Market access impact on individual wages: evidence from China

Laura Hering* and Sandra Poncet†

Abstract

We study the effect of geography and in particular of market access on wages by working with individual data from 56 Chinese cities in 11 different provinces. By applying the theory of the New Economic Geography on individual survey data, we contribute to the explanation of growing disparities within the country, and even within provinces. We examine to what extent proximity to markets can explain inter-individual wage heterogeneity and growing wage disparities within Chinese provinces. Using a New Economic Geography style model, we derive an econometric specification relating wages to market access. The latter is calculated as a transport cost weighted sum of the surrounding locations' market capacity. Based on data from 1995 on around 10,000 Chinese workers, and after controlling for individual skills and factor endowments, we find that a significant fraction of inter-individual differences in terms of return to labor can be explained by the geography of access to markets. Moreover, our study investigates whether the relationship between market access and wages holds for all types of workers equally and shows that the magnitude of the impact depends on the firm type and the level of qualification.

JEL Codes: F12, F15, R11, R12.

Keywords: Economic geography, International Trade, Regional Integration, Wage, China, Inequality.

*Corresponding author: Panthéon-Sorbonne-Economie, Université Paris 1. Address: Maison des Sciences Economiques, Bureau 405, 106 bld de l'Hôpital 75013 PARIS. Tel: +33 1 44 07 82 42 Fax : +33 1 44 07 82 47. Email: laura.hering@gmail.com

†We would like to thank Thierry Mayer, Rodrigo Paillacar, Jean-Louis Arcand and Agnès Bénassy-Quéré for their comments. We have also benefited from remarks from participants at the LICOS China Seminar as well as at the Final CEPR Conference of the Research Training Network "Trade, Industrialisation and Development".

1 Introduction

Since the foundation of the People's Republic of China in 1949, income inequality has always been an important and sensitive political issue. But although the country is still influenced by Maoist egalitarian ideology, its unprecedented income growth over the last two decades has come with large and still increasing income level differences within China (Meng et al., 2005). Coastal provinces are known for their wealth and development, inland regions for their backwardness and poverty. A core-periphery structure of the income gradient across regions, i.e. regions with low per capita income are predominantly located at the geographical periphery while the richest are at the center, can already be observed in many countries for a long time. Today this structure is also clearly apparent in China (Lin, 2005). Most studies and policies on spatial inequalities concentrate on the gap between eastern and western provinces. The literature on heterogeneity within these regions has concentrated mainly on the contrast between rural and urban areas (Wei et al, 2002). But with the ongoing growth of intra-provincial disparities, researchers and policy-makers have become increasingly interested in the origins of this new phenomenon.[1]

A number of reasons may prevent convergence of income levels. Combes et al. (2005) note that three broad sets of explanations are frequently invoked to explain large spatial wage disparities: spatial differences in the skill composition of the work force, differences in non-human endowments, and differences in interactions among workers or among firms. While controlling for the first two mechanisms, this paper aims at explaining the intra-national and intra-provincial disparities by one particular strand of literature of this last argument, which is the New Economic Geography (NEG) theory.

Theoretical developments and growing empirical evidence in the field of NEG theory (Head and Mayer, 2004a; Redding and Venables, 2003 and 2004; Hanson and Xiang, 2004) explain the emergence of a heterogeneous economic space on the basis of increasing returns to scale and transport costs (Krugman, 1991 and Krugman and Venables, 1995). One central proposition of NEG theory is the importance of the proximity to consumers, represented by the region's "market access" which is usually defined as the distance-weighted sum of the market capacity of surrounding locations (Fujita et al., 1999). The NEG "wage equation" models nominal wages as a function of the region's market access. According to this relationship, wages are predicted to be higher at the economic center while the periphery is confronted with lower wages. The idea is that the closer a particular location is to consumer markets, the lower are the transport costs and the fiercer the demand linkages. This leads to higher economic activity and thus to a higher employment and/or wage level.

In the case of a change in the demand of goods of a particular location, the structural model of NEG proposes

[1] See for example Lee (2000) for a work on spatial disparities in Jiangsu; Weng (1998) in Guangdong or Lyons (1998) in Fujian.

two different mechanisms by which the local economy can adjust to the new situation: either the adjustment is achieved through a change in prices, which means that wages will go up with a higher market access (Redding and Venables, 2004), or the adjustment is quantitative. In the latter case, the level of employment adapts to the new demand but MA stays unchanged (Davis and Weinstein, 1999). Consequently, we can expect that any change in MA will translate into a change in wages.

In this paper, we look at the price channel by applying the NEG wage equation to the Chinese economy.[2] China is an interesting country for testing the wage equation since migration is officially restricted and the urban labor market is strictly segmented into urban and rural workers.[3] A change in demand is therefore much likely to lead to an adjustment by wages. So if a city faces an increase in its demand, an adjustment by the means of the level of employment would not be possible and wages have to go up.[4] Consequently, great variations across Chinese cities in terms of magnitude and rate of increase of market access are likely to fuel the current growth in wage differentials.

By applying NEG theory on individual survey data, we can describe this transmission mechanism and thereby contribute to the explanation of growing disparities within the country, and even within provinces.

Recent empirical work on the impact of market access on nominal wages validates the hypothesis of a positive relationship between these variables, mainly focusing on developed countries. Major contributions include Feenstra and Hanson (2000) on US counties, Mion (2004) for Italy, De Bruyne (2003) for Belgium, Brakman et al. (2004) for 114 German districts and Head and Mayer (2006) and Breinlich (2006) for the European Union. However, since individual skills are known to account for a large fraction of existing spatial wage disparities, these studies, which all rely on aggregate data, fail to control for cross-regional variation in skill composition.[5] This exposes them to the risk of wrongly attributing wage disparities to economic geography explanations (Combes et al., 2005).[6]

Because of the use of individual data, this paper can go further than most of the other studies in this domain: It

[2] For a detailed description of both adjustment channels, see Head and Mayer (2006).

[3] For more details on the Chinese labor market, see Bannister (2005).

[4] Although the absence of jobs in the countryside forces each year millions of people to abandon agriculture and to migrate to richer parts of the country, mainly to towns, to find work in urban industries (Ping and Pieke, 2003), their impact on urban wages might be negligible: The labor market is strictly segmented and impedes migrants to have access to urban jobs or social security (Bannister, 2005).

[5] Head and Mayer (2006) try to take into consideration the skill composition through a human capital augmented version of the wage equation.

[6] Because workers with better unobservable characteristics are more likely to be located in more central areas, wage disparities across areas tend to be wrongly attributed to agglomeration economies. This would induce an upward bias in the estimated coefficient on MA. Also, beside a direct effect, remoteness may hamper the accumulation of human capital, aggravating the wage disadvantage of peripheral regions (Redding and Schott, 2003). As argued by Breinlich (2006), this is the case if intermediate and transport cost-intensive goods use relatively more skilled labor. Then, more central locations will offer higher wages for skilled labor, which increases the incentives for human capital accumulation.

is one of the first to account for individual worker heterogeneity (in terms of education, age and sex among others) in NEG models.[7] This permits us to control for skill composition before investigating whether the wage of workers responds to the economic geography of their location.[8] Also, this approach can alleviate the potential problem of endogeneity of market access since it is less likely that a shock on an individual wage translates into a change in local market access.[9]

Our analysis uses a household survey conducted by the Chinese Academy of Social Sciences (CASS), which provides detailed information on labor income as well as on individual and household characteristics for the year 1995. Using data on around 10,000 Chinese workers from 56 cities in 11 provinces, we find that a significant fraction of inter-individual differences in terms of return to labor can be explained by the geography of access to markets.

