Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 09-23

Regional determinants of FDI in China: A new approach with recent data

Martijn Boermans
Hein Roelfsema
Yi Zhang

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Janskerkhof 12
3512 BL Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics.
It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to J.M.vanDort@uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors

Please direct all correspondence to the first author.

Hein Roelfsema*#
Martijn Boermans#
Yi Zhang*
* Utrecht University
Utrecht School of Economics
Janskerkhof 12
3512 BL Utrecht
The Netherlands.
E-mail: h.j.roelfsema@uu.nl
HU Business School of the Utrecht University of Applied Science
Research Group for International Business and Innovation
Padualaan 101
3508 AA Utrecht
The Netherlands.

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 09-23

Regional determinants of FDI in China: A new approach with recent data

Martijn Boermans[b]
Hein Roelfsema[ab]
Yi Zhang[a]

[a]Utrecht School of Economics
Utrecht University

[b]HU Business School
Utrecht University of Applied Sciences

September 2009

Abstract

We empirically investigate the factors that drive the uneven regional distribution of foreign direct investment (FDI) inflows to China's 31 provinces from 1995 to 2006. The aim of this paper is to explain the investment patterns in (partly) foreign funded firms across these provinces. We use factor analysis and derive four factors that may drive FDI: institutions, labor costs, market potential, and geography. The factor analysis then structures our dataset to concentrate on these four clusters consisting of 42 province specific and time -varying items. Factor analysis not only helps us to identify the latent dimensions which are not apparent from direct study, but also facilitates econometrics with reduced number of variables. We apply fixed effects panel estimation and GMM to account for endogeneity. In line with theoretical predictions we find that foreign investors choose and invest more in provinces with better institutions, lower labor costs, and larger market size. Nonlinear results denote that the positive effects of infrastructure and market potential on FDI are complementary to each other, which is in line with the economic geography literature. In particular the effect of market size on FDI is larger in provinces with better institutions. Sub-sample study confirms the existences of a large disparity between East and West. In the poorer large western provinces FDI is strongly driven by the geographical factor in contrast to the east of China where institutions play a significant role to build the 'factory of the world'. Robustness tests indicate that two sub-dimensions of institutions, namely infrastructure and governance, are important to determine the location choice of FDI in China.

Keywords: FDI, China, factors analysis, regional and spatial distribution of FDI, location choice

JEL classification: F21, F23, O18, O53, R11

1. Introduction

Over the last decades, foreign direct investment has been an important engine for Chinese growth. However, there are large differences in FDI patterns across Chinese regions. For example, the five special economic zones account for 80 percent of total FDI, whereas the combined five provinces in the North-West account for only 10 percent. Moreover, regions differ in the type of FDI they attract. Urban growth centers increasingly are magnates for market seeking FDI, whereas other regions are the factory of the world. Clearly, differences in FDI patterns across regions also explain internal discrepancies in economic development.

Most papers that study Chinese FDI patterns take a traditional route of analyzing FDI from a specific theoretical angle and therefore focus on a limited number of determinants to explain the variation across regions. Some focus on geographical factors and agglomeration effects, labor costs or institutional quality. Further, as is often stressed in factor analysis, traditional empirical methods often use proxies for the underlying more general determinants that are potentially related to omitted variables, which hampers causal inference. Given these restrictions in focus and method, evidence on what explains the variation in FDI across Chinese regions is still incomplete.

But there are more identification problems in the papers that deal with FDI in China. The obvious is reverse causality, since FDI inflows affect regional characteristics. Clearly, panel analysis can deal with this effectively, but such methods are difficult with for example firm level data. If one uses aggregate data at the provincial level, for fixed effects one needs a sufficiently long period in which many things happen, whereas for random effects one ideally would like a large number of cross sectional observations. In addition, when one prefers fixed effects (for example because the Hausman test would point that way) with limited cross sectional observation (regions) one has limited degrees of freedom, which restricts the inclusion of variables, so that omitted variable bias may be rampant or at least results rely heavily on the specifications used. If both time and number of regions are limited, there is a heavy trade-off. But even when one succeeds in running fixed effects, it then is very likely to exclude many potentially important fixed factors that affects the distribution of FDI across regions, for example geographical characteristics. Clearly, with

random effects one may counter the endogeneity problems, but the omitted variable return with a vengeance.

In this paper, we aim to provide a more eclectic approach to analyzing FDI patterns and to deal with omitted variables and endogeneity problems by combining conventional empirical methods and (less conventional) factor analysis. Let us briefly explain our line of thinking, without claiming that it solves all the problems mentioned above. We use data on FDI at the provincial level for the period 1995-2006. This is a period in which FDI spread from highly concentrated in the Pearl River Delta (PRD) and hence the Guangdong region to include more coastal regions as well as recently a move to the Western and Northern provinces. Before we regress provincial characteristics for which we have theoretical priors that they are correlated with FDI, we first ask to what extent provinces actually differ in their economic and social characteristics. To this end, we perform a factor analysis where we include 42 variables common in the literature (see the next section on related literature), where the analysis shows which factors (clusters of variables) explain a large part of regional variance. Certainly we hope that a subset of factors cluster in a factor that can be related to economic theory: new economic geography, regional comparative advantage, new institutional economics and the like. We have to keep in mind that the factors are clusters of variables that change over time, although some of the variables are rather static. We have included many variables to explain a significant part of regional variance, so that we can be confident to indirectly control for many potentially omitted variables.

After that, we run traditional panel estimations where we control for endogeneity by using GMM. Broadly speaking, the following results stand out. First, institutions, comparative advantage, and market size all matter, but there are important differences with respect to coastal and inner provinces and with respect to interaction effect among these factors. However, as a single factor, differences in comparative advantage and especially labor costs seem to matter most in explaining the FDI flows between 1995-2006. Hence, from a policy perspective one may argue that the efforts to spread investments towards regions with lower labor costs have succeeded. Second, although governance and infrastructure cluster into one factor, especially infrastructure seems a precondition for comparative advantage in labor costs and market size to have a sizeable effect on FDI inflows. This calls for support of

policies that promote (massive) infrastructural projects and support for local authorities in regions where FDI is low, such as the westerns and northern provinces. Lastly, we find no strong individual effect of better governance on FDI, other than its connection with an increased supply of public goods.

The paper commences as follows. The next section discusses related literature with the aim of providing a theoretical foundation for our empirical research. Section 3 introduces the data and empirical strategy in more detail, with a special emphasis on the role of factor analysis in this paper. Following that, section 4 presents the core results. Then, section 5 performs robustness checks on the main findings. Section 6 concludes the paper.

2. Related literature

FDI inflows into China are a widely studied subject. From the academic perspective, studying FDI to China attracts great interest because flows are high – so much is happening – and by focusing on a single large country one account for many variables that would may otherwise be omitted or at least imperfectly captured. In addition, FDI inflows have created much policy debate within China because of its close links to growth diversion across regions, see e.g. Chan, Henderson and Tsui (2008).

The start of the academic debate on FDI inflows in China is related the emergence of the new economic geography literature, associated with the work of Paul Krugman, Richard Baldwin and many other leading international economists in the 1980s. The central thinking is that firm location choice involves a trade-off between making use of positive externalities that come from agglomeration and the negative effects that agglomeration has on factor costs. Given that China in the 1980 opened up to foreign capital, agglomeration was (and still is) low, it provided an ideal study ground for studying the forces of the new economic geography.

