomes from the fixed effects estimation with cross-section weighted least squares (CSWLS), which is a special case of the generalized least squares (GLS) and helps to control the heteroskedasticity impact.[28] The estimation results for the fixed-effects model (9) can be specified as

$$LQCHANGE_{it} = 0.65 + 2.38 * TRCHANGE_{it} - 0.07 * LnIS_{it} - 0.64 * LnHM_{it} + \xi_i + u_{it} \quad (9)$$

(9.8) (123.5) (-15.6) (-5.0)

R^2=0.99, DW=2.01, SSR=0.24

All estimated variables pass the significance test (t-tests) and the estimated values can fit the actual values well because of excellent R^2 statistic. The F-statistic is highly significant; it shows that all coefficients are non-zero and that the regression on LQCHANGE can significantly explain the samples. The Durbin-Watson statistics implicates no autocorrelation in the random error term. The value of sum-squared-residuals is so small that the heteroskedasticity problem is under control. Let's make a comparison with the results of normal OLS fixed effects estimation in table 6: the CSWLS results have improved the performances of t-statistics significantly, especially in the term of TRCHANGE; while the goodness of fit and sum squared residuals statistics are kept at similar levels. That is to say, this CSWLS estimation method is somehow more suitable in this model than the OLS one.

The coefficient for TRCHANGE is positive and statistically significant, indicating a strong effect of TRCHANGE on LQCHANGE. In terms of empirical economic

---

[28] Generalized least squares (GLS) is use to avoid inefficient estimations by OLS when the variances of variables are not the same or variables interact with others in some measure. Weighted least squares technique is one typical case of GLS in the condition that no correlations between variables are observed.

interpretation, increases in international trades boost the agglomerations in Guangdong. If the international trades grow 1%, the agglomerations are strengthened accordingly by 2.38%; as both variables TRCHANGE and LQCHANGE measure variations from situations of the initial year. This result confirms the hypothesis of this paper that the international trade has a positive influence on industrial agglomeration in Guangdong. At the same time, this result can be illustrated by the facts that good performances of foreign trades can benefit firms and thus firms are easily attracted to locate intimately in identical areas.

However, the coefficient value for LnIS which reflects the changing tendency of internal economies of scale is of few differences from zero; while the absolute value of t-statistics is statistically high enough to explain the effectiveness of this estimated coefficient. Even though the level of influence is very low, the adverse effect of internal economies of scale on industrial agglomeration can be found. Surprisingly it is against the conclusions in some previous literature that internal economies of scale push spatial concentration in industries. Possible explanations might be: (1) the processed data for IS might be not appropriate to denote the actual conditions of internal economies of scale in this case. The average scale of firms in a manufacturing sector has been used as a proxy for this sector's economies of scale. However the yearly data for sector i's GDP may not match the yearly data for number of firms in sector i. For example, even though some firms disappear at the end of a year, the data for GDP is created by all firms (including those disappearing firms) during the whole year. However, the data for the amount of firms is just obtained at the end of a certain year. Obviously, there is a chance that the final number of firms in a year is smaller than the actual number of firms generating the GDP. This limitation on collecting data might cause the inconformity to preceding research. (2) The concentration of activities in Guangdong is dramatically induced by policies, resulting in this negative consequence. This point could be inferred from the above-discussed explanations of trade affecting concentration: some sectors with higher integrating levels might not benefit from greater economies of scales but actually only from contributions of trade-oriented policies. It means that if plants in a sector were supported strongly by political benefits, they would concentrate on that area even these plants have lower economies of scales. But it is no doubt that internal economies of scale should be put into the model because of their important roles in affecting agglomeration. On the other hand, this paper stresses on the effect of trade on integration so that it does not matter whether the internal economies of scale motivate or restrict the centripetal process.

Similarly the estimation outcomes of LnHM are not in line with preceding research's claim of home market effects accelerating industrial clusters because LnHM, the indicator for home market effects, stays negative with significant effects on the LQCHANGE. This result does support the linkage between market size and integration, but oppose the positive or strong effects on integration. Compared with the proxy for internal economies of scale, this indicator has a slightly stronger relationship with concentration. Given other variables, if the home market effects change by 1%, the agglomeration goes down by nearly 1% (the accurate percentage is 0.64%). Before running the estimation of this model, we assumed that plants locate to exploit the advantages of proximity to larger markets and plenty of reasons believe that home market effects should enhance agglomeration performance. This unexpected result is probably caused by potential defects of the measure of home market effects, such as neglecting the characteristics of products and the strong induces of policies, etc. It is worth to point out that the open door policies in China varied over the observed periods. These policies tended to other areas beyond Guangdong in later periods of the samples, possibly driving down the agglomeration level to some extents. If this is the case, it is not hard to understand why firms enjoying stronger home market effects suffer dispersions across Guangdong. Although the negative sign of measure of home market effects will not be further explored in this study, the reasons of continuing using this indicator should be clarified. First of all, adding important factors into the model can reduce autocorrelation; secondly this proxy is widely used in studies of home market effects.

The estimation results presented above have not completely wiped out the endogeneity; further estimations have to run despite that the endogeneity problem is under control by transforming measures of LQ and TR.

