Agglomeration of industry in China: Does location matter?

M.Sc. Economics extended essay

EC421 International Economics

Author: Anders Akerman

Supervisor: Professor Anthony Venables

Table of content **page 0**

List of figures and tables **page 1**

Abstract **page 2**

1. Introduction **page 3**

2. Agglomeration theory **page 6**

2.1 Traditional theory *page 6*
2.2 'New economic geography' *page 8*
2.3 Geography and economic growth *page 11*
2.4 Creating a link *page 12*

3. The economics of location in Chinese history **page 13**

4. Regression models and results **page 16**

5. Ideas for further research **page 24**

6. Conclusions and implications **page 26**

7. Bibliography **page 28**

8. Footnotes **page 30**

List of figures and tables:

Figure 1:	**Map of China**	**page 2**
Table I:	**OLS Results for equation 1**	**page 16**
Table II:	**OLS Results with heteroskedasticity adjustment for equation 1**	**page 17**
Table III:	**List of conditioning variables in equation 2**	**page 19**
Table IV:	**OLS Results with heteroskedasticity adjustment for equation 2**	**page 22**

Abstract

This paper is concerned with agglomeration (a process of firms congregating in a certain region because of certain advantages in doing so) and location of industry. Literature on Chinese agglomeration has so far only looked at maps when examining the geographic structure of agglomeration. My paper tests the general observation that there is a 'coastal' bias in agglomeration against the observation that there is a bias towards the southeast and the financial centre of Hong Kong. There is a significant positive relationship between agglomeration and proximity to Hong Kong. However, conditioning for other variables that influence agglomeration, I find that the only significant location factor is proximity to the sea. This not only supports literature's theoretical prediction of a 'coastal bias' empirically, but also gives important information when looking at other transition countries. When market forces take an increasing part in economic decision-making and transport costs fall, there might be a coastal bias in the location of industry in a country. If agglomeration has positive growth effects, then regions with sea access will experience increasing income levels, possibly at the expense of landlocked regions in the country.

1. Introduction

In the late 1970's when China's reform period was started, Chinese leaders set the goal of doubling the size of the economy by 2000. Many economists pointed out the difficulties in the project and the impossibility of its achievement. China reached the target with a margin of almost five years, sometime in 1995.

The average annual increase in per capita GNP during 1978 – 2000 was 8.3%[i]. However, for a country of China's size, regional economic disparities are unavoidable. So an uneven development in China's different regions and a corresponding growth in interregional income disparity have characterised this period.

A frequent aim of literature on contemporary China has been to identify those factors that have caused the present situation. Some are policy oriented and point at a long tradition in China to favour development in the coastal provinces to the interior and date back to the tenth century Song Dynasty up to present day with modern policy drives such as the Open Door Policy and the Coastal Development Strategy[ii]. Others highlight interregional disparities in classical growth determinants such as investment in human capital and foreign direct investment and the resulting divergence of productivity[iii].

However, a common ground is that location matters and that a certain degree of agglomeration of economic activity has taken place in specific regions of China.[iv]

Figure I

This paper takes the term "location" seriously. Indeed, at a glance on the map of China it is evident that there is an eastward bias in income per capita among the provinces. It is also evident that the very top income region is centred in the South East of the country, close to Hong Kong and the Taiwan Strait. Literature on Chinese economic geography has so far taken this conclusion as given. Usually, China is divided into the West, Central and Coastal/East (see Figure 1). Sometimes, the Southeast is singled out as an important

subset of the East. And sometimes certain provinces are excluded because of lack of data, which is in fact a result of endogeneity since absence of data is likely to be correlated with lower income[v]. This paper aims to make the analysis of location more rigorous. Accounting for all the provinces in China excluding Hong Kong, I will examine the relationship between location (interpreted here as the distance of the province from the sea or Hong Kong) and agglomeration of industry. Thus, the paper aims to test the common conclusion. Its aim is simply to add a more rigorous structure to the discussion.

An important point to be made is that this paper takes the presence of an agglomeration process in China as given, although I provide proof for this from previous literature. What I am concerned with is if and how location matters.

Section 2 aims at providing an understanding of the economic theory of agglomeration, or 'economic geography'. I will briefly describe the disparity between regions and give a historical background in section 3. And in section 4 empirical tests of the impact of location on agglomeration are presented. I try to identify some issues for further research in section 5 and discuss possible conclusions and policy implications in section 6.