The seminal paper in this approach is Head and Ries (1996) who, controlling for geographical factors, find strong agglomeration effects in FDI decisions, concentrated in the coastal areas' export processing zones. Many would follow in their footsteps. For example, recently Amiti

and Javorcik (2008) use firm level data to analyze location decision and find effects of agglomeration and costs advantages on FDI decisions.[2] Ng and Tuan (2006) study the mainland investment decision at the provincial level of firms from Hong Kong and also find agglomeration effects outside the nearby PRD region. The paper also provides a good overview of other studies on the new economic geography in China. The main conclusion from these papers is that (market) size, the presence of other firms and infrastructure, as well as labor costs are the main determinants of explaining the spatial dispersion of FDI. With respect to FDI inflows, Sethi and colleagues (2003) explore the Dunning model related to FDI using a factor analysis. Their results based on principal components shows two important determinants of FDI, namely "regional characteristics" and "market attractiveness".

In the 1990s, there emerges a new line of thinking that is much more skeptical on the powerful effects geography and the forces of the new economic geography may have on economic prosperity. The work of Daron Acemoglu, Anver Greif, and other instead stress the importance of institutions in economic development. Taking up this point, Cole, Elliott and Zhang (2006) show that when controlling for factors such as labor costs and geography, institutional variables such as control of corruption have a positive effect on attracting FDI. Local institutions may also refer to good property right protection (Cheung & Lin, 2004) and to local absorption capacity (Fu, 2008). In general, these studies stress that local institutional conditions play an important role in attracting FDI.

A current wave is to put more emphasis on firm heterogeneity. Zhao and Zhang (2005) study different motives for source countries to become engaged in FDI to China. Where Zhao and Zhang (2005) concentrate on the macro motives (differences in labor costs, for example), Hu and Owen (2005) analyze firms level data. They show that firms from Hong Kong, Macau. and Taiwan (HMT) have different motives than firms from OECD countries. More specific, agglomeration effects are especially important for firms from OECD countries, whereas labor costs attract FDI from HMT firms. In addition Belderbos en Carree (2002) analyze investment behavior of Japanese firms in China and conclude that agglomeration effects are important

[2] With firm level data it is important to note that often they restrict the analysis to cross section only, since there are no investment patterns at the firm level recorded over time. But clearly reverse causality is a limited problem when using firm level data.

for small firms, whereas large firms pay more attention to cost advantages. For our results it is important to keep in mind that over time FDI flows are driven by the fact that firms from OECD countries enter, existing firms become more acquainted in doing business in China and may be compared to firms from HMT, and that increasingly China is 'discovered' by medium sized firms.

3. Data and Methodology

3.1 Factor Analysis

In order to identify a broad structure within dataset we perform a factor analysis. Using this method we extract and exhibit the chief core from the explanatory variables without any prejudice. The goal of the factor analysis is to study interrelationships between the 42 explanatory variables and specify a new set of (latent) variables that expresses the 'communality' among the original variables. It is widely applied in psychology, medicine, geology, biology, sociology, marketing and becoming more popular in economics and management studies (Boivin & Ng, 2006; Jöreskog, 2007).[3]

It has several advantages in our context. Factor analysis basically discerns patterns of association among the data. A complete set of interdependent relationships is examined such that the technique can describe the variability among observed variables in terms of fewer (unobserved) factors. So the data is reduced to a small set which accounts for most of the variance in the initial dataset and is translated to factors.

Most other studies have a limited set of variables, derived from a theoretical angle, whereas our study takes advantage of the diversity of various variables. In addition, factor analysis decreases the degree of correlation (multicollinearity) between independent variables by reducing the number of variables to smaller set of uncorrelated (orthogonal) factor scores. Related to the reduction of variables is another distinction of factor analysis, namely that it produces neutral determinants of FDI measures, such that we overcome the selection bias

[3] As Rummel (2008) states: "factor analysis can simultaneously manage over a hundred variables, compensate for random error and invalidity, and disentangle complex interrelationships into their major and distinct regularities… [it] divides the regularity in the data into its distinct patterns."

typical in hypothesis testing research. For instance, Easterly (2008) explains that with sufficient variables, you will always find an effect, because of problems of finding the true measures. The constructs of the factor analysis partial mitigate these types of problems. Because selection criteria in regression analysis easily leads to the conclusion that adding another variable does not add any explanatory power – conditional on the already included variables - factor analysis is unique in the sense that it a priori includes all variables. Actually for these reasons Hendry proposed to model from general to specific, however, this still cannot overcome the selection bias (Sala-i-Martin, 1997).

Many studies in economics, for example those using VAR models, rely on a few pre-selected variables instead of applying large-scale models, because of restrictive assumptions about the joint distributions of all included variables. Likewise, inclusion of irrelevant information can have costs. Factor analysis uses a common-idosyncratic decomposition such that the empirical framework is kept small. As Bouvin and Ng (2006:170) state: "factor analysis provides a formal way of defining what type of variation is relevant for the panel of data as a whole." They cite a number of macroeconomic studies that "successfully" applied factor analysis in order to reduce large datasets (see Forni et al., 2001, Stock & Watson, 2002; Bernanke et al., 2005).

The identification strategy using factor analysis is neutral and in this respect can be viewed as an eclectic way of constructing explanatory variables.[4] Moreover, factor analysis partially overcomes measurement problems. It involve an "un-measurable" dimension or corresponding latent variables that underlie them which a single variable cannot capture, unless using predetermined indices build up of scaled indicators. For instance, the choice of a specific data series for the concept economic activity is "often arbitrary to some degree"

[4] In matrix notation we have $x - m = LF + e$, where x is a vector of random variables (items) that each have an average score m, L is a vector [matrix of basis vectors] of estimated constants or the factor the established factors are the factor loadings L. Because any rotation of the solutions given by factor analysis is also a solution, understanding of factors is difficult (e.g. we rewrite: $x = LF + e$ with the covariance structure $S = LTL'+P$ st. any L can be chosen, see Jennrich (2007). In addition to this rotation issue, many different conceptualizations of factor analysis have been established for various purposes. The most broadly employed techniques are common factor analysis (exploratory and confirmatory, see also global and ecological) and principal components analysis. The approaches differ because the diagonal of the relationships matrix is replaced with communalities (here: the variance accounted for by several variables) in common factor analysis. In practice, the results from various methods are closely related (Velicer & Jackson, 1990).

(Bernanke et al., 2006). Researchers normally use a proxy which can be correlated with an omitted variable, which in turn hampers causal interference.

Unfortunately there is no unique way to identify the number of factors (Jöreskog, 2007). One commonly refer to method is the Kaiser little jiffy, which states that the number of eigenvalues of the correlation matrix that are above unity reflects the number of factors. Another way to determine the number of factors is by Cattell's scree-plot, which plots the eigenvalues against their rank and number of factors is derived from the "elbow" of the curve. Maximum likelihood procedures also have been developed, but there is always a theoretical foundation needed for the naming of factors.

In order to obtain factors, first an un-rotated factor matrix is estimated. The next step is to estimate a rotated factor matrix, which is the object of interpretation. The factor loadings measure which variables or items are involved in which factor and to what extent variations influence the factor, such that they have a similar interpretation as the correlation coefficients. The communality (h^2) displays the proportion of a variable's total variation that is involved in the patterns and thus delineates a measure of "uniqueness". It is calculated for each variable by summing up the squared factor loadings. The percent of common variance indicates how the data pattern is allocated among the different factors. The first factor or component accounts for a maximum amount of variability in the data, and each succeeding one comprises as much of the remaining variability. The observed variables are modeled as linear combinations of the factors with additional error terms (non-linear methods have been developed, e.g. Wall and Amemiya (2007).