We confirm that individual skills account for a large fraction of existing spatial wage disparities in China. After controlling for worker heterogeneity, non-human endowments on industry level and the clustering of error terms, the elasticity of individual wages to market access is 0.14. When including additional controls for provincial specificities with province fixed effects, the elasticity of individual wages to market access decreases to 0.06. On average, a one standard deviation over mean increase in market access of a prefecture would lead to an average increase in individual wages by 5.6%.[10]

Having found a positive and significant impact of market access on wages also within provinces, we can conclude that intra-provincial wage disparities could be partly due to the big differences in the cities' market access. Growing differences in trade costs or market sizes between cities can therefore increase wage disparities within provinces.

Our study moreover investigates whether the relationship between MA and wage holds for all types of workers equally, or whether this relationship is valid only for certain firm types or skills profiles. It emerges that high skilled workers profit from a higher MA, while wages for unskilled workers do not seem to be affected by any change in MA. Also, only wages earned in certain firm types react to changes in MA: While wages in private firms react strongly, wages in state-owned enterprises are largely insensitive to them.

The paper is organized as follows. The next section outlines the theoretical framework from which the econo-

[7] Paillacar (2006) is the only paper to our knowledge that uses individual wage data and explores the implications of workers' individual characteristics in the determination of wages. It applies to Brazilian data.

[8] The literature on market access has so far addressed this issue by resorting to geography-based instruments of market access and/or relying on regional fixed effects regression.

[9] Hanson (2005) argues that when the dependent variable is at the finest level possible, shocks in the error term will be less likely to affect the dependent variable. Moreover, if the explanatory variables are more aggregated, endogeneity is again less likely since shocks to individual variables affect regional variables only slightly.

[10] The value of 5.6% is obtained by multiplying the coefficient of $lnMA$ with the ratio of the standard deviation of MA over its mean.

metric specifications used in the subsequent sections are derived. It presents the two key relationships that are estimated in this paper, namely the "trade equation" and the "wage equation". Section 3 describes the data and calculates market access measures for Chinese regions. Section 4 investigates to which extent wages in China respond to these measures and undertakes several robustness checks. Section 5 provides a conclusion.

2 Theoretical framework

The theoretical framework underlying the empirical analysis is a reduced version of a standard New Economic Geography model of monopolistic competition based on Dixit and Stiglitz (1977), similar to the one used by Fujita et al. (1999) and Redding and Venables (2004).

We consider a world with R locations, composed of firms that operate under increasing returns to scale and produce differentiated manufactured products. Consumers' utility increases with the number of varieties. Demand for differentiated products is modeled in the usual symmetric constant elasticity of substitution way with σ ($\sigma > 1$) being the elasticity of substitution between any pair of products.

Final demand for goods in location j is derived via utility maximization of the representative consumer's CES utility function.[11] Country j's demand for a variety produced in r is:

$$demand_{rj} = p_{rj}^{-\sigma} \frac{E_j}{G_j^{1-\sigma}} \tag{1}$$

where E_j is location j's total expenditure on manufactured goods and p_{rj} is the price of varieties from location r sold in j (consisting of the mill price p_r and iceberg transportation costs T_{rj} between the two locations: $p_{rj} = p_r\ T_{rj}$). G_j is the aggregate price index for manufactured goods, $G_j = \left[\sum_{r=1}^{R} n_r\ p_{rj}^{1-\sigma}\right]^{1/1-\sigma}$, with n_r the number of firms in r. Taking into account that T_{rj} units must be shipped in order for one unit to arrive, we obtain the effective demand x_{rj} faced by a firm in r from location j :

$$x_{rj} = T_{rj}\ p_{rj}^{-\sigma}\ G_j^{\sigma-1}\ E_j = T_{rj}^{1-\sigma}\ p_r^{-\sigma}\ G_j^{\sigma-1}\ E_j \tag{2}$$

As put forward by Redding and Venables (2004), the own price elasticity of demand is σ, and the term $G_j^{\sigma-1}\ E_j$ gives the position of the demand curve faced by each firm in market j. It is referred to as the "market capacity" of country j. It corresponds to local expenditure E_j adjusted for the "market crowding" effect G_j, summarizing the number of competing firms and the prices they charge.

[11] For a complete statement of the underlying model, see Fujita et al. (1999).

Intuitively, a larger number of competitors and thus a lower value of G_j will reduce the attractiveness of j for any firm exporting there.

Equation (2) highlights that trade costs influence demand more when the elasticity of substitution is high. We follow the literature in referring to $\phi_{rj} = T_{rj}^{1-\sigma}$ as the "phi-ness" of trade (see Baldwin et al., 2003). It can take a value between 0 (when trade costs are prohibitive) and 1 (when trade costs are negligible). Summing over all products produced in location r, we obtain the "trade equation" (Redding and Venables, 2004).

The value of total exports of region r to region j is therefore:

$$n_r p_r x_{rj} = n_r p_r^{1-\sigma} \phi_{rj}\, G_j^{\sigma-1}\, E_j \tag{3}$$

As emphasized by Redding and Venables (2003), this equation for bilateral trade flows provides a basis for the estimation of a gravity trade model. While the last term in the right hand side of equation (3) contains the "market capacity" of region j, $m_j = G_j^{\sigma-1}\, E_j$, the first term, $n_r p_r^{1-\sigma}$, measures what is referred to as the "supply capacity" of the exporting region, $s_r = n_r p_r^{1-\sigma}$. It corresponds to the product of the number of varieties with their price competitiveness.[12]

Equation (3) of bilateral trade flows, which will serve as the basis of the gravity equation estimated in Section 4, can therefore be expressed simply as:

$$n_r p_r x_{rj} = s_r \phi_{rj}\, m_j \tag{4}$$

Turning to supply, we follow standard assumptions of the Dixit-Stiglitz-Krugman (DSK) model. Differentiation costs of varieties are supposed to be so small that each variety is produced by only one firm.

Increasing returns at the firm level come from the combination of a plant-specific fixed cost, f_r, and a marginal cost of production, c_r, which are region-specific (Head and Mayer, 2004a). The cost of producing q_r in each region is assumed to take the form $c_r q_r + f_r$. Each firm maximizes the following gross profit function in each market. The gross profit function of region r on each market j is therefore: $\pi_{rj} = (p_r - c_r)\, T_{rj}\, q_{rj}$. The resulting mill price for each origin r is a simple mark-up over marginal costs:

$$p_r = \frac{c_r \sigma}{\sigma - 1} \tag{5}$$

All varieties produced in a given region r are thus valued at the same price (before transport costs). The gross

[12] Redding and Venables (2003) discuss the concepts of market and supply capacity in greater depth.

profit earned in each market j for a variety produced in region r is given by $\pi_{rj} = p_r\, x_{rj}/\sigma$.

Substituting in equation (5), summing up the profits earned in each market and subtracting the fixed costs, f_r, we obtain the net profit to be earned in each potential location r:

$$\Pi_r = \sum_j p_r\, x_{rj}/\sigma - f_r = \lambda\, c_r^{1-\sigma} \sum_j \left[\phi_{rj}\;\; G_j^{\sigma-1}\, E_j\right] - f_r \tag{6}$$

where $\lambda\; = \frac{1}{\sigma}(\frac{\sigma}{\sigma-1})^{1-\sigma}$. Following the literature we refer to

$$\sum_j \phi_{rj}\, G_j^{\sigma-1}\, E_j = \sum_j \phi_{rj}\, m_j = MA_r \tag{7}$$

where MA_r is the "market access" of region r. This is simply the sum of the market capacities of all destinations j, m_j, weighted by the measure of bilateral trade costs, ϕ_{rj}, between r and j. It summarizes how well a location is endowed with access to markets for the goods it produces. The literature on New Economic Geography highlights that firms in locations with higher market access incur less overall transportation costs and are able to pay higher wages (Fujita et al., 1999).[13]

We adopt the standard Krugman (1980) assumptions that labor is the only factor and there is both a fixed and a variable component to firm-level labor requirements. Free entry of firms forces firm economic profits to be zero at the equilibrium. It implies that the equilibrium production quantity of each firm in a given region r is equal to $q = \frac{f_r(\sigma-1)}{c_r}$, so that the total regional production is $q_r = n_r q$.