2. Agglomeration theory

2.1. Early theory

Traditional economics predict differences in economic activity across space to be explained by differences in technologies, saving rates and investment, endowments in resources of production or preferences. However, this fails to explain today's very uneven distribution of economic activity, often between two very similar geographic regions. Much literature has attempted to explain this inequality of development across regions, starting already with Adam Smith. and later on promoted in the middle of the last century by for example Myrdal calling it 'cumulative causation'. Its main conclusions are that, firstly, industrial development will occur at a spatially uneven pace. Secondly, growth will be higher in those regions experiencing industrial agglomeration. It diverges from standard neo-classical theory in that growth rates of GDP and GDP per capita will be unequal across regions, as there is no full factor price equalisation. I will briefly present the general idea of early agglomeration literature in this section of the paper.

Early literature recognises that there are advantages for firms to agglomerate in a certain location. Marshall (1920) points at three types of externalities: labour pooling, intermediate inputs (for both, increasing returns to scale characterise the industry) and knowledge spillovers[vi]. He also notes that external economies are more common in manufacturing than in agriculture. This difference has been an illustratory example in

later work on the subject. Myrdal (1957) concludes that the process tends to increase intranational economic inequalities. In what he describes as 'cumulative causation', one industry locates itself in a region initiating a process of agglomeration there. As employment increases, firms, which were present before, can sell more and raise profits because of the increase in demand. This attracts more factors of production from outside, such as labour and capital. In addition, the increase in profits increases investment through savings as well.

Geography matters in determining where such processes begin and coastal access is an important factor: "... naturally economic geography sets the stage. Commercial centres are, of course, usually located in places where there are reasonably good natural conditions for the construction of a port..."[vii]. However, 'historical accident' is seen as the fundamental cause. It may be that the initial investment would have resulted in greater success elsewhere, but because of some event it did not occur there. And as long as the externalities are strong enough, the process continues once it has begun. Myrdal also describes the forces of backwash effects, e.g. when an agglomerating region attracts factors of production from elsewhere: "By themselves, migration, capital movements and trade are rather the media through which the cumulative process evolves – upwards in the lucky regions and downwards in the unlucky ones"[viii]. Counteracting these are 'external diseconomies' or 'spread effects' when increases in factor payments, especially immobile factors reach a certain level, thereby halting the process of agglomeration. This benefits the surrounding regions: "It is natural that the whole region around a nodal centre of

expansion should gain from the increasing outlets of agricultural products and and be stimulated to technical advance all along the line".[ix]

This is linked to the discussion on the existence of cities for which a similar reasoning is employed. For example Mills (1967) produces a model for the occurrence and existence of cities by assuming increasing returns to scale and non-homogenous land, he writes that the two assumptions of a traditional economic model without cities "most in conflict with reality are that land is homogeneous and that production functions all have constant returns to scale"[x], thereby using a geographic determinant to agents' decision-making in the location of industry.

2.2 New economic geography

Only recently, the literature has attempted to formalise these ideas. These attempt to explain large-scale agglomerations by looking at pecuniary rather than technological externalities. A common thread throughout the literature is the usage of monopolistic competition by Spence (1976) and Dixit and Stiglitz (1977) in combination with a Samuelson (1954) 'iceberg' cost of trade. In this section, I will briefly describe the main conclusions of the literature[xi]. However, I will not go into detail or formally describe any specific model. For the purpose of this paper, it is more important to get an overview of the literature, its assumptions and results.

The fundamental question is why a priori similar countries with the same comparative advantages can develop different production structures because of their different market access, by assuming increasing returns to scale and imperfect competition. One of the first models is that of Krugman and Venables (1990) who assume two regions, one large core and one smaller periphery. There are two industries, one that is monopolistically competitive with differentiated goods, each produced under increasing returns to scale. The other sector has one homogenous good under constant returns to scale and is perfectly competitive. There is no difference comparative advantage as relative endowments of factors of production are the same, but the larger country has more of both factors.

The result of this fundamental 'new economic geography' model is, with finite trade costs, an agglomeration of the imperfectly competitive sector's firms in the core, i.e. the core's share of industrial output is larger than its share of endowments. When trade costs are very high, proximity to the product market is the primary determinant of firm location. As trade costs gradually decrease, factor market competition grows stronger and an agglomeration process starts in the core when firms from other regions seek the higher profit levels in the core, which increases the core's share of industry. But with even trade costs approaching zero, factor market competition becomes so hard in the core that spread effects overtake and there is disintegration again back to the original shares of industry in the regions.