Table 8: Estimation Results of Fixed Effects model Using CSWLS

Dependent Variable: LQCHANGE
Method: Panel EGLS (Cross-section weights)
Date: 04/03/12 Time: 18:45
Sample: 2000 2009
Periods included: 10
Cross-sections included: 11
Total panel (balanced) observations: 110
Linear estimation after one-step weighting matrix

	Coefficient	Std. Error	t-Statistic	Prob.
TRCHANGE	2.380282	0.019274	123.4982	0.0000
LNIS	-0.070008	0.004482	-15.62074	0.0000
LNHM	-0.641006	0.127158	-5.041016	0.0000
C	0.646567	0.066106	9.780792	0.0000
Effects Specification				
Cross-section fixed (dummy variables)				
Weighted Statistics				
R-squared	0.995883	Mean dependent var		1.795002

Adjusted R-squared	0.995325	S.D. dependent var	1.037203
S.E. of regression	0.049562	Sum squared resid	0.235811
F-statistic	1786.152	Durbin-Watson stat	2.011297
Prob(F-statistic)	0.000000		
	Unweighted Statistics		
R-squared	0.994331	Mean dependent var	1.415768
Sum squared resid	0.236372	Durbin-Watson stat	1.734323

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

5.2.2 Instrumental Variables Estimation

The following part presents the consequences from estimations after controlling the endogeneity. We first turn our attentions to 2SLS procedure that takes care of the endogeneity of international trade with respect to industrial concentration. Next, taking the influences of lagged integration on current-period integration into account, we modify the model (9) and use the GMM estimators accordingly to solve the endogenous problem.

5.2.2. (1) 2SLS

Based on the methodology of 2SLS in section 3.3, we start to look for a good instrument for the TRCHANGE. As I addressed before, the doubt of endogeneity comes from the bilateral causality between economic activity clusters and trade. Conversely in this particular case of Guangdong, the growth of international trade has been encouraged by the conscious open door policies at least at the beginning of implementation and then followed by the upwards-trend of industrial integration. Thus the foreign trade values obtained from the initial year of implementing the policies can be regarded as exogenous factors. We will focus $TR_{i,1985}$ on the panel data of international trade in year 1985 when the open door policies started in Guangdong and take this policy-induced variable as an instrumental variable to do the 2SLS estimation of our model.

In order to reduce the side effects of endogeneity, 2SLS method is additionally applied in the earlier estimation by the fixed effects model. Table 9 below summarizes the outcomes of this 2SLS regression using $TR_{i,1985}$ as an IV. Compared with results of table 8, all estimated coefficients for variables except constant term in the model become larger and the signs of the coefficients stay the same. All estimated coefficients are still significant while their absolute values of t-statistics all decrease slightly. From a global perspective, the fitness and the significance of equation seem as good as those in table 8; yet bigger SS_r and smaller DW value indicate increasing risks in heteroscedasticity and autocorrelation respectively, even though the values of these two measurements are acceptable. Because of no considerable changes in estimated exogenous variables, this selected instrumental variable, $\mathrm{TR}_{\mathrm{i},1985}$, could be inferred to restrict the endogeneity well at a certain level. This 2SLS estimation further supports our hypothesis that the increasing international trade can enhance industrial integration in Guangdong. In addition, the results of this 2SLS estimation don't differ substantially from the results of table 8. It suggests that the endogeneity has been controlled in our model thanks to the transformations of measures of agglomeration and trade. The analysis obtained in the former estimation seems more reliable and reasonable.

Table 9: 2SLS Estimation Results of Fixed Effects model Using CSWLS

Dependent Variable: LQCHANGE

Method: Panel Two-Stage EGLS (Cross-section weights)

Date: 04/03/12 Time: 18:56

Sample: 2000 2009

Periods included: 10

Cross-sections included: 11

Total panel (balanced) observations: 110

Linear estimation after one-step weighting matrix

Instrument list: LQCHANGE IV LNIS LNHM C

	Coefficient	Std. Error	t-Statistic	Prob.
TRCHANGE	2.404648	0.019516	123.2114	0.0000
LNIS	-0.069578	0.004492	-15.49093	0.0000
LNHM	-0.627150	0.129083	-4.858507	0.0000
C	0.628539	0.067149	9.360368	0.0000

Effects Specification

Cross-section fixed (dummy variables)

Weighted Statistics

R-squared	0.995887	Mean dependent var	1.781609
Adjusted R-squared	0.995330	S.D. dependent var	1.002952
S.E. of regression	0.049883	Sum squared resid	0.238882
F-statistic	2.22E+25	Durbin-Watson stat	1.948458
Prob(F-statistic)	0.000000	Second-Stage SSR	1.93E-23
Instrument rank	14.000000		

Unweighted Statistics

R-squared	0.994263	Mean dependent var	1.415768
Sum squared resid	0.239194	Durbin-Watson stat	1.683405
Second-Stage SSR	9.80E-24		

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

5.2.2. (2) GMM

When it comes to the endogeneity, our focus turned to the relationship of international trade and integration; then $TR_{i,1985}$ as an instrument went through the experiment of 2SLS. However, does the potential endogeneity in our model only result from TR? Is it likely that some other factor exists? We are hinted to the possibility of self-reinforce of agglomeration. Spatial concentration does not emerge all of a sudden; it is an accumulating process as a result of several observable and unobservable factors. In fact, previous accumulations can affect the agglomeration in present period. Thus a dynamic model concerning a lagged measure of concentration as an independent variable should be tried. Therefore, our fixed effects model (9) can be modified as

$\mathrm{LQCHANGE}_{it} = \beta_0 \mathrm{LQCHANGE}_{i,t-1} + \beta_1 \mathrm{TRCHANGE}_{it} + \beta_2 \mathrm{LnIS}_{it} + \beta_3 \mathrm{LnHM}_{it} + \eta_i + \epsilon_{it}$ (10),

where $\mathrm{LQCHANGE}_{i,t-1}$ represents LQCHANGE in the period of (t-1), η_i indicates individual fixed effects, ϵ_{it} is a random disturbance and other variables' specifications are the same as those in model (9). The central characteristic of model (10) is that β_0 can explain how earlier concentrations affect current ones.