We modify the model to allow skills to be different for each individual i. For each region r we assume that labor requirement, ℓ_r, depends on both output per region and the workers' skills, z_i:[14]

$$\ell_r = (\alpha + \beta q_r) \exp(-\rho z_i), \tag{8}$$

where α and β measure fixed and variable labor requirements in "effective" (skill-adjusted) labor units. The

[13] As pointed out by Head and Mayer (2006), the "market access" bears a close resemblance to Harris (1954)'s "market potential". The difference lies in the fact that Harris' market potential implicitly treats G_r, the price index, as constant and ϕ_{rj} is approximated with $1/distance_{rj}$. In this sense, the MA_r is real, not nominal, since it incorporates the notion that large markets that are extremely well-served by existing firms might offer considerably less potential for profits than smaller markets with fewer competitors around them.

[14] This hypothesis implies that regional labor force is not proportional to the number of enterprises, as the original version of the model has held. Regions with an abundance of skilled workers are expected to have a smaller number of employees.

parameter ρ corresponds to the return to skills.[15] These assumptions imply a fixed cost $f_r = \alpha \exp(-\rho z_i) w_i$ and a marginal cost $c_r = \beta \exp(-\rho z_i) w_i$, where w_i is the wage rate.

Replacing f_r and c_r in the profit function by their expressions, we derive what Redding and Venables (2004) call the "wage equation".

$$\lambda(\beta \exp(-\rho z_i) w_i)^{1-\sigma} MA_r = \alpha \exp(-\rho z_i) w_i \tag{9}$$

Rearranging and replacing MA_r by its expression, we obtain our second key relationship that will be estimated in the next section:

$$w_i = \left[\sum_j \phi_{rj} \, G_j^{\sigma-1} \, E_j \frac{\lambda \beta^{1-\sigma}}{\alpha} \right]^{1/\sigma} \exp(\rho z_i) \tag{10}$$

In this equation, we can see the two different ways by which a region r can adjust to a shock, for example an increase in its local demand, E_r: either the number of firms and workers goes up, which is reflected in a change in the price index, G_j, (quantity adjustment). In this case, the adjustment takes place inside MA since G_r compensates the change in E_r and the total market access does not change. Or the adjustment is a price adjustment, where the number of firms and workers remains unchanged and MA therefore increases. The higher demand drives prices up and is to be compensated by an increase in wages to ensure that the zero-profit condition is maintained.

With the use of survey data on the individual level and the incorporation of skills and other individual characteristics in our model, we try to take into account some basics of labor economics, which have been very much neglected before in the NEG wage equations. It is only recently that authors of the NEG (Redding and Schott, 2003; Mion and Naticchioni, 2005) have started to incorporate features of this strand of literature in their models.

In the next section we will empirically test the impact of MA_r on w_i.

3 The Data

The aim of the empirical part of this paper is to explain the variation of Chinese individual wages by recourse to the differences in the cities' market access. In this section, we first describe the data sources of the explained variable, wages at the individual level, as well as the independent variables of our model. After that, we work out in detail

[15] In the empirical section, we will proxy z_i with information on gender, age, education level and the number of years of experience at the individual level i. In this latter case, ρ will show the percentage increase in productivity from an extra year of experience. We will explore potential diminishing returns to the impact of skills (such as education, experience...) and allow quadratic terms in the empirical section.

how our variable of interest, the market access, is constructed.

3.1 Dependent variable

Our database is from the 1995 survey of the China Household Income Project (CHIP). The data set was collected in 1996 by a team headed by the Institute of Economics, Chinese Academy of Social Sciences (Riskin, Zhao and Li, 2001). It covers 6,931 households and 21,694 individuals in urban China[16] across 56 cities in 11 provinces.[17] The sample we use in this study is composed of around 10,000 workers, aged between 16 and 60, for which wage and basic characteristics such as sex, activity and education are available.[18] Table B1 (Appendix B) reports descriptive statistics on the wage variable.

The wage variable w, expressed in yuan, is provided directly by the survey where it is defined as the sum of the basic salary, bonuses, allowance, subsidies (housing, medical, child care and regional subsidies) and other wages (overtime wages and wages for special circumstances).[19] As a robustness check, we will verify that our results are not driven by differences in working hours. We will rely as the explained variable on hourly wages calculated as the ratio between wages and the number of declared hours worked per week.

3.2 Independent variable: Market access

In our analysis, the explanatory variable of particular interest is market access, expressed as in equation (7), for each city c, yielding: $MA_c = \sum_j \phi_{cj} G_j^{\sigma-1}\ E_j$. Since neither the market access itself nor its components G_j and E_j are directly observable, we have to estimate these two variables first and then use the obtained estimates to calculate the market access. This two-step procedure has been pioneered by Redding and Venables (2004) and is detailed in Head and Mayer (2004b).[20]

Redding and Venables (2004) suggest that the unobserved country-specific variable $(G_j^{\sigma-1}\ E_j = m_j)$ could be obtained from the estimation of the gravity equation described in equation (4) using importer fixed effects.

Taking the natural logarithms in equation (4) yields the basic econometric specification used for the trade equation, so the total value of exports to region j from all firms based in region r is given by:

[16] The Chinese Household Income Project is a joint research effort sponsored by the Institute of Economics, Chinese Academy of Social Sciences, the Asian Development Bank and the Ford Foundation. Additional support was provided by the East Asian Institute, Columbia University (Riskin, Zhao and Li, 2000).

[17] The sample includes the following provinces: Beijing, Shanxi, Liaoning, Jiangsu, Anhui, Henan, Guangdong, Yunnan, Sichuan, Gansu and Hubei. Refer to Map 1 in the Appendix A for localization.

[18] We loose about half of the total available data sample because 1) the individuals are not aged between 16 and 60, 2) no wage information is given or 3) the place of residence is not provided.

[19] Studies examining earnings differentials based on this data include among others Démurger et al. (2005).

[20] These authors refer to the market access as the real market potential. The reader is referred to that paper for more detail.

$$\ln(X_{rj} = n_r p_r x_{ij}) = \ln s_r + \ln \phi_{rj} + \ln m_j \tag{11}$$

Where $\ln s_r$ and $\ln m_j$ are exporter and importer fixed effects.

We follow the existing literature in calling FX_r the exporter fixed effect and FM_j the importer fixed effect. Exporter fixed effects capture the log of the exporter's supply capacity, $FX_r = \ln n_r - (\sigma - 1) \ln p_r$ while importer fixed effects correspond to the market capacity of the importer region j, $FM_j = \ln(G_j^{\sigma-1}\, E_j)$. Thus, for the calculation of MA_r, we will need to predict FM_j.

Another important information retrieved from the estimated gravity equation is the freeness of trade between each pair of partners.[21]

In the next subsection (3.2.1) we estimate the gravity equation, using values of all bilateral trade flows and distances involving Chinese provinces as well as various other countries. Parameter estimates of this regression are then used to compute market access for each Chinese city (Subsection 3.2.2).

3.2.1 Estimation of the trade equation

For the estimation of the trade equation, we rely on several data sources to build our bilateral trade flows data set covering intra-provincial, inter-provincial, international and intra-national trade flows.[22]

This approach is necessary in order to estimate the market capacities, m, of international and national trading partners as well as transport costs, ϕ, which will be used to compute the market access of Chinese cities. Since the most disaggregated level of bilateral trade data is at the provincial level, we will first estimate the trade equation on international and domestic trade flows of Chinese provinces and international countries to predict their respective market capacities.