3.2 Econometics

Taken from the National Bureau of Statistics of China, a panel dataset for 31 provinces from 1995 to 2006 is employed to examine the location choice of FDI across China. We consider the investment decision of a foreign firm in a two-stage game, which is pointed out to be an important aspect of choosing conceptually appropriate FDI variable in Navaretti and Venables (2004), by investigating two FDI related dependent variables. The number of foreign funded firms (FFE) represents the stage that firms decide whether or not to invest in

a province, while the amount of total investment of foreign funded enterprises helps to explore how these firms choose production levels if production is established. Dynamics of dependent variables are deployed in Figure 1 (Figures in Appendix).

As for explanatory variables, we derive four latent factors: institutions (F1), labor costs (F2), geography (F3), and market potential (F4), based on factor analysis which captures variability among a large number of observed variables in terms of fewer dimensions.[5] Table 1 (in Appendix) lists items and their loadings to subjective factors. The higher the loading the more variation of the item is explained by a specific underlying factor. Proportion of variation explained by each factor is presented by a pie chart in Figure 2. Specifically, in this paper we use a relatively wide concept of institutions which covers infrastructure of transportation and communication, as well as quality of government and rule of law. Although many studies are focused on the latter, our data support La Porta and others (1999) in which as an important output of public goods infrastructural quality measures government performance. (See more discussion on labor costs factor in Appendix).

Following the standard process of empirical research, we first test panel unit root and panel cointegration. Tests show that all the series are I (1) and coint egrated in the long run. With reduced number of variables from factor analysis we apply the fixed effects estimation to control for time-invariant province characteristics. Between estimation is also used to show the difference across provinces in attracting FDI on the average level. Given the potential existence of reversal causation, we then employ GMM to solve the problem of endogeneity. For example, since for the same productivity level foreign firms usually pay more to attract labor force, foreign investment may raise the local labor costs. When low labor costs help to draw more FDI, methods like the fixed effects estimation are likely to underestimate the impact of labor costs on FDI. With the assumption that current endogenous independent variables are not correlated with the future realization of the error term, internal instruments which generally satisfy instruments relevance are valid to obtain reliable estimation results. Given the first-order autocorrelation in our data, we use the lagged two years variables as internal instruments. Finally, we perform various robustness checks on

[5] We applied the two discussed criteria, namely the Kaiser little jiffy based on eigenvalues, and the Cattell scree-plot, which both indicated the use of four factors.

sub-sample study of the eastern and western China, extended factors with specific items, and alternative dependent variables.

4. Estimation results

Table 2 presents the fixed effects estimation (with time effects) and between estimation results of using both dependent variables. Explanatory variables have different effects on two stages of FDI investment. Over years, higher institutional quality and larger market size in a province attract more foreign investors to establish firms there. When the location decision is made, however, all factors are irrelevant to the yearly amount of investment. Cross provinces, all other factors except for labor costs determine both the province chosen decision of foreign investors and the amount of investment. Such results seem not very plausible. For instance, the insignificant effect of labor costs is not consistent with the fact that a large proportion of FDI to China is driven by vertical specialization. One explanation for this result is: although foreign firms choose China as host country for its low labor costs, they are less concerned about this factor when locate investment in Chinese provinces that overall have sufficient low production costs. However, the impact of labor costs is also possible to be underestimated if reversal causality is present. Not only labor costs can affect FDI, location choice of foreign firms may also change the local labor costs. Without controlling for such issue, regression of using endogenous labor costs gives biased results. In our case the second reason is more promising, because results in Table 2 show similarly downward biased effects of market size and institutions.

Taking endogeneity into account, we apply lagged explanatory variables as internal instruments and show less biased GMM estimation results in Table 3. All regressions control for time and province-specific effects. In line with theoretical predications foreign investors choose and invest more in provinces with better institutions, lower labor costs, and larger market size. Significantly negative impact of labor costs and positive impact of market potential in Columns (1) and (3) provide empirical evidence of the coexistence of vertical FDI and horizontal FDI in China. Both the magnitude and the significance level of coefficients indicate that labor costs are the most important determinant of FDI across China. Although geography seems not to be a significant FDI determinant, its impact may be captured by

other factors. For example, whether or not a province is on the coast is also represented by the preferential policy indicator in the institutional factor. Furthermore, the effect of institutions is found to be dependent on other factors like labor costs and market size. First, Column (2) shows that in the absence of good institutions the change of labor costs does not matter for attracting FDI. Since vertical FDI relies on both infrastructure and labor costs, the impact of low labor costs is more predominant in provinces with better infrastructure. Conditional on local business environment labor costs are significant to determine the production level in Column (4). Second, the positive effects of institutions and market size on FDI are complementary for each other. The effect of market size is larger when institutional quality is improved; meanwhile institutions are more important when market size is enlarged. Specifically, infrastructure is crucial for distribution of products sold in the local market, and foreign investors care more about local rule of rule if they have larger volume of local trade. Finally, provincial institutions have larger impact on attracting more foreign firms because it is the first-stage of FDI that foreign investors choose investment environment. After commencing production, institutions have to work with labor costs and market size to affect the amount of foreign investment.

Table 2: Fixed and between estimation results

	Number of FDI firms (log)		Amount of FDI (log)	
	Within	Between	Within	Between
Institution	0.1299	1.0588***	0.0148	1.1360***
	(0.0805)	(0.2038)	(0.1068)	(0.2079)
Labor costs	-0.3118	0.2251	-0.3275	0.2168
	(0.3018)	(0.1792)	(0.3961)	(0.1828)
Market	0.1480*	0.7209***	0.1023	0.7906***
	(0.0768)	(0.2089)	(0.0954)	(0.2131)
Geography	-0.0162	0.4955***	0.1724	0.4487**
	(0.1304)	(0.1558)	(0.2007)	(0.1589)

Table 3: GMM results

	Number of FDI firms (log)		Amount of FDI (log)	
	Basic	Interactions	Basic	Interactions
Institution	0.0871***	0.1339***	-0.0812	-0.0733
	(0.0287)	(0.0463)	(0.0645)	(0.0750)
Labor costs	-0.9996***	-0.7708***	-1.2290***	-1.1596***
	(0.2106)	(0.2792)	(0.4789)	(0.4471)
Market	0.2777***	0.1814***	0.2181***	0.1463
	(0.0370)	(0.0625)	(0.0856)	(0.0995)
Inst*Costs		-0.0331**		-0.0810***
		(0.0168)		(0.0313)
Inst*Market		0.0308**		0.0194
		(0.0136)		(0.0226)
Geography	-0.0455	-0.0481	0.0023	-0.0150
	(0.0743)	(0.0741)	(0.1272)	(0.1134)
Jointly significance		All significant		All significant
Endogeneity test (null: exogenous)	p: 0.0368	p: 0.0773	p: 0.0000	p: 0.0000

5. Robustness

Given the huge geographic and economic disparities between Chinese eastern and western parts, we further explore regional distribution of FDI by sub-sample study. Then specific items of factor 1 are incorporate into regressions for robustness. Tests results also confirm that our basic findings hold for various dependent variables.

5.1 East and West

For sub-sample study we group provinces into East (13 provinces: Beijing, Fujian, Guangdong, Hainan, Hebei, Heilongjiang, Jiangsu, Jilin, Liaoning, Shandong, Shanghai, Tianjin, Zhejiang) and West (18 provinces). Although these two groups have same common factors which are institutions, labor costs, market, and geography, they have different factor structures (see Appendix). Therefore, we generate factors for the east region and the west region, respectively. Figure 12 demonstrates the dynamics of each factor over time for both the east and west. With similar trend of development, eastern provinces have advantage in

better institutions and larger market size. The labor costs are initially lower in the west of China but increase dramatically with a steeper slope in recent years.