Ex ante discussions on the estimating methodology have pointed out those GMM estimators help clearing the endogeneity and heteroscedasticity off the dynamic model at the same time. So we are going to estimate the dynamic model (10) in GMM way[29] with the help of EViews 6.0 and get the results in table 10. The significant positive coefficient for TRCHANGE suggests again international trade's motivation to industrial integration—trade performance varies 1% resulting in an increase of 2.35% in agglomeration. Estimations for other explanatory variables are not distinct from those in preceding estimations. Signs of coefficients for LnIS and LnHM are still negative, while the absolute values for these coefficients go up a little bit. However, concentration does not vary strongly by internal economies of scale because the estimated coefficient for LnIS is nearly zero. The t-statistics' absolute value increases in LnHM's estimation whereas decreases to 8.62 in LnIS's. It is necessary to point out that the lagged LQCHANGE does not seem to have a close relationship with present LQCHANGE surprisingly. This conclusion is made from the fact that the estimated value of coefficient for the lagged dependent variable is close to zero and this estimated

[29] The specifications of the GMM estimation follow those mentioned in section3.4.

coefficient is significant at 95% confidence level. We should notice that the transformation of LQ into LQCHANGE may weaken the impact of pervious LQ on current LQ, because LQ is the direct measure of industrial agglomeration while LQCHANGE is defined as a difference which is weakly correlated with its previous value. Probably if we had LQ_t in the estimation, it would be more correlated with LQ_{t-1}. What's more, all estimations above turn out that industrial concentration in Guangdong rests on policy-induced trades severely. That is to say, current-period concentration is affected less by foregoing concentration, at least not by all lagged ones. This can explain the very small estimated coefficient for LQCHANGE (-1). From the prospective of estimation's validity, J-statistics[30] implies the validity of this result at 95% confidence level; however, sum-squared-residuals are larger in this estimation method.

Table 10: GMM Estimation Results of Fixed Effects model Using CSWLS

Dependent Variable: LQCHANGE
Method: Panel Generalized Method of Moments
Transformation: First Differences
Date: 04/03/12 Time: 18:52
Sample (adjusted): 2002 2009
Periods included: 8
Cross-sections included: 11
Total panel (balanced) observations: 88
White period instrument weighting matrix
White period standard errors & covariance (d.f. corrected)
Instrument list: @DYN(LQCHANGE,-2)

	Coefficient	Std. Error	t-Statistic	Prob.
LQCHANGE(-1)	0.004310	0.001526	2.823895	0.0059
TRCHANGE	2.351533	0.016476	142.7255	0.0000
LNIS	-0.073128	0.008478	-8.625701	0.0000
LNHM	-0.757146	0.092913	-8.149003	0.0000

30 J-test also called a test for over-identifying restrictions is used in GMM method to test the validity of model as a whole; if the J-statistics$< q_{0.95}^{\chi^2_{k-l}}$, the null hypothesis that the model is valid cannot be rejected. For more details can be referenced to Hansen and Singleton (1982).

Effects Specification

Cross-section fixed (first differences)			
Mean dependent var	-0.106546	S.D. dependent var	0.337955
S.E. of regression	0.065824	Sum squared resid	0.363960
J-statistic	10.37548	Instrument rank	11.000000

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

5.2.3 Comparisons of Estimation Results

Generally speaking, all the above-discussed results account for the significant positive effect of TRCHANGE on the explained variable LQCHANGE, while statistically significant negative influences of other variables except the lagged dependent variable in GMM method. Additionally various statistic evaluations of estimations, from individual assessments such as t-statistics to integral ones such as R^2 and so forth, show that the models can well explain features of panel data in this study and that the estimation results are creditable as well.

In details, estimated coefficients for all variables are similar in three techniques: the coefficients for TRCHANGE and LnHM differ from 2.35 to 2.40 and from -0.82 to -0.49 respectively; whereas the coefficients for LnIS and LQCHANGE(-1) are statistically close to zero, fluctuating between -0.0731 and -0.0696 for LnIS and reaching 0.0043 for LQCHANGE(-1). The latter two approaches belong to instrumental variables approach used for removing the endogeneity problem; I would like to compare estimation results of these two methods firstly and then compare them with the outcome of the first method.