Intra-provincial trade flows for China and the intra-national trade flows for foreign countries are computed, following Wei (1996), as domestic production minus total exports.

All trade flows are merged into a bilateral international trade flows data set covering 29 Chinese provinces and around 200 countries of the rest of the world (ROW). The estimate of equation (12) based on this complete data set will allow us to compute market capacities of Chinese provinces and foreign countries based on their exports to all destinations (domestic and international).

[21] By using the gravity equation rather than simply assuming bilateral transport costs to depend in a particular way on bilateral distance, we follow Head and Mayer (2006).

[22] See appendix C for details on data sources of trade flows and production indicators for Chinese provinces and international countries used to estimate the trade equation.

According to previous studies on border effects of Chinese provinces, we allow for impediments to domestic trade to be different from impediments to international trade (Poncet, 2003).[23] Therefore, in our gravity equation, transport costs are assumed to depend on bilateral distances and a series of dummy variables indicating what type of border is crossed.

Allowing for border effects to vary according to trading partners, equation (12) yields the following trade regression, where B denotes dummies:

$$\ln X_{rj} = FX_r + FM_j - \delta \ln dist_{rj} - \varphi \, B^{foreign}_{rj,\ r\ or j \in\ China} - \varphi^* \, B^{foreign}_{rj,\ r\ \&j \in\ ROW} \qquad (12)$$

$$+\psi Contig_{rj} \, B^{foreign}_{rj} + \vartheta \, B^{provincial}_{rj,\ r \neq j,\ r\ \&j \in\ China} + \xi \, B^{intranational}_{rj,\ r=j,\ r\ \&j \in\ ROW} + \zeta_r$$

Literally this equation says that we allow for differentiated transport costs depending on whether trade occurs between a Chinese province and foreign countries ($-\delta \ln dist_{rj} - \varphi + \psi Contig_{rj}$), between two foreign countries ($-\delta \ln dist_{rj} - \varphi^* + \psi Contig_{rj}$), between a Chinese province and the rest of China ($-\delta \ln dist_{rj} + \vartheta$), within foreign countries ($-\delta \ln dist_{rj} + \xi$) or within Chinese provinces ($-\delta \ln dist_{rj}$). In these last two cases, only internal distance affects trade freeness. Accessibility of a Chinese province or a foreign country to itself is modeled as the average distance between producers and consumers in a stylized representation of regional geography, which gives $\phi_{rr} = distance_{rr}^{-\delta} = (2/3\sqrt{area_{rr}/\pi})^{-\delta}$, where δ is the estimate on distance in the trade equation.

Being neighbors dampens the border effect ($Contig_{rj} = 1$ for pairs of partners that are contiguous).

Equation (13) is estimated for the year 1997[24], yielding country/region-specific estimates to construct Chinese provinces' market access. Results appear in Table B2 in Appendix B.[25] The coefficient of distance is similar to those found in related literature, the impact of contiguity as well. We confirm findings from Poncet (2003) that the border effect inside China is important. Furthermore we find impediments to trade to be greater between China and the rest of the world than between countries included in our sample (which are mostly members of the WTO and are therefore much more integrated in the world economy than China in the 1990s).

---

[23] This author finds domestic and international border effects of Chinese provinces to be around 27=[exp(3.30)] and 400=[exp(6)], respectively in 1997. Refer to Poncet (2003 and 2005), for more on the existence, level and evolution of impediments to inter-provincial trade flows in China.

[24] As explained in the appendix C, data on inter-provincial trade flows is limited to the year 1997. For the sake of coherence, we decided to rely on trade flows for a single year, thus 1997, to compute estimates used to calculate provincial market access. Section 4 will regress individual wages for 1995 on market access for 1997. We argue that the associated time discrepancy should not be a problem because of the relative time persistence of market access. The potential problem of reverse causation between wage for 1995 and market access for 1997 will be addressed in the empirical section.

[25] Importer and exporter fixed effects are included in the regression so that the border effect within foreign countries ($-\delta \ln dist_{rj} + \xi$) is captured by their fixed effects. The reference category in the regression is within Chinese province trade.

3.2.2 Market Access computation

The data set on wages gives the location of each individual by reporting his/her province and prefecture level or county level city of registration.[26] To compute market access of the cities, we follow Head and Mayer (2006)'s allocation rule. FM_j, the estimated province-level market capacity $m_j = G_j^{\sigma-1} E_j$, is then allocated to cities c inside the province j according to the GDP share[27] of each constituent city:

$$m_c = G_c^{\sigma-1} E_c = (y_c/y_j)m_j = (y_c/y_j)G_j^{\sigma-1} E_j = (y_c/y_j) \exp(FM_j) \tag{13}$$

Note that while the lack of sub-provincial trade data forces us to choose an allocation rule for provincial competition-weighted expenditure m, the other component of market access, ϕ, uses genuine city-level information.

Finally, we can compute each city's market access which consists of four components, corresponding to the four sums in equation (14): local market access, provincial market access (from all other cities in the same province), national market access (from all other Chinese provinces), and rest of the world (ROW) market access. Formally, the market access of city c to all regions/countries j (including itself) is given by:

$$\begin{aligned} MA_c &= \phi_{cc}G_c^{\sigma-1} E_c + \sum_{j \in Province} \phi_{cj}G_j^{\sigma-1} E_j + \sum_{j \in China} \phi_{cj}G_j^{\sigma-1} E_j + \sum_{j \in ROW} \phi_{cj}G_j^{\sigma-1} E_j \\ &= dist_{cc}^{-\delta} (y_c/y_j)\exp(FM_j) + \sum_{j \in Province} dist_{cj}^{-\delta} \exp(\vartheta) \exp(FM_j) \\ &+ \sum_{j \in China} dist_{cj}^{-\delta} \exp(\vartheta) \exp(FM_j) + \sum_{j \in ROW} dist_{cj}^{-\delta} \exp(\varphi_{+}\psi Contig_{rj}) \exp(FM_j) \end{aligned} \tag{14}$$

where FM_j and the parameters δ, ϑ, φ and ψ are estimated in the trade equation and $dist_{cj}$ are great circle distances between c and j.

Table B3 (Appendix B) provides some information on the composition of each city's market access. For the 56 cities in our data set, national market access is by far the dominant part of overall access. On average, the national component accounts for 78% of the total market access. The final column displays the ratio of each city's market access to the highest (Foshan, Guangdong province). The market access of the last ranking city (Guangyuan, Sichuan province) in our sample is only 4% of Foshan's.

Figure 1 in Appendix A graphs the market access of the 56 cities as a function of their average wage.[28] As

[26] The entire country is divided into 27 provinces plus four "supercities" with province-status – Beijing, Chongqing, Shanghai and Tianjin. In each province (or supercity), the population is further divided into prefecture level cities and lower level rural counties and cities. Our data set on urban individuals consists of information on individuals in prefecture level cities or county cities.

[27] The GDP shares are obtained from the China's State Statistical Bureau.

[28] Average wage is computed based on the household database as the average of individual workers' wages.

is apparent from the graph, high market access is found for high-wage cities which is in line with the theoretical prediction of the NEG model based on the wage equation.

Figure 2 in Appendix A graphs market access as a function of their average distance to the two main cities of the country (i.e. Beijing and Guangzhou). As expected, high market access is found for more central cities.

3.3 Individual variables and variables of control

Individual workers' information include age, gender, educational level, number of years of schooling as well as number of years of work experience. We introduce as control variables in our regressions dummy variables to account for the 9 types of ownership[29] as well as for the 13 sectors[30] in the data.

4 Empirical Results for the wage equation

4.1 Wage equation - baseline specification

Having calculated market access at the city level, MA_c, we can now run the regressions of our human capital augmented version of the wage equation for around 10,000 individuals i, located in 56 cities c in 11 provinces. Tables B4 and B5 (in Appendix B) show the correlation matrix and the summary statistics of all variables used in this section.