The model develops a framework in which agglomeration can be explained in terms of market access but it crucially assumes that the regions are of different size. Consequently, several models have tried to describe agglomeration between two a priori identical regions of the same size. I will describe briefly. Krugman (1991) assumes that certain factors of production are mobile across countries which leads to migration induced demand linkages when an agglomeration process, for whatever reason, starts in some region. For example, when one firm in a sector with increasing returns to scale relocates between two a priori identical regions, workers will benefit from migrating along with it (because of a rise in the number of product local varieties and labour demand and wages), provided their labour supply is sufficiently elastic. The workers' spending creates additional demand in that region and more firms will benefit from relocating, creating a 'core' and a 'periphery'.

However, factors of production cannot always be assumed to be mobile across regions (see discussion on China in Section 3). For example, only 1.5% of EU citizens live in a member state where they were not born[xii]. To deal with this, Venables (1996) formalises Hirschman's 'forward' and 'backward' linkages between firms operating in sectors under increasing returns to scale. He assumes vertical linkages through intermediate goods between upstream and downstream firms. The result is that an increase in production of one of the firms is beneficial for the other. This creates a process of agglomeration even with perfectly immobile labour. Finally, Baldwin (1997) models a case of endogenous growth and factor accumulation which creates a possibility for agglomeration in a way

similar to that in the migration induced demand linkages model and also similar to what Myrdal wrote.

The models contribute to our understanding of the agglomeration process. However, for the case of China, there is little in the models to tell us where agglomeration will start. I will describe in section 3 where and why agglomeration may have started. The models do not tell us why this was the case, only that a small asymmetry, possibly caused by historical accident, becomes amplified through a process of agglomeration. From what we learnt in the first section, China still seems to be in the stage of development where agglomeration is increasing. But according to theory, agglomeration might some day reach a peak and then start decreasing again, either when trade costs decrease far enough or when 'spread' effects take over because of 'congestion' or unsustainable levels of factor costs in the regions experiencing agglomeration.

2.3 Geography and economic growth literature

Although this paper does not examine economic growth but agglomeration, a general result from the models mentioned so far is that the result of agglomeration is a higher income level in the 'core' region. Therefore, it is useful to briefly describe the literature that focuses on geography and economic growth.

Gallup and Sachs (1998) address the issue of coastal location and economic growth. They formalise the theory by accounting for transport costs in a Harrod Domar model (also

known as the AK model) framework. But more importantly, they establish empirically that none of twenty-eight landlocked non-European countries is a high income country as well as finding a significant gap in income levels between coastal and landlocked countries[xiii]. They also perform regressions finding, firstly, a significant relationship between distance from "core" areas of the world economy and costs of trade. And, secondly, they find evidence for cost of trade having a significant impact on the log level of per capita GDP in a cross section regression including 150 countries in 1995. Indeed, they find the results so striking that they argue in favour of inserting geographic variables in cross-country analysis of growth.

2.4 Creating a link

Clearly, the two different lines of research that I have just described differ in some important ways. Firstly, agglomeration is a self organising behaviour of economic agents. And in theory this process can take place between geographically identical regions. Secondly, agglomeration theory describes the process of firms locating close to each other, and is not primarily focused upon economic growth levels per se. However, I believe that there are two ways in which the two lines can complement each other. Firstly, the result of agglomeration in all the models that I have described is an increase in real per capita income due to the increase in production in the 'core' region. Secondly, agglomeration theory does not go into detail describing the initiating 'spark' from which the process begins. However, as some literature suggests, coastal location can constitute exactly such an event. The agglomeration process can then continue in a gradual way

when transport costs start to fall, the country enters international trade or because of economic liberalisation. In fact, Radelet and Sachs (1998) show that "almost all countries with macroeconomic success in labor-intensive manufacturing
exports have populations almost totally within 100 kilometers of the coast"[xiv].

3. The economics of location in Chinese history

Agglomeration in China has a long history. In fact, centuries of such change beginning in at least the Song Dynasty (960 – 1279)[xv] had moved China's economic centre toward the coast. Consequently events and policy have historically seemed to favour coastal provinces. China's gradual opening and subsequent confrontation during the Opium War with foreign powers resulted in even greater access for foreigners. And importantly, the restricted areas which became accessible were, with Shanghai as the one exception, along the southeast coast. And in recent days, the Open Door Policy and the establishment of Special Economic Zones continue a long process where government policy has favoured foreign investment along the coast.