Both 2SLS and GMM techniques are designed to reduce the endogeneity problem; yet, they are usually applied in different types of model, static and dynamic models respectively. In this paper, model (10) was specified as a dynamic model and the lagged agglomeration was considered as an additional independent variable. Then we estimated model (10) in GMM technique only to find a small impact of lagged dependent variable. This result questioned the need to use the dynamic model. First, it is obvious from the results that the number of observations shrank and that with the lagged explained variable, the model lost some degrees of freedom. Perhaps these results would end up with bias of estimation and losses in statistical significance. Second, in GMM way, the estimation is reported with fewer statistics evaluations such as R^2 and standard error of regression that can be given in 2SLS technique. It will therefore be more difficult to review the efficiency and effectiveness of estimation results. Third, no criteria or acquired previous information offered support for deciding how many lags of LQCHANGE should be considered as explanatory variables. Although in 2SLS approach, effects of previous agglomeration were not reflected from the regression, nearly-zero coefficient for LQCHANGE (-1) could make up this shortcoming as this lagged LQCHANGE was not significant to explain the samples. Hence 2SLS way did not worsen the key features of data. Fourth, a GMM estimator that takes all available lagged values as instruments is likely to hurt computational efficiency and estimation's effectiveness. Even if we had added the instrument $TR_{i,1985}$ to the GMM estimation, the result was not improved apparently, compared with the results of 2SLS estimation and of GMM estimation without $TR_{i,1985}$.[31] To sum up, 2SLS' estimation compared favorably to GMM's estimation in terms of fewer limitations of structuring a model with suitable explanatory variables, lower risks of invalid estimations and more statistics assessments.

31 GMM estimation results with the instrument in 2SLS method can be found in Appendix.

Comparing estimations obtained by CSWLS method and by instrumental variables ways, we should pay attentions to differences between CSWLS and 2SLS approaches. Because of the transformations of measures of concentration and foreign trades in the models, the endogeneity has been controlled even in CSWLS method. The values and signs for all estimated coefficients in CSWLS technique have not differed considerably from those in 2SLS way. Both estimation results are good, but the estimation outcome by 2SLS outperforms that by CSWLS; as this 2SLS method further shows good controls of endogenous-variable problems and can be widely considered as creditable and practical. The advantage of using 2SLS method is more authentic but not better estimated results.

6. Conclusions

The objective of this paper is to find out whether international trade influences industrial agglomeration, based on the empirical case of manufacturing industry in Guangdong Province of China. The motivations for this study are the prevailing performance of agglomeration in Guangdong since 1990s and few discussions about industrial agglomeration from the perspective of international trade affecting agglomeration. What is more, quite few debates were based on empirical analysis of Guangdong Province. Guangdong has been piloted by open door policies in China and featured by open economy whose typical case is high degree of international trade; it is thereby of interest to study the topic of this paper. The contribution of this paper is to make up some blanks in previous research and to propose some theoretical and empirical grounds for policy decision makers in China.

Apart from international trade, this paper has taken two other critical incentives to industrial agglomeration, internal economies of scale and home market effects, into account, following the preceding literature. According to characteristics of acquired data from Guangdong and the practical computations of indicators for agglomerations' factors, variables LQCHANGE, TRCHAGNE, LnIS and LnHm were transformed in the models to represent industrial agglomeration, international trade, internal economies of scale and home market effects respectively. The empirical part of analysis was three-piece. The analysis model was set as a fixed-effects model and estimated by CSWLS. Then the endogeneity problem was further under consideration: the fixed-effects model was estimated with a 2SLS technique and latter transformed into a dynamic model with GMM estimation. As the comparisons discussed above, the second method is superior to other ways. The estimation results support the hypothesis of this paper that international trade affects industrial agglomeration positively. Conversely both internal economies of scale and home market effects were shown to restrain integrated performances in the manufacturing industry in Guangdong; these results are in contradiction to the finding of preceding studies. The relationships between industrial agglomeration and internal economies of scale as well as home market effects were unclear in this study and these contrary findings may be caused by limitations of data. Due to changing statistics ways of original data in the period of observation, some manufacturing sectors with high level of agglomeration and international trade were not selected. Besides, the processed approaches of data might have potential self-restrained factors so that the calculated data could not reflect the actual condition of industrial agglomeration and its corresponding factors accurately. Furthermore, the open door policies have exerted great roles in the manufacturing industry in Guangdong; the powers of other less policy-induced factors might be weaken. Except from imperfect data, some other limitation exists in variables. LQ and TR were replaced by their corresponding differences in models to remove the endogeneity; yet there is no way to test whether the endogeneity problem is definitely controlled. However, based on the results of estimations, internal economies of scale and home market effects are vital in studying industrial agglomeration in Guangdong case and the models are acceptable.

In sum, the international trade ought to be considered seriously as a motivation for the industrial agglomeration. Since the agglomeration has contributed to industrial competiveness and economic transitions that help economy development, the benefits of agglomeration have been consistently concerned by regions or countries. This paper concludes that international trade enhances industrial agglomeration, at least in some manufacturing sectors in Guangdong; so boarder views for pursuing the industrial agglomeration have been provided. Further studies can work on empirical analysis at the Guangdong provincial level in this economic geography field. For example, how to make effective policies related with industrial agglomerations in the typical situations of Guangdong Province in China. What is more, it will be interesting to test the findings of the positive effects of international trade on industrial agglomeration in larger sample data.

7. Acknowledgements

I thank my supervisor-professor Pertti Haaparanta for his guidance and excellent comments during the whole process of working at this paper. I also appreciate professor Pekka Ilmakunnas helping me with using EViews to do analysis on panel data and providing very useful comments on my paper.

8. References

Bai, Z., 2008. Econometric Analysis of Panel Data. Tianjin: Nankai University Press.