Taking the natural logarithms of equation (11), the econometric baseline specification becomes:

$$\ln w_{ic} = a + b \ln MA_c + \rho z_{ic} + \epsilon_{ic} \tag{15}$$

where $a = -(\frac{1}{\sigma}) \ln \left[\alpha\sigma(\frac{\sigma}{\sigma-1})^{\sigma-1}\beta^{\sigma-1} \right]$ and $b = \frac{1}{\sigma}$.

As discussed by Head and Mayer (2006), the intercept, a, depends on the input requirement coefficients α and β. These are likely to vary across sectors, in part because of variations in capital intensities. For this reason we will estimate equation (15) with industry and later also industry-province specific constants. Our benchmark estimates are obtained by OLS. To ensure that our estimation results are not driven by outliers we rely on the method of Hadi (1994) for all regressions to identify multiple outliers in multivariate data.

The structure of our data confronts us with the problem of clustering of errors. It is to be expected that observable and unobservable characteristics of the various wage earners within the same location and industry are

[29] Ownership types cover state owned enterprises (SOEs) at central or provincial level, local publicly-owned enterprises, urban collective enterprises, private enterprises, individual enterprises, foreign-invested enterprises, Sino-foreign joint ventures, township and village enterprises and others.
[30] Sectors include agriculture, industry, mining, construction, transport and communication, commerce, social welfare, public, culture, research, finance and insurance, government and others.

correlated. At the statistical level, the issue is that the variance of our errors is no longer spherical and failure to account for this will lead to biased estimates of standard errors and erroneous inferences. Moulton (1986, 1990) emphasizes that the typical OLS measures of variance could understate errors by a potentially large factor, leading to poor inferences.

In this paper we correct for clustering using the Moulton correction. We therefore correct for the correlation of errors between individuals within a specific province and industry.

A final econometric problem arises from the two-step calculation of the market access. This variable is calculated with parameters that have been estimated with standard errors in a first regression. We correct for biased standard errors by applying the "bootstrap" procedure on each of our regressions.

Before we regress the individual wages on our calculated market access, we look at our basic wage determinants: school years, experience and gender, which are expected to explain a large proportion of existing spatial wage disparities. The first column of Table 1 reports the regression of wages over these three variables plus industry dummies. All variables have a highly significant impact on wages, explaining 16% of wage differences between individuals. The coefficient on school years seems quite small in comparison with findings from studies on wage determinants in other countries, but the magnitudes are similar to other studies on China (Démurger et al., 2005).

Adding market access as an independent variable in Column 2 leads to an increase by 4 percentage points in the R^2, which means that market access significantly contributes to the explanation of wage differentials. While the other parameters stay similar to the first column, the estimator of MA_c is positive and significant. Its magnitude is comparable to results found by Paillacar (2006) for Brazilian individual wages.

Column 3 compares the impact of our market access to that of the Harris market potential, defined as $\sum_j E_c/dist_{cj}$. This measure of market potential displays a higher estimated coefficient than our market access, which is consistent with existing literature (Head and Mayer, 2006). A further test of robustness consists in dividing our market access into its original components: Column 4 repeats the regression from Column 2, but instead of using total market access it uses only non local market access (excluding market access of the city), Column 5 relies on non provincial market access, excluding market access of the whole province and Column 6 shows only foreign market access. We find similar coefficients for all parts of the MA. These results confirm that our estimates in Column 2 do not depend on a single component of the market access. Since these variables are determined simultaneously, they are highly correlated, which prevents us from estimating the specific impact of each of the components of MA on wages either through their simultaneous inclusion in the wage regression or through the estimation of separate regressions.[31]

In Columns 7 and 8 of Table 1, we repeat the first two estimations but include a dummy for the provincial capital. By doing so, we follow Brülhart and Koenig (2006) who analyze employment and wage structures in Eastern Europe. They formulate a "Comecon hypothesis", according to which wages are not systematically related to market access except for capital regions. We could expect a similar pattern in China since this country is considered as a socialist regime until today. The positive and significant coefficient of this dummy shows that in China, workers in a provincial capital city earn more than in the rest of the province. The impact of MA_c remains positive and significant, which suggests that the estimated significant impact of MA on wages does not only result from a capital city premium: wage sensitivity to MA is not restricted to the capitals but applies to all cities. Consequently, we can reject the "Comecon hypothesis" for China.[32]

4.2 Instrumentation

Our baseline specification (15) is clearly subject to a potential simultaneity problem. Market access, on the right hand side of the estimated equation, is a weighted sum of all potential expenditures including local ones. Those expenditures depend on incomes, and therefore on wages, raising a concern of a reverse causality in the estimation. As explained by Head and Mayer (2006), a positive shock to w_i will raise E_c and consequently increase MA_c.

A promising approach to this simultaneity problem with MA_c is to isolate variations in market access that can reasonably be assumed to be exogenous to potential shocks on wages. Geographic variables seem to be the most suitable candidates for such an instrumental variables estimation. A good instrument for MA_c is i) a variable that is not influenced by wages but impacts MA_c, ii) a variable that does not enter the wage equation directly.[33] Therefore, following Head and Mayer (2006), we use measures of "centrality" of each city as instruments. These instruments are obtained by dividing the surface of the globe in approximately 11,700 squares. The "national centrality" of each city is the sum of the inverse distances to the center of all other Chinese squares, j, with a population density above 2 persons per square kilometer: $ln \sum_{j \in China} dist_{cj}^{-1}$. The "international centrality" of each city is the sum of the inverse distances to all non Chinese squares, j, with a population density above 2 persons per square kilometer: $ln \sum_{j \in ROW} dist_{cj}^{-1}$. This measure does not incorporate any information on market

[31] In this latter case, the coefficient on the included component would capture the impact of the omitted components of MA.

[32] The capital city being normally the biggest city of a province, this dummy captures also the size which might have as well an influence on wages. We will discuss this issue in the Section 4.4.

[33] Redding and Venables (2004) use the distance to the nearest central place (Brussels, New York City, or Tokyo) as instrument for the market potential of each country in the world. As argued by Head and Mayer (2006), although physical geography variables of this kind seem indeed to be instruments meeting the criteria above, the choice of the reference points raises an endogeneity issue since these three cities are chosen because of their extremely high market potential.

sizes of regions, which are affected by the spatial distribution of wages.

We systematically check the validity of our instruments with the Hansen's J test of overidentifying restrictions. Insignificant test statistics indicate that the orthogonality of the instruments and the error terms can not be rejected, and thus that our choice of instruments is appropriate.[34]

The next step is to perform the Wu-Hausman test, which tests for the endogeneity of the market access indicator in a regression estimated with IV.[35] Both test statistics are reported in the last two lines of the estimation tables. Since the Wu-Hausman test does not reject the null hypothesis of exogeneity of the market access (at the 10% confidence level), we report OLS estimates since they are more efficient than IV estimates (Pagan, 1984).

4.3 Within province estimations

Having shown that market access has a significant impact on Chinese wages, we want to go one step further and investigate this relationship within the provinces. We want to know if when comparing two cities of the same province, the one nearer to its markets (local, national or international) pays higher wages. If this is not the case, the impact found above only measures the differences between big cities in Guangdong and small villages in Yunnan and we would not observe a wage gradient within a province, but only over the whole country.[36] Also, this approach allows us to better control for differences in non-human endowment, as demanded by Combes et al. (2005).

Therefore, the regressions from now on are all performed with province-industry fixed effects and we also control for clustering of error terms at this level. China, being a country with significant differences between provinces in terms of policies, endowments and development, the province-industry fixed effects will take into account this heterogeneity and purge the market access impact of the effects of particularities of a province or an industry on individual wages.