GMM results in Table 4 indicate that the east and the west have different comparative advantages to attract FDI. Although labor costs are important in both regions, foreign firms located in the west are driven by geographical factor like natural resources while institutions and market potential have large impact on FDI in the east. In the east better institutions facilitate vertical FDI, which attract more foreign investors and induce them to increase investment level. Local market size is not significant in Column (2) to attract foreign firms in western provinces, because foreign investors may produce in the west and transport goods produced to the east either for larger market or exports. However, once production has been set up, larger local market raises the production level to meet the existing local demand. Interestingly, negative effect of institutions on the amount invested shows that foreign investors may give incentives to local governments to provide better institutions by increasing the amount of investment. If local institutions are already very good, they do not have to invest extra money to enhance it.

Table 4: Sub-sample study

	Number of FDI firms (log)		Amount of FDI (log)	
	East	West	East	West
Institution	0.4805***	0.0010	-0.1791***	-0.0225
	(0.1150)	(0.0574)	(0.0651)	(0.0760)
Labor costs	-1.2669***	-0.9329**	-1.7975***	-1.0220***
	(0.2317)	(0.4276)	(0.3187)	(0.3828)
Market	0.3999***	0.1020	0.3645***	0.6259**
	(0.0642)	(0.2407)	(0.0905)	(0.2752)
Geography	-0.0081	-0.3762***	0.1920	-0.6926***
	(0.0869)	(0.1121)	(0.1325)	(0.1641)

Table 5 illustrates interactive effects of FDI determinants. When foreign investors choose the west to produce for domestic trade, in the absence of good institutions such as good transportation and communication labor costs in the west have minor effect on attracting FDI in Column (2). After the location is chosen, the amount of investment is affected by market potential and labor costs in both east and west regions. The negative within-sample effect of initial institutional quality on incentives to the local government through FDI is

again found in the east (the second stage), which is opposite to the positive impact of good institutions on attracting FDI firms (the first stage).

Table 5: Sub-sample study with interactions

	Number of FDI firms (log)		Amount of FDI (log)			
	East	West	East		West	
Institution	0.0730	0.5998	0.0828	-0.0981	-0.0850	0.2574
	(0.0916)	(0.4601)	(0.0874)	(0.1196)	(0.0915)	(0.1801)
Labor costs	-1.1494***	0.1046	-1.1232***	-1.5981***	-1.0111***	-0.3723
	(0.3278)	(0.8979)	(0.3551)	(0.4590)	(0.3433)	(0.5854)
Market	0.2097**	0.8341	0.1238	0.4682***	0.4552**	0.8378**
	(0.1038)	(0.6025)	(0.0922)	(0.1061)	(0.2295)	(0.3449)
Inst*Costs	-0.0879**	-0.1287	-0.2069***		0.0569	
	(0.0387)	(0.1028)	(0.0472)		(0.0400)	
Inst*Market	0.0538	-0.1985		-0.0635		-0.1214
	(0.0480)	(0.1499)		(0.0597)		(0.0739)
Geography	-0.0755	-0.3381**	0.0990	0.2249*	-0.6794***	-0.6512***
	(0.0824)	(0.1395)	(0.1145)	(0.1282)	(0.1544)	(0.1501)
Jointly Significance	Institution, Labor costs, Market size significant	Labor costs significant; Institution and Market not	Labor costs, institutions significant	Market size, institutions significant	Labor costs significant; institutions insignificant	Market size significant; institutions insignificant

5.2 Specified factor institutions

Factor analysis indicates that factor 1 comprises variation of infrastructure and that of governance variables (Table 1.1 in Appendix). Table 6 shows detailed information on institutional impact by analyzing the two sub-dimensions of factor 1.

First, interactions in Column (2) imply that the significantly positive impact of institutions on the number of FDI firms comes from the aspect of infrastructural quality. Given the significant position of China's domestic and overseas vertical integration, foreign investors are more concerned with local transportation and communication. On the contrary, if the locational choice has been made, investment and production level are more influenced by the quality of government and rule of law. It is reasonable that governance especially plays a great role in the second stage of FDI, since property rights protection and contract enforcement environment are crucial to alleviate externalities, such as inefficient production level caused by hold-up problem, in joint production. Finally, in Column (4) we find an unexpected interaction from governance and market. Contrary to a complementary relation

between institutions and markets in basic results, market power and governance are substitute for each other in coordinating economic activity. When market has sufficient power to tackle with the hassles in contracts, external enforcement from the government loses its importance.

Table 6: GMM results of specified Factor 1

	Number of FDI firms (log)		Amount of FDI (log)	
	Basic	Interactions	Basic	Interactions
Infrastructure	0.1569***	0.2102**	-0.0609	0.1306
	(0.0482)	(0.1026)	(0.0770)	(0.1349)
Governance	0.1366	-0.0280	0.2380	0.2189
	(0.1558)	(0.3085)	(0.2462)	(0.3688)
Labor costs	-0.9343***	-0.7935**	-1.1894***	-1.1823***
	(0.2305)	(0.3239)	(0.3479)	(0.4068)
Market	0.2441***	0.4337**	0.1609*	0.5882**
	(0.0557)	(0.1934)	(0.0944)	(0.2529)
Infra*Costs		-0.1225*		-0.1254
		(0.0744)		(0.0872)
Gov*Costs		-0.1997		-0.1119
		(0.1707)		(0.2043)
Infra*Market		0.0089		-0.0106
		(0.0227)		(0.0337)
Gov*Market		-0.0993		-0.2079***
		(0.0637)		(0.0766)
Geography	-0.0561	0.0185	-0.0358	0.0534
	(0.0923)	(0.1256)	(0.1377)	(0.1585)
Jointly significance		All significant but Policy		All significant but infrastructure

5.3 Alternative FDI variables

Table 7 shows GMM estimation results of using various FDI dependent variables which are FDI inflows, registered capital of foreign funded firms, number of people employed by FDI firms, and a factor based on all FDI related variables. Effects of labor costs and market potential are consistent across all panels. However, the impact of institutions depends on the choice of dependent variable. First, the quality of institutions has different impact on different stages of FDI process, which cannot be reflected by using variables like FDI inflows. Second, more complex nonlinear relation between institutions and FDI is expected. If foreign firms strategically react to local institutional quality by providing incentives to local government, we find insignificant or even negative relationship between institutional quality and shares of foreign investors.

Table 7: Alternative dependent variables

	FDI inflows	Registered Capital FFE	Employed people FFE	An overall FDI factor
Institution	-0.2053	-0.0611	0.1767***	0.4890***
	(0.2907)	(0.0389)	(0.0413)	(0.0779)
Labor costs	-4.2849**	-1.2621***	-0.3590*	-1.2185***
	(1.9532)	(0.2928)	(0.2153)	(0.4487)
Market	0.35240*	0.2447***	0.1263***	0.4430***
	(0.1884)	(0.0511)	(0.0484)	(0.0879)
Geography	-0.2888	-0.0549	-0.2001**	0.2090**
	(0 .2540)	(0.0958)	(0.0834)	(0.0983)

5.4 Looking deeper into regional comparative advantage

Given loadings of items in factor analysis, we identify factor 2 as labor costs which account for both productivity and wage. Curves of factor 2 on wages and labor productivity in foreign related firms support this argument. Graphic results indicate that labor costs are jointly determined by wages and productivity. First and not surprisingly, Figure 3 shows a negative relationship between labor costs and productivity. Moreover, in Figure 4 the increase of productivity may dominate the growth of wages in the low wage level, and therefore factor 2 (labor costs) decreases with wages. When wages are high, however, the effect of wages outweighs that of productivity and causes high labor costs. Finally, the similar dynamics of our factor 2 and unit labor costs manufacturing index of China by Dullien (2005) in Figure 5 further prove that categorizing factor 2 as labor costs is convincing.