Baltagi, B. H., 2008. Econometric Analysis of Panel Data. 4th ed. Chichester: John Wiley & Sons.

Brooks, C., 2008. Introductory Econometrics for Finance. 2nd ed. Cambridge: Cambridge University Press.

Brülhart, M. and Torstensson, J., 1996. Regional integration, scale economies and industry location in the European Union. CEPR Discussion Paper No. 1435, [online] Available at :< http://ssrn.com/abstract=291178> [Accessed 20 May 2011].

Brülhart, M., 2000. Evolving Geographical Specialization of European Manufacturing Industries. Lausanne: University of Lausanne Press.

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

Ellison, G. D. and Glaeser, E. L., 1997. Geographic concentration of in US manufacturing industries: A dartboard approach. Journal of Political Economy, 105, pp. 889–927.

Engle, R. F. and Granger, C. W. J., 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica, Econometric Society, 552, pp. 251-276.

Feenstra, R. C., Markusen, J. R. and Rose, A., 1998. Understanding the home market effect and the gravity equation: the role of differentiating goods. NBER Working Paper No. 6804. [online] Available at: < http://ssrn.com/abstract=140605> [Accessed 1 March 2011].

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

Granger, C. and Newbold, P., 1974. Spurious regression in econometrics, Journal of Econometrics, [online] Available at :< www.climateaudit.info/pdf/others/granger.1974.pdf> [Accessed 22 May 2011].

Haaparanta, Pertti, 1998. Regional concentration, trade, and welfare. Regional Science and Urban Economics, 28(1998), pp. 445-463.

Hanson, G. H. and Xiang, C., 2002. The home market effect and bilateral trade patterns. American Economic Review, 944, pp. 1108-1129

He, C., 2009. Concentration and Agglomeration in Chinese Manufacturing Industry. Beijing: Science Press.

He, Y., 2002. Industrial cluster, upgrade and transformation—take digital manufacturing industry in Dongguan, Guangdong as an example. Nankai Management Review, [e-journal] 2002(01). Available through: Wan Fang Data database [Accessed 12 June 2011].

Hoover, E., 1936. Spatial price discrimination. Review of Economic Studies, 4, pp. 182-191.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control, pp.122-3 and pp.231-254.

Johansen, S., 1991. Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. Econometrica, 59, pp. 1551–1580.

Kao, C. and Chiang, M., 1999. On the estimation and inference of a cointegrated regression in panel data. Nonstationary Panels, Panel Cointegration and Dynamic Panels, [online] Available at :< http://faculty.maxwell.syr.edu/cdkao/working/chap.pdf> [Accessed 13 June 2010]

Kao, C., 1999. Spurious regression and residual-based tests for cointegration in panel data. Journal of Econometrics, 90(1), pp. 1-44.

Kim, S., 1995. Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860-1987. Quarterly Journal of Economics, 110, pp. 881-908.

Krugman, P., 1991a. Geography and Trade. Cambridge, MA: MIT Press.

Krugman, P., 1991b. Increasing returns and economic geography. Journal of Political Economy, 99(3), pp. 483-499.

Li, X. and Li, E., 2002. Comparisons of industrial agglomeration mechanism. Zhongzhou Studies, 13(04), pp. 5–8.

Li, X., 2000. Entrepreneurship’s coordination and industrial agglomeration. Strategic Management, 2002(13), pp. 85-92.

Li, Z., 2000. Advanced Econometrics. Beijing: Tsinghua University Press.

Liu, S., 2003. Industrial cluster and its implications to economic development. Industrial Economics Research, (3), pp. 64–68.

Liu, Z., 2002. Research on interdependence of FDI and regional industrial cluster. Industrial Technology & Economy, 6, pp.1-3.

Luo, R., 2002. Comparisons and enlightenments of industrial agglomeration in regions in China. Nankai Management Review, 2002(6), pp.64-72.

Maddala, G. and Kim, I., 1998. Unit Roots, Cointegration and Structural Change. Cambridge: Cambridge University Press.

Marshall, A., 1895. Principles of Economics. London: Macmillan and Co.

Midelfart, K.H., Overman, H.G., Redding, S.J. and Venables, A.J., 2000. The location of European industry. [online] Directorate General Economic and Monetary Affairs, European Commission. Available at :< http://ec.europa.eu/economy_finance/publications/publication11136_en.pdf> [Accessed 20 May 2011].

Ottaviano, G.I.P. and Puga, D., 1997. Agglomeration in the global economy: a survey of the 'New Economic Geography', The World Economy, Wiley Blackwell, [online] Available at : <http://eprints.lse.ac.uk/20324/1/Agglomeration_in_a_global_Economy_A_Survey.pdf> [Accessed 1 March 2011].

Overman, H.G., Redding S., and Venables, A.J., 2003. The economic geography of trade, production and income: a survey of empirics. [online] Blackwell Publishing. Available at :< http://onlinelibrary.wiley.com/doi/10.1002/9780470756461.ch12/summary>[Accessed 23 July 2010]

Pedroni, P., 1999. Critical values for cointegration tests in heterogeneous panels with multiple regressors. Oxford Bulletin of Economics and Statistics, [online] Available at :< http://web.williams.edu/Economics/CV/PedroniCV.pdf> [Accessed 20 May 2011]

Porter, M. E., 1990. The Competitive Advantage of Nations. New York: Free Press.