Column 1 of Table 2 allows for a non-linear influence of experience and schooling through a quadratic term. Column 2 further introduces age and its square term. These two indicators display a strong and significant impact on wages.

In the last column of Table 2, we introduce additional control variables. The coefficients of the dummies for firm ownership reproduce the results found by earlier studies, in particular Démurger et al. (2005), who use the

[34] Significance is judged at the 10% confidence level.

[35] The rejection of the null hypothesis (at the 10% confidence level) that an OLS estimator of the same equation would yield consistent estimates means that endogenous regressors have a statistically relevant effect on coefficients and we have to rely on the IV estimation.

[36] In Table 1, no province dummies are introduced so that estimations focus on differences between the 56 cities whatever their province of location.

same data set as we do: Foreign and Sino-foreign owned enterprises as well as state owned firms pay higher wages than other owners.[37]

We see that even after introducing province-industry fixed effects, in all three columns of Table 2, market access remains statistically significant at the 5% level and has a coefficient of around 0.06. This means that on average, doubling market access increases average wage by approximately 6%. Since no other study known to us examines market access on individual wages while controlling for endowments at the regional and industry levels, we cannot say if this is particularly high or low.[38]

As a robustness check, we repeat the estimations of Table 2, using wage per hour as the dependent variable.[39] Results do not change much, although there is indication of a slightly stronger market access impact. Considering that with the information available in our data set, calculating exact working hours per year is not actually possible and in order to avoid losing observations, we keep using total individual wages per year as our dependent variable in the remainder of the article.

Having found a positive and significant impact of market access on wages also within provinces, we can conclude that wage disparities within the provinces could be partly due to the big differences in the cities' market access. Growing differences in trade costs or market sizes between cities can therefore increase intra-provincial wage disparities.

4.4 Disentangling size and market access effects

According to the estimates above, we could conclude that spatial variation of Chinese wages is consistent with our NEG model, based on increasing returns internal to firms. However, since our only variable at the city-level is market access, its impact could also capture features consistent with urban agglomeration theories. Hanson (2000) distinguishes three additional mechanisms to market access linking agglomeration and wages: (1) factor endowments, (2) increasing returns external to firms and (3) human capital externalities.

Concerning the first mechanism, as explained before, we continue to assume that differences in factor endowments are more likely to occur at the provincial and industry level and do not control for local differences.

The second and third mechanisms correspond to two main competing dynamics relating to agglomeration. On the one hand, the bigger and/or denser an agglomeration is (in terms of labor), the more knowledge spillovers there are between firms and workers. This leads to a higher productivity of each worker and therefore to higher wages.

[37] Estimation in Column 3 is made in reference to the group "others".

[38] Paillacar (2006) introduces only industry fixed effects, finding results similar to ours in Table 1. Mion (2004) does not test a structural model, but finds a similar coefficient as we do.

[39] For calculation of the variable see Section 3.1.

On the other hand, large agglomerations often exert a downward pressure on the price level because of tougher competition between a greater number of producers who all want to sell their products. A lower price level drives wages down. Big cities are therefore exposed to these two contradictory forces.

So far, we controlled only insufficiently for these aspects by including a capital dummy, which account for the specificity of some of the biggest and most human capital abundant cities. So it is possible that our results on the significance of MA captures the size effect caused by spillovers between firms or human capital externalities. The regressions in Table 4 introduce the natural logarithm of the local population and of the local GDP as a proxy for size, respectively, in Column 1 and Column 2. Results indicate that size significantly influences Chinese wages while keeping the impact of MA_c unchanged. Comparing two cities within a province with the same population, we find that the city with the higher market access pays higher wages.

In the third column, we introduce the city average percentage of individuals having finished upper middle school as a proxy for city-level human capital stock. Its coefficient is positive and highly significant while the coefficient of MA stays unchanged, attesting the presence of positive human capital externalities on individual wages.

In the last column, we combine the impact of city size and human capital. Also here, the impact of MA remains significant. Our results confirm the validity of NEG theory even after scale and human capital effects are accounted for.

4.5 Wage equation - augmented specification

This section investigates potential conditionality in the relationship between market access and wages depending on workers' skills and the ownership respectively. To test whether the impact of market access varies according to these characteristics, we create interaction variables between market access and educational level and certain firm ownership types, respectively.

4.5.1 Market access impact: skills matter

In the first two columns of Table 5, we introduce a dummy variable to distinguish between high skilled and low skilled workers and create two interaction variables between market access and educational level. A person is considered as high skilled if he or she has finished at least upper middle school. Otherwise the individual is classified as low skilled workers. The coefficient on the dummy variable shows that high skilled workers have higher earnings than low skilled. Furthermore, these regressions show that high skilled workers, in contrast to low skilled workers, benefit significantly, in terms of wages, from a higher city market access. Since it is possible that the impact of the other variables varies according to the local educational level, we divide the data set according to the qualification of workers and run our estimations on the two sub-samples separately. Results confirm the high sensitivity of skilled workers to market access.

Why exactly we have such a strong market access impact for one group and no impact at all for the other group cannot be easily explained by our data. An intuitive explanation for the Chinese economy derives from the fact that on the one hand, skilled workers are still relatively rare in comparison with developed countries, while on the other hand, the pool of unqualified workers is huge, so that a higher demand for low-skilled workers is more easily adjusted by a higher employment.

Another explanation could be that, even though our data set contains only urban registered persons and these are less likely to be in direct competition with the masses of illegal migrants, they might not be completely indifferent to the arrival of a great number of cheap unskilled workers in their city driving unskilled wages down.

4.5.2 Market access impact: the influence of firm ownership

In 1995, while the number of private firms started to grow rapidly, a very high portion of registered employment was still in publicly owned enterprises. When looking at the interaction between MA and firm ownership, we would expect different impacts for different types of ownership: private, recently created firms should be more flexible in terms of wages and react more to shocks on the market than state-owned. This intuition is confirmed by the results reported in Table 6. In this table, we add MA interacted with firm ownership dummies. In the first two columns, the reference group includes TVE, individual, locally-publicly owned and urban collectively owned enterprises as well as others. The wage elasticity to MA by firm ownership type can be retrieved by adding the common coefficient on MA to the coefficient on the appropriate interaction. We find that, on the one hand, wages in private firms react more strongly to changes in market access than those in other firm types. On the other hand, there is no impact of MA on wages for state owned enterprises (SOE). Results for two tests of hypothesis are reported at the bottom of table 6. The first test does not reject the hypothesis that the sum of the coefficient on the common MA and that on the SOE specific MA is significantly different from 0. The second test at the bottom of Table 6 verifies that the impact of MA on private firms is significantly different from the one on state-owned firms.

Foreign and Sino-foreign owned firms, although having relatively high coefficients, are not significantly different from the reference category. However, the lack of significance might be due to a small number of firms in these two groups (15 foreign and 136 Sino-foreign versus over 3000 state owned firms) or, particularly in the case of the 100% foreign firms, the location of these enterprises which are already in places with high market access like Beijing or Guangzhou in Guangdong.

In the last column, we add an interactive term between MA and a dummy variable for urban collectively owned firms. These enterprises are often considered to be similar to private firms. However, their wage sensitivity to market access does not appear to be different to that of the firms of the reference group.

5 Conclusion

This study has examined the impact of economic geography on the spatial structure of wages in China. It has attempted to explain inter-individual wage differences by the geography of access to markets of the individuals' location. We control for gender and age as well as individual skills in terms of experience and education, which can already explain an important part of spatial wage variations. To account for the literature pointing out the importance of factor endowments in spatial wage disparities, province-industry fixed effects are introduced. Even after controlling for these factors, the relationship found between the city's market access, computed based on a gravity equation, and the individual Chinese wages is significant and positive.