There are exceptions, for example when production moved westwards during the Second World War by nationalist forces and when major investment in physical capital for heavy industry in the interior was made by the Mao government during the 'Third Front Policy'. The latter reached a peak during the Third Five-Year Plan when 71% of Chinese investment in fixed assets was made in the interior[xvi]. But despite its intentions, the state's administrative organisation seems to have failed. The investments were subject to

bureaucracy and lack of good management and employee incentives. In addition, the neglect of the coastal area during this period resulted in a bias towards light industry and manufacturing in the area. This, in light of recent history, turned out to be in its favour.

There has been some debate over the trend of interregional disparity over the reform period which started in 1978. Lu and Song (2002) found, among other things, a narrowing of inter-provincial output disparity from 1978 - 1990 but after 1990 a widening of inter-provincial output disparity and a general widening of livelihood indicators of development between regions. Overall, they found an increase in output disparity since 1980. The per capita GDP ratio for the three regions was (Coast; Central and West) 1.00:0.50:0.45 in 1980, the very beginning of the reform period, and 1.00:0.44:0.35 in 2000.

There are also signs of industrial agglomeration in China. Golley (2002), firstly, compares gross value of industrial output (GVIO) between regions. By ranking all the provinces in terms of growth rates, she identifies five provinces in the southeast, Guangdong, Zhejiang, Jiangsu, Fujian and Shandong as the core of growth. Their combined share of national GVIO rose from 24% in 1978 and 45% in 1998. Interestingly, there is also a correlation of as much as 0.93 between real GDP growth and change in shares of GVIO. Secondly, she also calculates locational Gini coefficients (a measure of the degree of agglomeration of a sector) for nineteen main manufacturing sectors. What she finds is that in fifteen of the nineteen sectors the Gini coefficient rose between 1985

and 1997. These two results are typical for what we would expect in an agglomerating economy.

As mentioned in the previous section, labour mobility plays a large role in agglomeration and an even more important role in income equalisation. Literature seems to tell us that inter-regional mobility is rather limited in China, at least compared to intra-regional rural-urban mobility[xvii]. This is in part a result of policy. To quote the PRC Minister of Labour, policy wants to "limit the inter-regional movement of workers to the current level and the majority of redundant rural workers should leave agriculture for new jobs locally[xviii]". Banister (1997) finds that only about 3.5% of the total rural population in China worked outside their native province. Much of inter-regional income inequality is in fact a result of rural-urban disparities although rural-urban (intra-regional) labour mobility has been observed to be rising. The fact that inter-regional labour mobility is low is incompatible with Krugman's migration-induced demand linkages model but does not prevent us from using Venables's vertical linkages model or Baldwin's endogenous growth model. However, this paper does not aim to discriminate between different agglomeration models.

We have seen that there are many signs of the existence of interregional income and industry concentration disparities. We have also seen that there are signs of a widening of this during the reform period up until today. In the literature, there is a general consensus that location matters for agglomeration of industry and growth, but different views on how it matters.

4. Regression models and results

As explained in the previous section, the process of agglomeration can start either by 'historical accident' (for example, one firm reallocates between two previously identical regions) or because of some conditions that favour one region over the other when costs of trade decrease. Indeed the Gallup and Sachs model can be interpreted in a way that coastal bias can cause agglomeration. The historical accident which initiates the agglomeration can simply be the coastal location of a province. What I want to examine is how geographic location explains the nature of agglomeration in China.

One way to do this would be to construct a growth model and use panel data to examine the effect location has had on growth during the reform period (the most recent 25 years when market forces have increasingly been allowed to determine investment and allocation of firms). However, not enough data is available. Several provinces would have to be excluded. This would cause a problem, since poor provinces tend to have less available data during this period. For example, Hainan for example is a less well off province in the coastal block, does not have available data over the full period[xix]. Alternatively, missing data points could be estimated but this would eventually put the overall credibility at risk. In addition, an analysis using panel data would be beyond the possible scope of this paper.