Pratten, C., 1988. A survey of the economies of scale. In: commission of the European communities. Research on the "Cost of Non-Europe" vol. 2: studies on the economics of integration, Luxembourg.

Rauch, J. E., 1991. Comparative advantage, geographic advantage and the volume of trade. The Economic Journal, 101(408), pp.1230-1244.

Redding, S., 2009. Economic geography: a review of the theoretical and empirical literature. CEP Discussion Papers, 904. Centre for Economic Performance, London School of Economics and Political Science, London, UK.

Ren, Z., 2005. Government policies in the forming of industrial agglomeration: theoretical research and empirical analysis. Jiaxing College Studies, 17（1）, pp. 120–125.

Shao, Y., 2010. FDI and industrial cluster—based on empirical analysis of Guangdong Province. World Economic Outlook, 2010(3), pp. 10-16.

Wang, J., 2001. Thinking about the development of local industrial agglomeration in outward area. The World Geography Research, 2001(3), p.98.

Wang, L., 2005. The interdependence of industrial agglomeration and transfer. Liaoning Economics, 2005(11), p.11.

Wang, Y. and Wei, H., 2006. The analysis of time-spatial characteristics of geographic concentration of industries--taking 28 two-digit industrial industries as examples. Statistical Research, 6, pp. 28-33.

Williamson, J.G., 1965. Regional inequality and the process of national development. Economic Development and Cultural Change, 13, pp. 3-45.

Yang, J. and Feng, G. 2002. Three— dimensional model for cluster orientation. Economic Survey, pp. 48-52.

Yin, D. and Tang, H., 2007. The determinative factors and change trend of industrial agglomeration in china's manufacturing industries: an analysis based on the three economic circles. World Economy Study, 12, pp. 3-9.

You, C. and Li, T., 2010. Empirical analysis of international trade affecting Yangzi Delta's industrial agglomeration. ZhiShi Jingji, 2, pp. 128-129.

Zheng, S., 2002. Copy of the social chain - Taiwanese investment in mainland China "cluster" Analysis. Economic Review, 5, p.13.

9. Appendix

Appendix 1: Economic Indicators for Industrial Agglomeration Zone in Guangdong Province (2004-2009)

Economic Indicators for Industrial Agglomeration Zone in Guangdong Province Year 2004-2009							
Year / Indicators		2004	2005	2006	2007	2008	2009
GDP(billion Yuan)		185.83	649.26	444.67	445.85	315.73	391.31
FDI (billion $)	Contracted	6.01	3.73	7.23	7.36	4.95	3.77
	Paid-in	3.01	2.26	4.56	5.29	4	3.13
Import-export (billion $)	Total	68.32	74.37	116.85	135.36	97.28	89.74
	Export	42.17	48.96	71.13	80.57	48.79	44.3

Source: Guangdong Provincial Department of Foreign Trade and Economic Cooperation, 2010. Conditions of development zones in Guangdong Province, 2004-2009. [online] Available at:<http://www.gddoftec.gov.cn/dept_sub.asp?deptid=1048&channalid=1293>[Accessed 23 September 2010].

Appendix 2: Ratio of international Trade Factor (TR) for Selected Sectors

Panel Data for International Trade Index										
Sectors in Manufacturing Industry	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Manufacture of Textile Garments, Footwear and Headgear	0.7282	0.6272	0.5900	0.4697	0.5691	0.2577	0.2306	0.2260	0.2089	0.2561
Leather, Fur, Feather, Down and Related Products	0.7701	0.6578	0.6269	0.4707	0.6341	0.2671	0.2476	0.2346	0.2285	0.2637
Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	0.5084	0.4762	0.4498	0.2840	0.2888	0.1266	0.1154	0.1095	0.0908	0.0938
Manufacture of Furniture	0.7181	0.7160	0.7051	0.6429	0.6653	0.3801	0.3428	0.3418	0.3091	0.3510
Printing and Record Medium Reproduction	0.5790	0.5841	0.5907	0.5529	0.3822	0.3529	0.3174	0.3029	0.2777	0.3372
Manufacture of Cultural, Educational and Sports Articles	1.1009	0.9696	0.8778	0.7258	0.8487	0.4402	0.4295	0.4336	0.3994	0.4695
Plastic Products	0.7923	0.7149	0.6790	0.5507	0.5041	0.3279	0.3108	0.3079	0.2797	0.3323
Nonmetal Mineral Products	0.3923	0.3590	0.3330	0.2549	0.1847	0.1531	0.1489	0.1421	0.1211	0.1297
Metal Products	0.6804	0.6485	0.6234	0.4726	0.5248	0.2991	0.2902	0.2828	0.2353	0.2795
Manufacture of Electrical Machinery and Equipment	0.7149	0.6840	0.6589	0.5740	0.5626	0.3641	0.3373	0.3211	0.2683	0.3012
Manufacture of Communication Equipment, Computers and Other Electronic Equipment	0.9133	0.9354	0.9653	0.7878	0.9152	0.4866	0.4555	0.4213	0.4001	0.4871

Sources: Author’s calculations, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Appendix 3: Location Quotient for Selected Sectors