Our results also highlight, that the impact of MA cannot be generalized since the relationship between market access and wage holds only for high-profile workers and for certain firm types. Wages earned in private firms react very strongly to changes in MA, while wages in state-owned enterprises are largely insensitive to them. This means that, probably due to the remnants of central economic planning, the impact of MA is limited in the Chinese economy.

Finally, it seems that inter-regional and intra-provincial wage disparities are partly due to differences in market access. Considering that market access also has an important international component, it is likely that with a further integration into the world economy, the disparities will grow, if access to new markets is not evenly distributed within the country. Therefore, the results of this study suggest that a further opening of the country, without an increasing liberalization of internal migration, may worsen the already pervasive spatial wage disparities.

Nevertheless, the impact of market access could change over time. More research with data covering several years is needed to conclude on the evolution of wages in response to rising market access and to see if eventually the ongoing internal migration will lead to an equalization of wages across the Chinese territory for given skills and factor endowments.

References

Baldwin, R., R. Forslid, P. Martin, G. Ottaviano and F. Robert-Nicoud (2003), Economic Geography and Public Policy. Princeton: Princeton University Press.

Bannister, J. (2005), Manufacturing Employment and Compensation in China, Consultant's report to US Departement of Labor, Bureau of Labor Statistics.

Breinlich, H. (2006), The Spatial Income Structure in the European Union - What Role for Economic Geography?, *Journal of Economic Geography*, 6(5), 593-617.

Brakman, S., H. Garretsen and M. Schramm (2004), The Spatial Distribution of Wages and Employment: Estimating the Helpman-Hanson Model for Germany, *Journal of Regional Science*, 44(3), 437-466.

Brülhart, M. and P. Koenig (2006), New Economic Geography meets Comecon: Regional Wages and Industry Location in Central Europe, The Economics of Transition, Economics of Transition, 14(2), 245-267.

Combes, P.-P., G. Duranton and L. Gobillon (2005), Spatial Wage Disparities: Sorting Matters!, *CEPR Discussion Paper* 4240, revised version.

De Bruyne, K. (2003), The Location of Economic Activity: Is there a Spatial Employment Structure in Belgium?, mimeo, K. U. Leuven.

Davis, D. and D. Weinstein (1999), Economic Geography and Regional Production Structure: An Empirical Investigation, *European Economic Review*, 43(2), 379-407.

Démurger, S., M. Fournier and C. Yi (2005), Earnings Differentials and Ownership Structure in Chinese Enterprises, *Economic Development and Cultural Change*, 53(4), 933-958.

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

Feenstra, R., W. Hai, W. T. Woo and S. Yao (1998), The U.S.-China Bilateral Trade Balance: Its Size and Determinants, with Short version published in the *American Economic Review*, May 1999, 89(2), 338-343.

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

Hadi, A.S. (1994), A Modification of a Method for the Detection of Outliers in Multivariate Samples, *Journal of the Royal Statistics Society*, 56(B), 393-396.

Feenstra, R. and G. Hanson (2000), Aggregation Bias in the Factor Content of Trade: Evidence from U.S. Manufacturing, *American Economic Review*, 90(2), 150-160.

Hanson, G. (2000), Firms, Workers, and the Geographic Concentration of Economic Activity. In G. Clark, M. Gerler and M. Feldman (eds) Handbook of Economic Geography, Oxford University Press.

Hanson, G. and C. Xiang (2004), The Home Market Effect and Bilateral Trade Patterns, *American Economic Review*, 94(4), 1108-1129.

Hanson, G. (2005), Market Potential, Increasing Returns, and Geographic Concentration, *Journal of International Economics*, 67(1), 1-24.

Harris, C. (1954), The Market as a Factor in the Localization of Industry in the United States, *Annals of the Association of American Geographers*, 64, 315-348.

Head, K. and T. Mayer (2004a), The Empirics of Aggl omeration and Trade, in V. Henderson and J.F. Thisse (eds.) Handbook of Regional and Urban Economics Volume 4, Amsterdam: Elsevier, chapter 59, 2609-2669.

Head, K. and T. Mayer (2004b), Market Potential and the Location of Japanese Investment in the European Union, *Review of Economics and Statistics*, 86(4), 959-972.

Head, K. and T. Mayer (2006), Regional Wage and Employment Responses to Market Potential in the EU, *Regional Science and Urban Economics*, 36(5), 573-594.

Krugman, P. (1980), Scale Economies, Product Differentiation and the Pattern of Trade, *American Economic Review*,70(5), 950-959.

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

Krugman, P. and A. J. Venables (1995), Globalization and the Inequality of Nations, *Quarterly Journal of Economics*, 110(4), 857-880.

Lee, J. (2000), Changes in the Source of Chinas Regional Inequality, *China Economic Review*, 11(3), 232-245.

Lin, S. (2005), Geographic Location, Trade and Income Inequality in China, in R. Kabur and A. J. Venables Spatial (eds.) Inequality and Development, Oxford University Press, London.

Lyons, T. (1998), Intraprovincial Disparities in China: Fujian Province, 1978-1995, *Economic Geography*, 74(4), 405-432.

Meng, X., R. Gregory and Y. Wang (2005), Poverty, Inequality and Growth in Urban China, 1986-2000, *Journal of Comparative Economics*, 33(4), 710-729.

Mion, G. (2004), Spatial Externalities and Empirical Analysis: The Case of Italy, *Journal of Urban Economics*, 56(1), 97-118.

Mion, G. and P. Naticchioni (2005), Urbanization Externalities, Market Potential and Spatial Sorting of Skills and Firms, *CEPR Discussion Paper* 5172.

Moulton, B. (1986), Random Group Effects and the Precision of Regression Estimates, *Journal of Econometrics*, 32(3), 385-397.

Moulton, B. (1990), An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units, *The Review of Economics and Statistics*, 72(2), 334-338.

Pagan, A. (1984), Model Evaluation by Variable Addition, in D.F. Hendry and K.F. Wallis (eds) Econometrics and Quantitative Economics, Blackwell: Oxford, chapter 5,103-134.

Paillacar, R. (2006), Market Potential and Worker Heterogeneity as Determinants of Brazilian Wages, mimeo, University of Paris 1.

Ping, H. and F. Pieke (2003), China Migration Country Study, paper presented at the DFID/ RMMRU Conference on Migration Development Pro-poor policy Choices in Asia, Dhaka, 22-24 June.

Poncet, S. (2003), Measuring Chinese Domestic and International Integration?, *China Economic Review*, 14(1), 1-22.

Poncet, S. (2005), A Fragmented China: Measure and Determinants of China's Domestic Market Disintegration, *Review of International Economics*, 13(3), 409-430.

Redding, S. and P. Schott (2003), Distance, Skill Deepening and Development: Will Peripheral Countries Ever Get Rich?, *Journal of Development Economics*, 72(2), 515-541.

Redding, S. and A. J. Venables (2003), South-East Asian Export Performance: External Market Access And Internal Supply Capacity, *Journal of the Japanese and International Economies*, 17(4), 404-441.

Redding, S. and A. J. Venables (2004), Economic Geography and International Inequality, *Journal of International Economics*, 62(1), 53-82.

Riskin, C., R. Zhao and S. Li (2000), Chinese Household Income Project, 1995 [Computer file]. ICPSR version. Amherst, MA: University of Massachusetts, Political Economy Research Institute [producer], 2000. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Riskin, C., R. Zhao and S. Li (2001), Chinas Retreat from Equality Income Distribution and Economic Transition. London: M.E. Sharpe.

State Statistical Bureau. Various issues. China Statistical Yearbooks. China Statistical Bureau. Beijing.

State Statistical Bureau (2000), Comprehensive statistical data and materials on 50 years of new China. Beijing.

Wei, S. J. (1996), Intra-National Versus International Trade: How Stubborn Are Nations in Global Integration, *NBER Working Paper* 5531.