Instead, I will use the data from the paper by Jane Golley, "Regional Patterns of industrial development during China's economic transition". She uses Chinese data from

1989 and 1994[xx] and examines the performance of 22 different manufacturing sectors in China's provinces. She uses a shift share analysis[xxi] to break down the components of growth of output in a certain sector in a certain province. Firstly, she examines if a sector in a particular province has grown at a different rate than national growth in GVIO. Secondly, she breaks down this divergence into a structural component and a locational component. The structural component captures the part of the divergence that is a result of the different growth of the specific sector on a national level and the general national growth rate of the economy. The locational component is then a sort of residual, a catchall for all factors that determine the remaining divergence in growth rates.

What is interesting for this paper is the case of a positive locational component. This happens when the specific sector in a specific province grows faster than would have been the case had the local industry experienced the national growth rate. This is a sign of agglomeration. I will use this data, which covers the main nineteen manufacturing sectors in China.[xxii] The number of sectors with a positive locational component will function as my variable for degree of agglomeration, *A*. It would have been useful to know the strength of the locational component in each case but this data is not available. Relevant geographic variables are distance to sea, *s*, and distance to Hong Kong, *hk*. A log-linear model would intuitively seem to be appropriate since it is reasonable to assume that the effect of distance to sea is concave, meaning that the effect of the distance variables can be assumed to be declining the larger the distance becomes. However, when examining the data I find that a standard linear regression model is supreme to the former. Also,

Gallup and Sachs (1998) use a linear model so I will proceed in line with previous literature. Firstly, I use the following simple model:

$$A = \beta_0 + \beta_1 s + \beta_2 hk + \varepsilon \quad (1)$$

The following results are found with standard Ordinary Least Squares:

Table I

Independent variable	Coefficient	Standard error	t-ratio	P-value
Constant	11.27 ***	2.27	4.95	0.0000
*H*k	-0.0026	0.0016	-1.60	0.12
*Se*a	-0.001608	0.0019	-0.84	0.41
R-squared	0.22			
Adj. R-squared	0.16			

*** significant at 1% level

** significant at 5% level

* significant at 10% level

Heteroskedasticity is very likely as we can assume *A* to be correlated with the size of the error. In fact, when using White's method to test for heteroskedasticity[xxiii] (as is appropriate for cross section data), I find that the significance level of the F-statistic is

5.2% so heteroskedasticity cannot be assumed away. To eliminate any heteroskedasticity which is there, I use White's correction method:

Table II

Independent variable	Coefficient	Standard error	t-ratio	P-value
Constant	11.27 ***	2.68	4.21	0.0003
Hk	-0.0026 *	0.0015	-1.77	0.088
Sea	-0.0016	0.0017	-0.95	0.35
R-squared	0.22			
Adj. R-squared	0.16			

*** significant at 1% level

** significant at 5% level

* significant at 10% level

The results are very interesting. When adjusting for heteroskedasticity I find that proximity to Hong Kong seems to cause agglomeration, although at a relatively weak significance level of 8.8%. Distance from sea seems to be insignificant. This does not support the theory of a 'coastal bias' of agglomeration in China. However, it supports the conclusions of the literature that argues that there is a southeastern bias instead of simply a coastal one.

Of course, we face a problem of multicollinearity in this model. By merely looking at a map, we realise that the distance of a point to Hong Kong must be correlated with distance to the sea. In fact the correlation between the variables is 0.46. This is a problem, but a necessary one unfortunately. There is a risk of a high variance when using least squares models which may cause variables to be individually insignificant but jointly they seem to explain the values of the dependent variable well.

However, the model is identified. We can see that by looking at three types of hypothetical observations: provinces (referring to figure 1) in the northwest (1), northeast (2) and southeast (and close to Hong Kong) (3) respectively of the country. If we consider the case of type 1, we see that the exogenous variables associated with both β_1 and β_2 will be influencing A. And for type 3, both s and hk will be very small. But for type 2, s is very small but hk large. By this observation, we see that there is sufficient variation in s and hk across provinces in different parts of the country for both β_1 and β_2 to be identified. So the model will be able to distinguish between the effects of s and hk.

There is a strong possibility for omitted variable bias in this model. Our results so far are merely an overview of how the degree of agglomeration is distributed across space. Of course, distance is not likely to be the only driver of the process. There may be many other factors that influence agglomeration and increases in GVIO. This will cause the least square model to show biased estimates. In order to prevent omitted variable bias I want to control for additional variables that might have an effect on the degree of agglomeration. This will show whether these factors can explain the process better than

our 'location' variables. If this is is the case, it means that agglomeration has happened because of other forces that for some reason have been linked with agglomeration.