Looking at variables loaded to classify our factor 2, we find that productivity is represented by education, and more interestingly, by different levels of education. Specifically, basic education (primary and junior high school education) and high education (senior high school and higher education) have different paths to affect labor productivity. Figure 6 and Figure 7 show that high education enhances efficiency in production, whereas basic education has negative or insignificant effect on productivity. Workers with higher education are able to use physical capital more efficiently and their capability to absolve and imitate new techniques allows for further improvement in productivity. However, such positive role of higher education may not be observed for basic education in China. First, low-educated people are hard to exert impact on technical progress by innovation. Second, since low-

efficient state-owned firms pay more to workers with low education, unskilled workers prefer to move out of non-stated-owned firms (Yue 2003; Zheng & Hu 2004). Low-educated workers in foreign related firms lack incentives to put efforts into production. Meanwhile, education has impact on labor costs through wages, which is illustrated in Figures 8 and 9. Wage compensation increases with high education, while a complex U-shape relationship between basic education and wages exists. The possible reason for such nonlinear relationship is that the negative effect of labor endowment on wages first dominates when the pool of labor (with minimum required skills) is small, but with more workers available it is replaced by the positive correlation between wages required and average education level. The overall effects of education on factor 2 are shown in Figures 10 and 11.

6. Concluding Remarks

In this paper we have analyzed recent FDI inflows in China at the provincial level. Our approach has been eclectic. Informed by a literature that stresses many variables which are correlated with FDI flows, we run a factor analysis to establish unbiased regressors for which Chinese provinces differ. Broadly speaking, on top of geographical fixed factors, regions differ in labor costs, market potential, and hard and soft institutions. We then perform a 'horserace' among these factors to see which factors matter most. We show that for the 1995-2006 period, labor costs and infrastructure (and especially when combined) are important for attracting FDI.

These results fit against a background of FDI diffusion away from the Pearl River Delta towards the Shanghai and Beijing region. Increasingly the Beijing region is able to capture a larger share of FDI by effectively tapping into cheap labor from the inner provinces. On top of that, it reflects a shifting towards inner provinces, especially by firms from Taiwan and Hong Kong. For these firms, cost advantages are important assets in competitive world markets, so that they shift to cheaper northern and western location when infrastructure is ready.

Our study certainly does not contradict the relations found in other papers. A main difference is that we focus on a time frame where the Chinese government has changed

course and the coastal regions became relatively less attractive for foreign investors. After setting up the export processing zones, the Chinese government in the 1990s has made great strides to diffuse FDI. First, this succeeded towards the other eastern provinces. However, according to very recent figures, economic growth is now higher in the northern and western provinces. In addition, our empirical findings indicate that over time improvements in infrastructure, or keeping labor costs low are becoming more important.

Can we draw lessons for the ongoing policy debate on the relative importance of geography, big push development, and institutions? Clearly, we have to be cautious here. However, from our analysis it becomes clear that geography is not all important if big push efforts in infrastructure are made. Foreign investors do not stick to location and agglomeration effects are not that strong that the inhibit the dispersion of FDI across regions. In addition, in China soft institutions (such as differences in local corruption and education) do not seem to play an important role other that they tend to go together with 'hard' institutions such as infrastructural improvements. This calls into question to what extent institutional reform alone in China as well as in other parts of the world is able to create FDI flows.

However, the analysis may also point to a more critical observation, one that is shared in much of the management literature on investing in China. In the data, there is the suggestion that labor costs and logistics remain the most important driving for foreigners to invest in China. This may also be because higher valued activities are still seen as too risky. The obvious reason is a lack of property rights protection, so that assembly based on higher skills (and, hence, higher labor costs and more schooling) remains unprofitable for foreign firms in the long run. A second reason is a lack of local management skills to perform integrated system production processes. Lastly, there is a often heard complaint that in joint ventures, ailing domestic firms are pushed by local politicians for inclusion in joint venture production. All these issues suggest that the dominant strategy for foreign firms still is to make use of cheap and disciplined labor, so that the next step towards high value added production is jet to come.

References

Amiti, M. and Javorcik B.S. (2008). Trade Costs and Location of Foreign Firms in China. *Journal of Development Economics, 85*(1), 129-149.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica, 70*(1), 191–221.

Belderbos, R. and Carree, M. (2002). The Location of Japanese Investments in China: Agglomeration Effects, Keiretsu, and Firm Heterogeneity. *Journal of the Japanese and International Economies, 16*(2): 194-211.

Bernanke, B., Boivin, J. and Eliasz, P. (2005). Factor augmented vector autoregressions (FVARs) and the analysis of monetary policy. *Quarterly Journal of Economics, 120*(1), 387-422.

Chan, K.W.J., Henderson, V. and Tsui, K.Y. (2008). Spatial Dimensions of Chinese Economic Development. In T.G. Rawski and L. Brandt (Eds), *China's Great Economic Transformation*. Cambridge: Cambridge University Press.

Cheng, L.K. and Kwan, Y.K. (2000). What Are the Determinants of the Location of Foreign Direct Investment? The Chinese Experience. *Journal of International Economics, 51*, 379-400.

Cheung, K. and Lin, P. (2004). Spillover Effects of FDI on Innovation in China: Evidence from the Provincial Data. *China Economic Review*, *15,* 25-44.

Cole, M., Elliott, R. and Zhang, J. (2006). Corruption, Governance and FDI Location in China: A Province-Level Analysis. *Department of Economics, University of Birmingham, Discussion Papers.*

Dullien, S. (2005). China's Changing Competitive Position: Lessons from a Unit-Labor Cost-Based REER. *International Trade 0502016*, *EconWPA*.

Du, J., Lu, Y. and Tao, Z. (2008). Economic Institutions and FDI Location Choice: Evidence from US Multinationals in China. *Journal of Comparative Economics, 36*, 412-429.

Easterly, W. (2008). Can the West Save Africa? *NBER Working Papers, 14363*.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2001). Coincident and leading indicators for the euro area. *Economic Journal, 111*, 82–85.

Fu, X. (2008). Foreign Direct Investment, Absorptive Capacity and Regional Innovation Capabilities: Evidence from China. *Oxford Development Studies, 36* (1), 89-110.

Head, K. and Ries, J. (1996). Inter-city Competition for Foreign Investment: Static and Dynamic Effects of China's Incentive Areas. *Journal of Urban Economics, 40*(1), 38-60.

Hu, A.G. and Owen, R.F. (2005). Gravitation at Home and Abroad: Regional Distribution of FDI in China. mimeo.

Jennrich, R.I. (2007). Rotation Methods, Algorithms and Standard Errors. In R. Cudeck and R.c. MacCallum (Eds.), *Factor Analysis at 100: Historical Development and Future Directions*. Lawrence Erlbaum.

Jöreskog, K.G. (2007). Factor Analysis and Its Extensions. In R. Cudeck and R.c. MacCallum (Eds.), *Factor Analysis at 100: Historical Development and Future Directions*. Lawrence Erlbaum.

La Porta, R, Lopez-de-Silanes, F., Shleifer, A and Vishny, R. (1999). The Quality of Government. *Journal of Law, Economics and Organization, 15*(1), 222-279.

Liu, T. and Li, K. (2006). Disparity in Factor Contributions between Coastal and Inner Provinces in Post-reform China. *China Economic Review, 17*, 449-470.

Navaretti, G.B. and Venables, A.J. (2004). *Multinational Firms in the World Economy.* Princeton University Press.