Panel Data for Location Quotient										
Sectors in Manufacturing Industry	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Manufacture of Textile Garments, Footwear and Headgear	2.0304	1.8122	1.6566	1.4099	1.8397	0.8562	0.7701	0.7750	0.7835	0.9250
Leather, Fur, Feather, Down and Related Products	2.1472	1.9005	1.7600	1.4128	2.0498	0.8874	0.8269	0.8042	0.8571	0.9525
Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products	1.4176	1.3757	1.2628	0.8524	0.9336	0.4205	0.3853	0.3753	0.3403	0.3389
Manufacture of Furniture	2.0023	2.0687	1.9798	1.9297	2.1509	1.2629	1.1445	1.1719	1.1590	1.2679
Printing and Record Medium Reproduction	1.6145	1.6875	1.6585	1.6594	1.2355	1.1724	1.0599	1.0385	1.0415	1.2180
Manufacture of Cultural, Educational and Sports Articles	3.0695	2.8014	2.4645	2.1784	2.7437	1.4625	1.4341	1.4865	1.4979	1.6961
Plastic Products	2.2090	2.0654	1.9065	1.6530	1.6297	1.0894	1.0379	1.0555	1.0491	1.2004
Nonmetal Mineral Products	1.0938	1.0373	0.9350	0.7651	0.5970	0.5086	0.4971	0.4871	0.4543	0.4687
Metal Products	1.8971	1.8735	1.7503	1.4186	1.6965	0.9935	0.9691	0.9697	0.8825	1.0098
Manufacture of Electrical Machinery and Equipment	1.9934	1.9762	1.8500	1.7230	1.8187	1.2095	1.1261	1.1009	1.0062	1.0881
Manufacture of Communication Equipment, Computers and Other Electronic Equipment	2.5465	2.7026	2.7103	2.3645	2.9588	1.6167	1.5208	1.4445	1.5005	1.7595

Sources: Author's calculations, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Appendix 4: GMM (With Additional Instrumental Variable $\mathbf{TR_{i,1985}}$) Estimation Results of Fixed Effects Model Using CSWLS

Dependent Variable: LQCHANGE
Method: Panel Generalized Method of Moments
Transformation: First Differences
Date: 04/03/12 Time: 18:51
Sample (adjusted): 2002 2009
Periods included: 8
Cross-sections included: 11
Total panel (balanced) observations: 88
White period instrument weighting matrix
White period standard errors & covariance (d.f. corrected)
Instrument list: @DYN(LQCHANGE,-2) @LEV(IV)

	Coefficient	Std. Error	t-Statistic	Prob.
LQCHANGE(-1)	0.004562	0.001547	2.949296	0.0041
TRCHANGE	2.350940	0.016508	142.4094	0.0000
LNIS	-0.073135	0.008487	-8.617176	0.0000
LNHM	-0.758378	0.093043	-8.150835	0.0000

Effects Specification

Cross-section fixed (first differences)

Mean dependent var	-0.106546	S.D. dependent var	0.337955
S.E. of regression	0.065854	Sum squared resid	0.364287
J-statistic	10.37258	Instrument rank	11.000000

Source: Results from EViews 6.0, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Appendix 5: Original Data for Variables in Regression

Sector	Year	LQCHANGE	TRCHAGNE	LnIS	LnHM
1	2000	2.030379	0.728203	-0.029888	0.482915
1	2001	1.812169	0.627228	0.002242	0.474117
1	2002	1.656555	0.590016	0.042000	0.491607
1	2003	1.409854	0.469717	0.047624	0.523732
1	2004	1.839688	0.569058	-3.028681	0.526094
1	2005	0.856219	0.257726	-0.077906	0.543932
1	2006	0.770084	0.230626	-0.085923	0.538940
1	2007	0.774975	0.226046	-0.069469	0.496915
1	2008	0.783480	0.208914	-0.056963	0.460913
1	2009	0.925007	0.256072	-0.041146	0.475978
2	2000	2.147164	0.770089	0.059815	0.482915
2	2001	1.900450	0.657784	-0.001327	0.474117
2	2002	1.760004	0.626862	-0.009580	0.491607
2	2003	1.412820	0.470705	-0.005873	0.523732
2	2004	2.049823	0.634058	-3.256123	0.526094
2	2005	0.887371	0.267103	-0.076682	0.543932
2	2006	0.826920	0.247647	-0.020159	0.538940
2	2007	0.804246	0.234583	-0.054291	0.496915
2	2008	0.857101	0.228546	-0.025404	0.460913
2	2009	0.952550	0.263697	-0.083491	0.475978
3	2000	1.417574	0.508419	0.496479	0.482915
3	2001	1.375739	0.476171	0.540496	0.474117
3	2002	1.262821	0.449780	0.545943	0.491607
3	2003	0.852393	0.283989	0.478250	0.523732

3	2004	0.933617	0.288789	-1.928655	0.526094
3	2005	0.420462	0.126561	0.293965	0.543932
3	2006	0.385305	0.115392	0.342915	0.538940
3	2007	0.375337	0.109479	0.324535	0.496915
3	2008	0.340344	0.090753	0.256303	0.460913
3	2009	0.338947	0.093832	0.143215	0.475978
4	2000	2.002251	0.718115	0.052542	0.482915
4	2001	2.068678	0.716011	0.108573	0.474117
4	2002	1.979799	0.705146	0.065429	0.491607
4	2003	1.929699	0.642913	0.223725	0.523732
4	2004	2.150860	0.665311	-2.413264	0.526094
4	2005	1.262885	0.380135	0.104201	0.543932
4	2006	1.144498	0.342756	0.053453	0.538940
4	2007	1.171920	0.341827	0.042423	0.496915
4	2008	1.159046	0.309059	0.033413	0.460913
4	2009	1.267854	0.350984	0.006037	0.475978
5	2000	1.614463	0.579034	0.489420	0.482915
5	2001	1.687536	0.584090	0.476602	0.474117
5	2002	1.658484	0.590703	0.431593	0.491607
5	2003	1.659446	0.552873	0.459495	0.523732
5	2004	1.235491	0.382166	-1.012427	0.526094
5	2005	1.172420	0.352904	0.264900	0.543932
5	2006	1.059872	0.317412	0.225064	0.538940
5	2007	1.038452	0.302897	0.125766	0.496915
5	2008	1.041500	0.277716	0.078051	0.460913
5	2009	1.217961	0.337172	0.103712	0.475978
6	2000	3.069507	1.100891	0.312167	0.482915