I will use data from the China Statistical Yearbook. As mentioned previously, the Golley data covers the period 1989 – 1994. Therefore, my preference has been to use data from 1989, the starting year of the period. However, when unable to do so, some variables are based on data from 1991 or 1992. I will use the following variables:

Table III

Variable	Notation
Foreign Direct Investment/GDP	*fdi*
Dependency Ratio[xxiv]	*dr*
Illiteracy	*illit*
Proportion of labour in agriculture	*agr*
Investment in fixed assets/GDP	*I*
Length of railroad network/Land area	*rail*
Length of highway network/Land area	*high*
Dummy: special economic zone dummy	d_{SEZ}

When doing this I attempt to filter out the effects of other factors that can possibly play a role in determining the location of agglomeration. Foreign direct investment (FDI) can have a large role in the pattern of spatial economics in an emerging market and can be a trigger and supporting external factor of agglomeration. The dependency ratio can be a

proxy for the demographic conditions of a province to attract new firms and can therefore influence agglomeration. Illiteracy and proportion of labour in agriculture can be proxies of the level of human capital and experience of working in non-agricultural sectors, respectively. As a proxy for infrastructure I chose the following crude variables: telephones per capita, length of railroad and highway network. I then divide the two latter variables by the land area. Investment will also influence the geographical structure of the economy. Finally I will introduce a dummy that will capture policy effects as it is often argued in literature that agglomeration has happened as a direct effect of preferential policy and treatment for South Eastern coastal provinces. The dummy is assigned value 1 if the province contains a special economic zone area and 0 if it does not.

$$A = \beta_0 + \beta_1\, sea + \beta_2\, hk + \beta_3\, fdi + \beta_4\, dr + \beta_5\, illit + \beta_6\, agr + \beta_7\, I + \beta_8\, rail + \beta_9\, high + \beta_{10}\, d_{SEZ} + \varepsilon \quad (2)$$

Table IV (adjusted for heteroskedasticity)

Independent variable	Coefficient	Standard error	t-ratio	P-value
constant	-12.68	14.41	-0.88	0.39
sea	-0.0055 *	0.0031	-1.76	0.095
*h*k	-0.00059	0.0030	-0.20	0.85
fdi	-1.42 **	0.58	-2.45	0.024
dr	-0.15	0.17	-0.87	0.40
illit	-0.156	0.14	-1.13	0.27
agr	0.20	0.14	1.40	0.18
I	0.068 ***	0.021	3.18	0.0049
rail	-0.24	0.17	-1.42	0.17
high	0.025	0.021	1.14	0.27
d_{SEZ}	8.3	6.84	1.2	0.24
R-squared	0.56			
Adj. R-squared	0.33			

*** significant at 1% level

** significant at 5% level

* significant at 10% level

It becomes apparent that there is a certain degree of multicollinearity as the R-squared is actually rather high for a cross section regression while some variables are individually insignificant. When applying a version of the extreme bounds analysis[xxv] it shows that there is a degree of multicollinearity. We searched for multicollonearity by observing the

consequences of taking out one of the conditioning variables at a time (in different combinations) of the regression. However, the conclusion drawn does not change when doing so since our estimates remain the *best linear unbiased estimator* (BLUE) even if we have strong multicollinearity. The largest problem we have is a high sample variance (also a result of small sample size).

The result is very interesting. What we see is that when conditioned for other factors the sea variable becomes more significant than the Hong Kong variable. Indeed, the Hong Kong variable becomes statistically insignificant, while the sea variable becomes significant. This revelation supports literature that argues coastal location favours agglomeration of industry. It shows that the case of Hong Kong and its successful neighbouring provinces can largely be explained by, firstly, their proximity to the sea and, secondly, other external factors that may favour agglomeration. Indeed, the results support the idea of a 'coastal bias' in agglomeration.

We also get the expected result that total investment drives agglomeration. However, there is a perverse result of a significant negative coefficient of the FDI variable which I find difficult to explain.

5. Ideas for further research

In terms of theory, it would be very exciting to formalise these results. Perhaps when modelling agglomeration and the inequality between two regions, one could take ‘access

to trade' into account. This could be done by using Samuelson's 'ice berg' trade cost model and link this to the distance over land there is between the region concerned and the sea or to a significant port. In fact, this would establish a link between the two lines of modelling.