Ng, L.F. and Tuan C. (2006). Spatial Agglomeration, FDI, and Regional Growth in China: Locality of Local and Foreign Manufacturing Investments. *Journal of Asian Economics, 17*, 691-713.

Rummel, R.J. (2008). Understanding Factor Analysis. mimeo.

Sala-i-Martin, X. (1997). I Just Ran Two Million Regressions. *American Economic Review, 87*(2), 178-83.

Sethi, D.S., Guisinger, E., Phelan, S.E. and Berg, D.M. (2003). Trends in Foreign Direct Investment Flows: A Theoretical and Empirical Analysis. *Journal of International Business Studies, 34*(4), 315-326.

Stock, J.H. and Watson, M.W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association, 97*, 1167–1179.

Velicer, W. F. and Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, *25*(1), 1-28.

Wall, M.M. and Amemiya, Y. (2007). A Review of Nonlinear Factor Analysis and Non Linear Structural Equation Modelling. In R. Cudeck and R.c. MacCallum (Eds.), *Factor Analysis at 100: Historical Development and Future Directions*. Lawrence Erlbaum.

Yue, C.J. (2003), Does Higher Educated People Earn More Money in the Labor Market in China? 4th International Conference on the Chinese Economy, The Efficiency of China's Economic Policy, 23-24 October.

Zheng, J.H. and Hu A.G. (2004), An empirical analysis of provincial productivity in China, 1979-2001. CCS Working Paper Series, Tsinghua University, N° 1, February 26.

Zhao, H. and Zhu, G. (2008). Location factors and country-of-origin differences: An empirical analysis of FDI in China. *Multinational Business Review*.

Appendix

Table 1: Rotated factor loadings

Variables	Factor1	Factor2	Factor3	Factor4	Factor5	Uniqueness
Capital	0.8434	0.2703				0.0847
City road length	0.7277	0.3637				0.1491
City road area	0.8298	0.3185				0.0976
Civil vehicle	0.8964	0.2998				0.0541
freight	0.6493	0.4972				0.2225
Gov Expenditure	0.9243					0.0400
Private vehicle	0.9004					0.0924
Ways (train, water, highway)	0.5331	0.4195		-0.3019	0.3236	0.2627
Exports	0.8666				-0.3747	0.0566
Imports	0.8010			0.3547	-0.3805	0.0712
Long telephone	0.9185					0.0741
Local telephone	0.9012	0.2832				0.0447
Mobile	0.9650					0.0415
Cable	0.6272			-0.3505	0.4000	0.1804
Patent registered	0.8882				-0.2647	0.0678
GRP per capita	0.5124	-0.3228		0.7091		0.0666
Wage	0.6450	-0.3610		0.5300		0.0816
Consumption household	0.5207	-0.3316		0.6983		0.0865
Tech market transaction	0.3244			0.8662		0.1183
Population	0.3394	0.8953				0.0301
Workers	0.3422	0.7798				0.0738
Primary school		0.8872				0.1078

Primary enrolment		0.9106				0.0539
Junior high school		0.9356				0.0562
Junior enrolment	0.3376	0.8573				0.0906
Senior high school	0.4405	0.8006				0.0837
Senior enrolment	0.6481	0.5298			0.4634	0.0622
Higher education institutions	0.6470	0.4633		0.2672	0.3995	0.0961
Higher education enrolment	0.7497				0.5354	0.0606
Humidity			0.9032			0.1321
Sunshine		-0.3354	-0.7836			0.2208
Temperature			0.8711			0.1514
Area			-0.4013	-0.2550		0.2192
Precipitation			0.8404			0.2164
Natural resource			-0.5645			0.2023
NERI index	0.6980		0.3581	0.4004		0.1099
Index property protection	0.3781			0.5757		0.1521
Index government intervention			0.4912		0.3469	0.2344
Index corruption						0.2553
Index contract enforcement						0.2229
PPI (Preferential Policy Index)	0.4026	-0.2636	0.4611			0.2464
Minority population				-0.2814		0.3903
(blanks represent abs(loading)<0.25); No. of observations: 309						

Factor rotation matrix

	Factor1	Factor2	Factor3	Factor4
Factor1	0.8758	0.3779	0.1495	0.1940
Factor2	-0.2418	0.8030	0.0956	-0.4952
Factor3	-0.1533	-0.0529	0.9568	0.1405
Factor4	0.2811	-0.3312	0.1500	-0.3551

Summary

Variable	Obs	Mean	Std. Dev.	Min	Max
Factor1	309	-2.03e-10	1	-1.0979	7.0008
Factor2	309	1.07e-09	1	-1.6994	3.1234
Factor3	309	-5.12e-11	1	-2.0398	2.2674
Factor4	309	1.18e-09	1	-1.5178	8.2555

Table 1.1: Specified Factor 1 (Institutions)

Variables	Factor 1 (Infrastructure)	Factor 2 (Governance)	Uniqueness
City road length	0.8459	0.2241	0.2343
City road area	0.9151		0.1328
Freight	0.7893		0.3644
Ways	0.6738	-0.5293	0.2658
Long telephone	0.9272		0.1395
Local telephone	0.9661		0.0625
Mobile	0.9367		0.1221
Cable	0.6854	-0.5204	0.2594
Patent	0.8509	0.2669	0.2047
NERI index	0.7276	0.4880	0.2324
Index property protection	0.3223	0.8060	0.2465
Index government intervention		-0.2028	0.9474
Index corruption		-0.3620	0.8376
Index contract enforcement		-0.3276	0.8921
PPI (Preferential Policy Index)	0.3601	0.5579	0.5590
Minority population		-0.6655	0.5433
(blanks represent abs(loading<0.2) ; No. of observations: 309			