6	2001	2.801366	0.969608	0.374570	0.474117
6	2002	2.464546	0.877799	0.326778	0.491607
6	2003	2.178412	0.725775	0.333408	0.523732
6	2004	2.743692	0.848687	-2.622957	0.526094
6	2005	1.462467	0.440210	0.237083	0.543932
6	2006	1.434110	0.429489	0.286241	0.538940
6	2007	1.486537	0.433595	0.324462	0.496915
6	2008	1.497918	0.399419	0.305079	0.460913
6	2009	1.696067	0.469527	0.268117	0.475978
7	2000	2.208994	0.792265	0.303534	0.482915
7	2001	2.065384	0.714870	0.269450	0.474117
7	2002	1.906489	0.679035	0.259112	0.491607
7	2003	1.652969	0.550715	0.234979	0.523732
7	2004	1.629727	0.504112	-2.594636	0.526094
7	2005	1.089432	0.327924	0.105214	0.543932
7	2006	1.037857	0.310819	0.128637	0.538940
7	2007	1.055493	0.307868	0.117224	0.496915
7	2008	1.049085	0.279738	0.090162	0.460913
7	2009	1.200367	0.332301	0.093923	0.475978
8	2000	1.093832	0.392307	0.269548	0.482915
8	2001	1.037256	0.359015	0.217428	0.474117
8	2002	0.934975	0.333010	0.187323	0.491607
8	2003	0.765145	0.254921	0.167597	0.523732
8	2004	0.597049	0.184681	-2.038616	0.526094
8	2005	0.508594	0.153089	0.133116	0.543932
8	2006	0.497129	0.148881	0.184227	0.538940
8	2007	0.487132	0.142087	0.183164	0.496915

8	2008	0.454325	0.121146	0.193886	0.460913
8	2009	0.468654	0.129739	0.137242	0.475978
9	2000	1.897083	0.680397	0.288474	0.482915
9	2001	1.873523	0.648463	0.249419	0.474117
9	2002	1.750308	0.623408	0.204668	0.491607
9	2003	1.418607	0.472633	0.157920	0.523732
9	2004	1.696519	0.524773	-2.357028	0.526094
9	2005	0.993539	0.299060	0.053237	0.543932
9	2006	0.969086	0.290223	0.074563	0.538940
9	2007	0.969680	0.282838	0.060984	0.496915
9	2008	0.882482	0.235313	0.055815	0.460913
9	2009	1.009809	0.279548	0.047905	0.475978
10	2000	1.993424	0.714950	0.399016	0.482915
10	2001	1.976150	0.683985	0.349516	0.474117
10	2002	1.850003	0.658917	0.262278	0.491607
10	2003	1.723006	0.574049	0.259302	0.523732
10	2004	1.818745	0.562580	-2.037266	0.526094
10	2005	1.209496	0.364064	0.173750	0.543932
10	2006	1.126139	0.337258	0.173035	0.538940
10	2007	1.100924	0.321119	0.145132	0.496915
10	2008	1.006175	0.268296	0.104627	0.460913
10	2009	1.088127	0.301229	0.058606	0.475978
11	2000	2.546461	0.913299	0.065370	0.482915
11	2001	2.702603	0.935425	0.160696	0.474117
11	2002	2.710317	0.965335	0.216196	0.491607
11	2003	2.364454	0.787759	0.215491	0.523732
11	2004	2.958844	0.915239	-1.445785	0.526094

11	2005	1.616690	0.486632	0.080473	0.543932
11	2006	1.520830	0.455460	0.118907	0.538940
11	2007	1.444473	0.421326	0.047523	0.496915
11	2008	1.500493	0.400106	0.078436	0.460913
11	2009	1.759505	0.487089	0.081501	0.475978

Sources: Author's calculations, based on data from China and Guangdong Statistical Yearbooks (2001-2010).

Note: Sector 1-11 are as follows: (1) Manufacture of Textile Garments, Footwear and Headgear; (2) Manufacture of Leather, Fur, Feather, Down and Related Products;(3) Manufacture of Timber Processing, Bamboo, Cane, Palm Fiber & Straw Products;(4) Manufacture of Furniture;(5) Manufacture of Printing and Record Medium Reproduction; (6) Manufacture of Cultural, Educational and Sports Articles;(7) Manufacture of Plastic Products;(8) Manufacture of Nonmetal Mineral Products;(9) Manufacture of Metal Products;(10) Manufacture of Electrical Machinery and Equipment;(11) Manufacture of Communication Equipment, Computers and Other Electronic Equipment.