Empirically, the data used in this paper is very specific in nature. It would be very interesting to extend the time period of the shift share analysis so that it covers the entire reform period from 1978 until present day. This would give us a better picture of the entire process even though the data used in this study should give us a clear indication.

Another point is that shift share analysis is a non-dynamic way of looking at the data. An alternative to this would be to create a dynamic growth model taking into account all factors of production in all provinces and examine the nature of any coastal bias in agglomeration by panel data. This would be restrained by data availability, especially for the earlier years of the reform period and for lower income provinces.

In a broader sense, it would be interesting to extend the tests beyond the borders of China. We have seen a statistically significant negative relationship between distance from sea and degree of agglomeration. This is probably true on an international level as well, simply by looking at an economic map. But we cannot be certain that this is the case. Transition economies provide an interesting laboratory for economic geography economists because of the rapidity of agglomeration that sometimes take place there. Useful research would be to examine how widespread the Chinese situation is

internationally and examine countries and regions such as Russia, India, Eastern Europe and others.

6. Conclusion and implications

In the 18th century Adam Smith predicted the low income levels of landlocked nations and he mentioned coastal access as a way to promote division of labour on an international level. Myrdal also mentions this as perhaps the most important geographic factor. This is supported by the Chinese data that I have used. When conditioning for external factors that might effect agglomeration in provinces, I find that the best geographic measure to use when explaining agglomeration in the Chinese case is distance to the sea, for which I find a significant negative linear relationship.

A 'natural' tendency for agglomeration to be initiated close to the sea has important implications. Firstly, it highlights one of the challenges transition economies face, an agglomeration process in certain parts of the country leading to an increase in regional inequality. In China's case there was certainly a significant difference in income prior to the reform period but it has been further increased by market reforms and the opening to international trade. All provinces have expanded during the period[xxvi] but theory tells us that a 'core' region can grow at the expense of the 'periphery' when firms relocate and factors of production are diverted.

Secondly, the introduction of policy favoured economic centres such as China's Special Economic Zones can be expected to be more successful when located close to the sea if there is a 'coastal bias' for agglomeration. This was certainly in the minds of the Chinese government when the zones were introduced, although another reason that has been mentioned was the will to 'contain' the experiment with market economics in a certain region.

And finally, it shows the tendency of firms to prefer to be located close to the sea because of better access to trade, possibly at the expense of landlocked countries. When keeping this in mind, the low income levels of countries and regions such as Afghanistan and Central Asia along with other landlocked areas could have been predicted long ago. Certainly, there are many exceptions to such a rule of thumb, but the economies of successful landlocked countries are by and large not based on industry producing tradable physical goods dependent on access to trade networks.

This paper finds that there is a positive significant relationship between sea access and agglomeration in China. This hypothesis has so far been assumed in the literature but without rigorous proofs. On the basis of the dataset I used, the hypothesis cannot be rejected.

7. Bibliography

Baldwin, R. E. (1997). 'Agglomeration and endogenous capital', Mimeo, Graduate Institute of International Studies, University of Geneva.

Banister, J. (1997). 'China: Internal and Regional Migration Trends', in Scharping, T. (ed) *Floating Population and Migration in China: The Impact of Economic Reforms*, p. 72-97.

China Statistical Yearbook, 1989, 1993 and 2001.

Chen, J. and B. Fleischer, B. (1996). 'Regional income inequality and economic growth in China', *Journal of Comparative Economics*, 22: 141-164.

Chen, J. and B. Fleischer, B. (1997). The Coast-Noncoast Income Gap, Productivity, and Regional Economic Policy in China', *Journal of Comparative Economics*, 25: 220-236.

Dicken, P. (1998). 'Global Shift: Transforming the World Economy'

Dixit, A. K. and Stiglitz J. E. (1977). 'Monopolistic competition and optimum product diversity', *American Economic Review*, 67: 297-308.

Golley, J. (2002). 'Regional patterns of industrial development during China's economic transition', *The Economics of Transition*, 10(3): 761-801.

Fujita, M., Krugman, P. and Venables, A. J. (2000). 'The spatial economy: Cities, Regions, and International Trade'.

Gallup, J. L. and Sachs, J. with Mellinger, A. D. (1998). 'Geography and Economic Growth', *Annual Bank Conference on Development and Economics*.