Figures

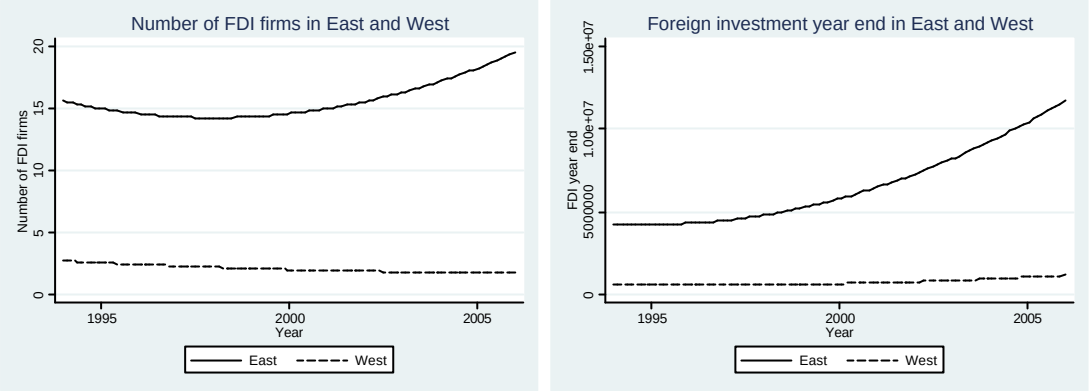

Figure 1

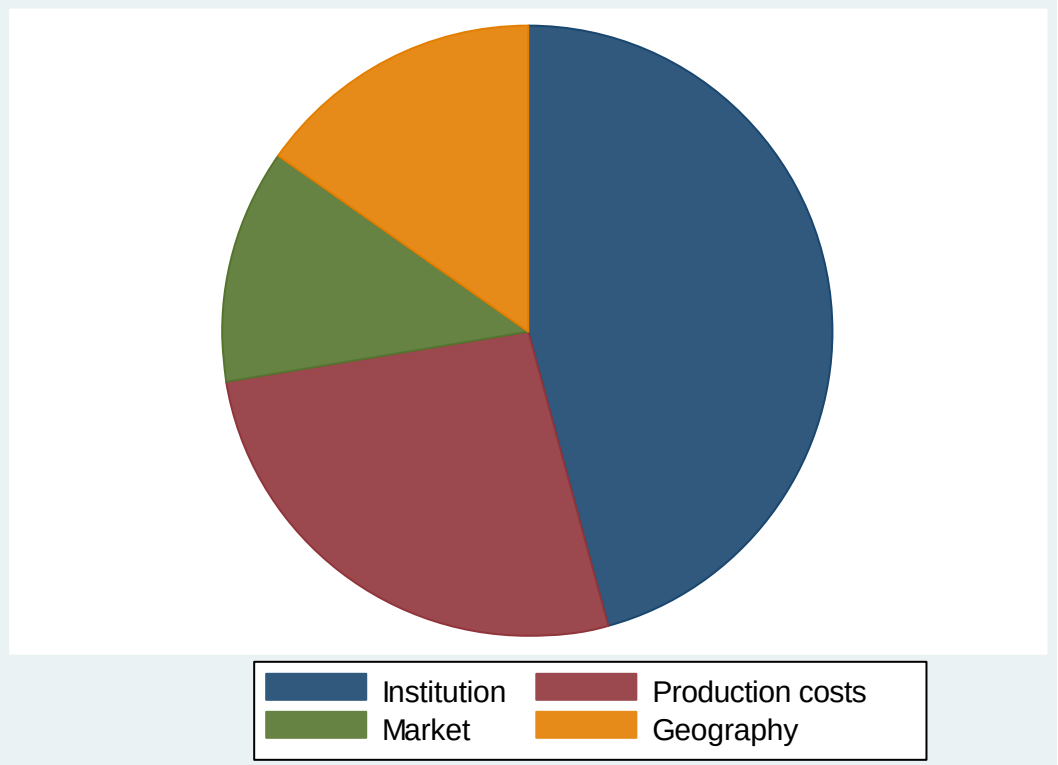

Figure 2

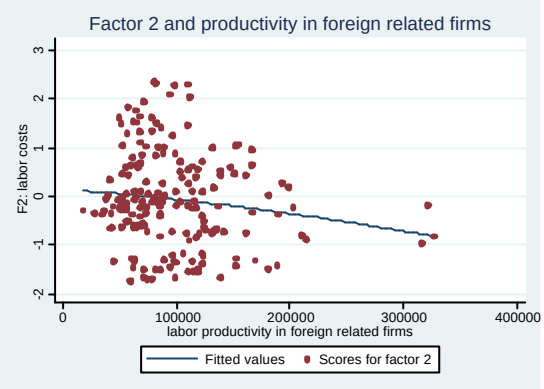

Figure 3

Figure 4

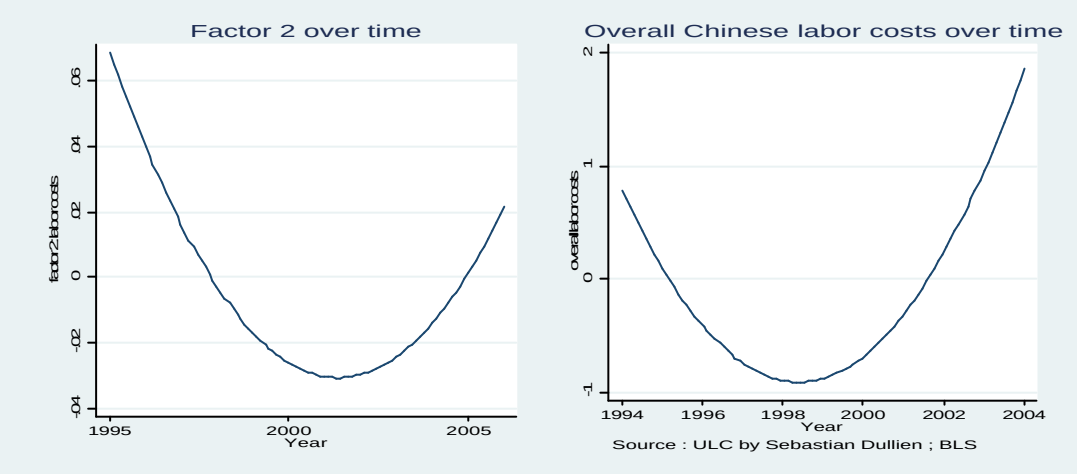

Figure 5

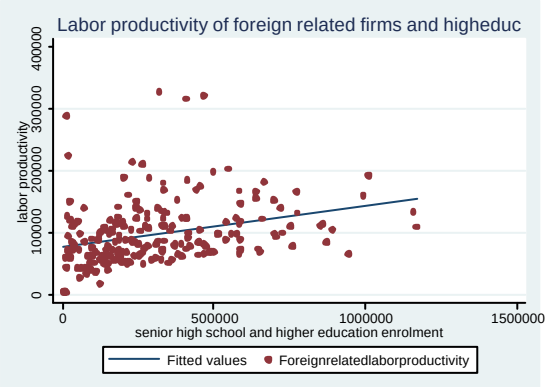

Figure 6

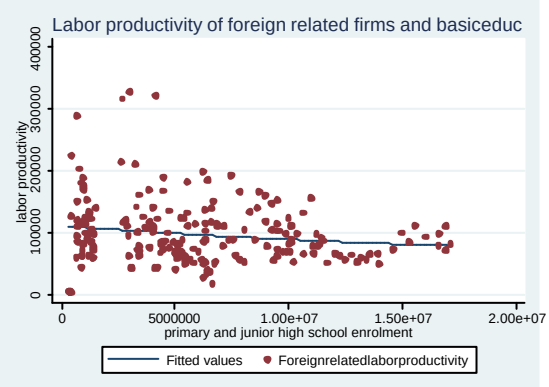

Figure 7

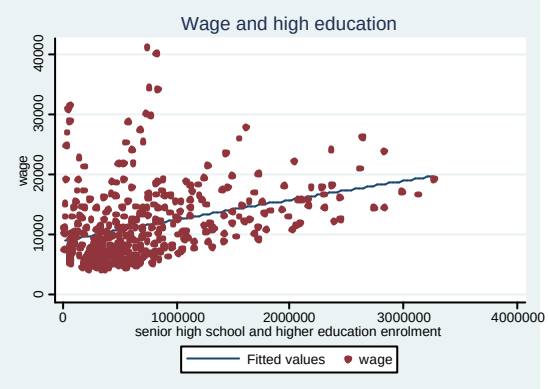

Figure 8

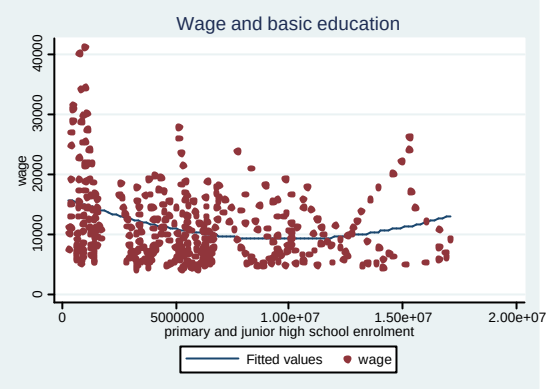

Figure 9

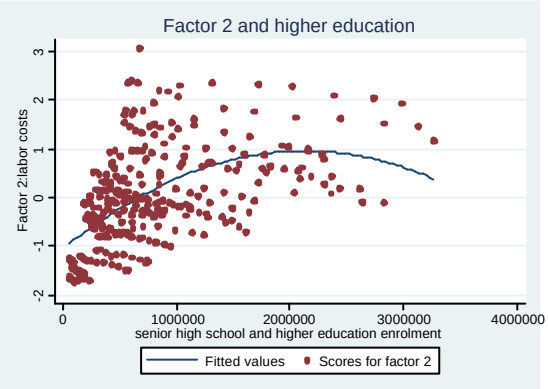

Figure 10

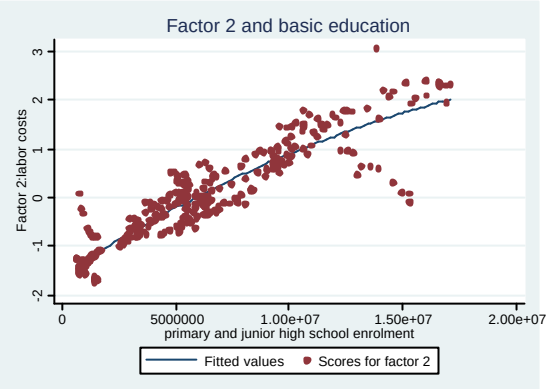

Figure 11

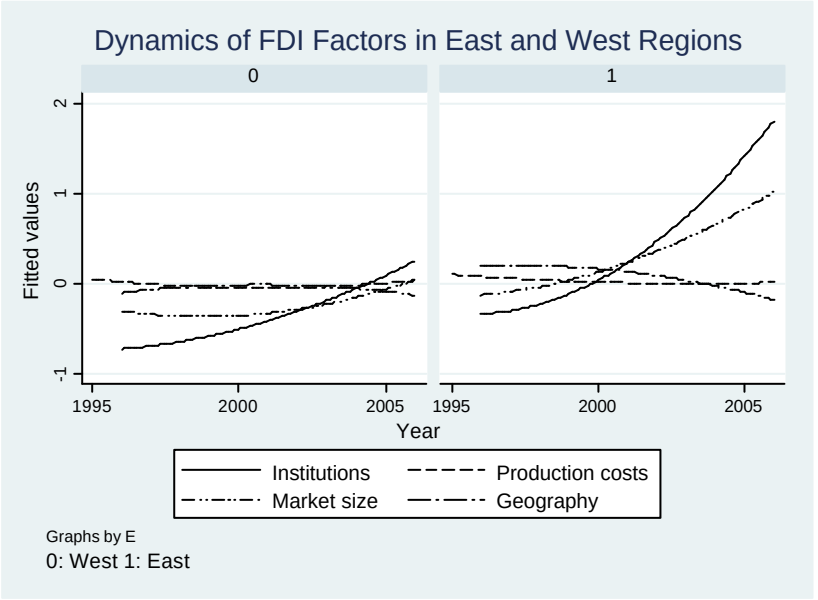

Figure 12

Factor loadings for East and West

East factor loadings

Variables	Factor1 (Institution)	Factor2 (Labor costs)	Factor3 (Geography)	Factor4 (Market)	Uniqueness
Capital	0.9014				0.0364
City road length	0.7869	0.4016			0.1479
City road area	0.8739	0.3353			0.0857
Civil vehicle	0.8800	0.2947			0.0535
freight	0.7104	0.5449			0.1464
Gov Expenditure	0.9114			0.3266	0.0400
Private vehicle	0.8769				0.1077
Ways (train, water, highway)	0.7030	0.4712			0.1293
Exports	0.7600		0.3924	0.3642	0.0727
Imports	0.6878		0.2720	0.5682	0.0817
Long telephone	0.8191	0.2666	0.2683		0.0676
Local telephone	0.9428				0.0340
Mobile	0.9223				0.0528
Cable	0.7595				0.1213
Patent registered	0.7924		0.2984	0.3854	0.0835
GRP per capita	0.3681	-0.3982		0.7145	0.0831
Wage	0.5804	-0.3901		0.5958	0.1023
Consumption household	0.3490	-0.3946		0.7471	0.1037
Tech market transaction	0.2609			0.7002	0.3003
Population	0.5085	0.8169			0.0208
Workers	0.3093	0.8234			0.1487
Primary school		0.9266			0.0886

Primary enrolment	0.3074	0.9055			0.0325
Junior high school	0.2911	0.8819		-0.2740	0.0457
Junior enrolment	0.4260	0.8379			0.0624
Senior high school	0.5482	0.7470			0.0947
Senior enrolment	0.8083	0.4569			0.0339
Higher education institutions	0.8060	0.2868	-0.3074		0.1093
Higher education enrolment	0.9416				0.0240
Humidity			0.8913		0.1597
Sunshine			-0.8873		0.1289
Temperature			0.8408		0.0912
Area					0.0536
Precipitation			0.9041		0.1532
Natural resource			-0.5162		0.1185
NERI index	0.6475		0.4314	0.3714	0.1169
Index property protection				0.8301	0.2452
Index government intervention	-0.2533	-0.7239		-0.3336	0.2374
Index corruption		-0.3441		-0.2800	0.1798
Index contract enforcement			0.4524		0.4063
PPI (Preferential Policy Index)			0.8479		0.2400
Minority population				-0.3741	0.2801
(blanks represent abs(loading)<0.25; No. of Observations: 140					

West factor loadings

Variables	Factor1 (Institution)	Factor2 (Labor costs)	Factor3 (Geography)	Factor4 (Market)	Uniqueness
Capital	0.8729	0.3794			0.0764
City road length	0.5027	0.4227	0.3272		0.2014
City road area	0.7058	0.3835			0.1344
Civil vehicle	0.8171	0.4928			0.0710
freight	0.5353	0.5522	-0.3391		0.2605
Gov Expenditure	0.9570			0.4348	0.0286
Private vehicle	0.8529	0.3209		0.2600	0.1075
Ways (train, water, highway)	0.7524	0.2546			0.2706
Exports	0.8671			0.3604	0.1763
Imports	0.8747			0.4511	0.1690
Long telephone	0.8649	0.3172			0.0923
Local telephone	0.8874	0.3696			0.0546
Mobile	0.9495				0.0686
Cable	0.8684				0.1595
Patent registered	0.7116	0.5163			0.1209
GRP per capita	0.8303	-0.2528		0.8090	0.1362
Wage	0.4457	-0.4569		0.8750	0.1450
Consumption household	0.3420	-0.2663		0.8099	0.1124
Tech market transaction	0.4699	0.3682			0.1464
Population	0.3432	0.8557	0.2789		0.0538
Workers		0.9145			0.0763
Primary school		0.9091			0.1265
Primary enrolment		0.8898	0.2772		0.0928

Junior high school	0.3148	0.9022			0.0528
Junior enrolment	0.4425	0.7839			0.1108
Senior high school	0.4819	0.7839	0.2582		0.0705
Senior enrolment	0.8294	0.4214			0.0897
Higher education institutions	0.7182	0.5103	0.2692		0.0960
Higher education enrolment	0.8527				0.1185
Humidity		0.3939	0.7823		0.1340
Sunshine		-0.4917	-0.6630		0.2242
Temperature		0.4023	0.6798		0.1752
Area		-0.2586			0.1514
Precipitation		0.2506	0.6839		0.3050
Natural resource					0.1643
NERI index	0.6963		0.2500	0.4456	0.1350
Index property protection				-0.3013	0.4122
Index government intervention			0.8231		0.2445
Index corruption				0.5701	0.2125
Index contract enforcement					0.2319
PPI (Preferential Policy Index)	0.5371		0.2755		0.5290
Minority population		-0.2535			0.1156
(blanks represent abs(loading)<0.25; No. of Observations: 169					