Kanbur R. and Zhang X. (1999).'Which regional inequality? The evolution of rural-urban and inland-coastal inequality in China from 1983 to 1995', *Journal of Comparative Economics*, 27: 686-701.

Krugman, P.R. and Venables, A. J. (1990). 'Integration and the competitiveness of peripheral industry', *Unity with Diversity in the European Community*, Cambridge University Press.

Krugman, P. R. (1991). 'Increasing returns and economic geography', *Journal of Political Economy*, 99: 484-499.

Krugman, P. and Venables, A. J. (1995). 'Globalisation and the inequality of nations', *Quarterly Journal of Economics*, 110: 857-880.

Levine, R. and Renelt, D. (1992). 'A Sensitivity Analysis of Cross-Country Growth Regressions', *The American Economic Review*, 82: 942-963.

Liu, A., Yao, A. and Zhang, Z. (1999). 'Economic growth and structural changes in Employment and Investments in China, 1985-94', *Economics of Planning*, 32(3): 171-190.

Lu, Z. and Song, S. (2002). 'China's Regional Disparities in 1978-2000'.

Marshall, A. (1920). 'Principles of Economics', London: Macmillan Press.

Mills, E. (1967). 'An aggregative model of resource location', *The American Economic Review*' 57(2): 197-210.

Myrdal, G. (1957). 'Economic Theory and Underdeveloped Regions', Essex: Gerald Duckworth.

Ottaviano, G. and Puga, D. (1997). 'Agglomeration in the global economy: A survey of the 'new economic geography', Centre for Economic Performance, London School of Economics and Political Science.

Radelet, S. and Sachs, J. D. (1998). 'The East Asian Financial Crisis: Diagnosis, Remedies, Prospects', *Brookings Papers on Economic Activity*, 1: 1-74.

Samuelson, P. A. (1954). 'The transfer problem and transport costs, II: Analysis of trade impediments', *Economic Journal*, 64: 264-289.

Smith, A. (1776). 'An enquiry into the nature and causes of the wealth of nations'.

Spence, M. (1976). 'Product selection, fixed costs, and monopolistic competition', *Review of Economic Studies*, 43: 217-235.

Venables, A. J. (1996). ' Equilibrium locations of vertically linked industries', *International Economic Review*, 37: 341-359.

White, H. (1980). 'A heteroskedasticity-consistent covariance estimator and a direct test for heteroskedasticity', *Econometrica*, 48: 817-838.

Yang, D. (1997). 'Beyond Beijing. Liberalization and the regions in China', New York: Routledge.

8. Footnotes

[i] China Statistical Yearbook (2001) as written in Lu and Song (2002)
[ii] Yang (1997)
[iii] Chen and Fleischer (1997)
[iv] The conclusion of Golley (2002), which includes data that I will discuss in Section 4
[v] Chen and Fleischer (1997), exclude Guangxi and Hainan, both coastal provinces but also relatively low income regions
[vi] Golley (2002) from Marshall (1920)
[vii] Myrdal (1957), p. 26
[viii] Myrdal (1957), p. 27
[ix] Myrdal (1957), p. 31
[x] Mills (1967), p. 198
[xi] This outline is partly drawn from the excellent review made in Ottaviano and Puga (1997)
[xii] Ottaviano and Puga (1997)
[xiii] Using 1995 World Bank PPP adjusted data
[xiv] As quoted in Gallup and Sachs (1998)
[xv] Yang (1997)
[xvi] Lu and Song, 2002
[xvii] Kanbur and Zhang (1999)
[xviii] As quoted in Kanbur and Zhang (1999)
[xix] It remained a part of Guangdong until 1988.
[xx] Very importantly, Golley notes that similar procedures have been done for different time periods with similar results.
[xxi] She adapts this method from Liu, Yao and Zhang (1999).
[xxii] Food, beverages and tobacco were combined into one unit to ensure time consistency and the sector producing meters, instruments and other measuring equipment was omitted from the dataset.
[xxiii] Halbert White, "A heteroskedasticity-consistent covariance estimator and a direct test for heteroskedasticity", *Econometrica*, vol. 48, pp. 817-838, 1980 is the reference given in the Eviews software manual.
[xxiv] Dependency ratio = (Number of people not aged between 15 and 65) / (Total population)
[xxv] See Levine and Renelt (1992).
[xxvi] Lu and Song (2002)