Wei, S. J. and Y. Wu (2002), Globalization and Inequality Without Differences in Data Definition, Legal System and Other Institutions, *International Monetary Fund Working Paper*, october.

Weng, Q. (1998), Local Impacts of the post-Mao Development Strategy: The Case of the Zhujiang Delta, *International Journal of Urban and Regional Research*, 22(3), 425-442.

Tables

Table 1: Benchmark Wage equation estimations

		Explained variable: individual wage							
		1	2	3	4	5	6	7	8
	ln(MA)		0.141 (0.01)***						0.134 (0.01)***
	ln(Harris MA)			0.722 (0.01)***					
	ln(Non local MA)				0.12 (0.01)***				
	ln(Non provincial MA)					0.137 (0.01)***			
	ln(Foreign MA)						0.117 (0.01)***		
	Provincial Capital							0.106 (0.02)***	0.072 (0.02)***
Indiv vari-ables	Female Gender	-0.118 (0.02)***	-0.119 (0.02)***	-0.119 (0.02)***	-0.119 (0.02)***	-0.119 (0.02)***	-0.119 (0.02)***	-0.120 (0.02)***	-0.121 (0.02)***
	School years	0.037 (0.00)***	0.035 (0.00)***	0.033 (0.00)***	0.035 (0.00)***	0.035 (0.00)***	0.035 (0.00)***	0.034 (0.00)***	0.033 (0.00)***
	Experience	0.022 (0.00)***	0.022 (0.00)***	0.022 (0.00)***	0.022 (0.00)***	0.022 (0.00)***	0.022 (0.00)***	0.022 (0.00)***	0.022 (0.00)***
	Constant	7.859 (0.05)***	8.357 (0.03)***	-8.366 (2.54)***	8.488 (0.04)***	8.352 (0.03)***	8.507 (0.05)***	7.865 (0.05)***	8.339 (0.04)***
	Fixed effects	Sector							
	Observations	9747	9747	9747	9747	9747	9747	9747	9747
	Number of sectors	14	14	14	14	14	14	14	14
	Within R-squared	0.16	0.20	0.20	0.22	0.20	0.22	0.17	0.20

Heteroskedastic consistent standard errors in parentheses, with ***, ** and * denoting the significance at 1, 5 and 10% level.
Moulton correction for industry cluster correlation.
The reported R-squared is the Within R-squared which indicates how much of the variation of wages within the group of sectors is explained by our regressors.

Table 2: Baseline Wage equation estimations

		Explained variable: individual wage		
		1	2	3
	ln(MA)	0.060 (0.030)**	0.061 (0.029)**	0.058 (0.029)**
Indiv variables	Female Gender	-0.136 (0.014)***	-0.139 (0.014)***	-0.117 (0.014)***
	School years	0.034 (0.010)***	0.035 (0.010)***	0.024 (0.010)**
	School years^2	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
	Experience	0.047 (0.003)***	0.020 (0.006)***	0.020 (0.006)***
	Experience^2	-0.001 (0.000)***	-0.000 (0.000)	-0.000 (0.000)
	Age		0.057 (0.009)***	0.060 (0.009)***
	Age^2		-0.001 (0.000)***	-0.001 (0.000)***
	Provincial Capital	0.074 (0.029)**	0.073 (0.029)**	0.054 (0.024)**
Firm Ownership	State-owned			0.149 (0.048)***
	Local public			-0.014 (0.045)
	Urban coll.			-0.203 (0.053)***
	Private			-0.216 (0.162)
	Individual			-0.183 (0.341)
	Sino-foreign			0.287 (0.056)***
	Foreign			0.803 (0.188)***
	TVE			0.070 (0.046)
	Constant	7.925 (0.117)***	7.068 (0.173)***	7.058 (0.190)***
	Fixed effects	Sector and province		
	Observations	9747	9747	9747
	Number of groups	140	140	140
	Within R-squared	0.19	0.19	0.24
Test	Hansen Over.	0.53	0.53	0.58
	Wu Hausman	0.29	0.28	0.86

Heteroskedastic consistent standard errors in parentheses, with ***, ** and * denoting the significance at 1, 5 and 10% level.
The reported R-squared is the Within R-squared which indicates how much of the variation of wages within the group of province/sector is explained by our regressors.
Moulton correction for province-industry cluster correlation.

Table 3: Baseline Wage equation estimations (Wage per hour)

		Explained variable: individual wage per hour		
		1	2	3
Indiv variables	ln(MA)	0.074	0.075	0.072
		(0.035)**	(0.034)**	(0.034)**
	Female Gender	-0.114	-0.116	-0.094
		(0.015)***	(0.016)***	(0.015)***
	School years	0.051	0.055	0.044
		(0.011)***	(0.011)***	(0.012)***
	school years^2	-0.001	-0.001	-0.001
		(0.000)	(0.000)	(0.001)
	Experience	0.047	0.019	0.019
		(0.003)***	(0.006)***	(0.006)***
	experience^2	-0.001	-0.000	-0.000
		(0.000)***	(0.000)	(0.000)
	Age		0.057	0.060
			(0.010)***	(0.010)***
	Age^2		-0.001	-0.001
			(0.000)***	(0.000)***
	Provincial Capital	0.093	0.092	0.071
		(0.032)***	(0.032)***	(0.028)**
Firm Ownership	State-owned			0.111
				(0.056)**
	Local public			-0.066
				(0.053)
	Urban coll.			-0.251
				(0.062)***
	Private			-0.474
				(0.166)***
	Individual			-0.406
				(0.377)
	Sino-foregin			0.130
				(0.058)**
	Foreign			0.586
				(0.180)***
	TVE			-0.005
				(0.055)
	Constant	0.157	-0.715	-0.671
		(0.131)	(0.226)***	(0.236)***
	Fixed effects	Sector and province		
	Observations	9559	9559	9559
	Number of groups	140	140	140
	Within R-squared	0.17	0.18	0.22
Test	Hansen Over.	0.21	0.21	0.24
	Wu Hausman	0.10	0.11	0.48

Heteroskedastic consistent standard errors in parentheses, with ***, ** and * denoting the significance at 1, 5 and 10% level.
Moulton correction for province-industry cluster correlation.

Table 4: Disentangling size and market access effects

		Explained variable: individual wage			
		1	2	3	4
	ln(MA)	0.084	0.058	0.067	0.063
		(0.039)**	(0.028)**	(0.026)**	(0.027)**
	ln(POP)	0.120			
		(0.040)***			
	ln(GDP)		0.075		0.066
			(0.024)***		(0.023)***
	ln(Qualified)			0.301	0.192
				(0.115)**	(0.091)**
Indiv variables	Female Gender	-0.123	-0.119	-0.118	-0.119
		(0.014)***	(0.014)***	(0.014)***	(0.014)***
	School years	0.017	0.022	0.021	0.020
		(0.013)	(0.011)**	(0.011)*	(0.011)*
	School years^2	0.001	0.000	0.000	0.000
		(0.001)	(0.000)	(0.000)	(0.001)
	Experience	0.021	0.021	0.020	0.021
		(0.006)***	(0.006)***	(0.006)***	(0.006)***
	Experience^2	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)
	Age	0.057	0.057	0.058	0.056
		(0.012)***	(0.009)***	(0.009)***	(0.009)***
	Age^2	-0.001	-0.001	-0.001	-0.001
		(0.000)***	(0.000)***	(0.000)***	(0.000)***
	Provincial Capital	-0.095	-0.071	0.023	-0.075
		(0.058)	(0.047)	(0.033)	(0.046)
	Constant	5.573	6.085	7.276	6.347
		(0.543)***	(0.349)***	(0.181)***	(0.309)***
	Fixed effects	Sector and province			
Dummies for firm ownership		yes	yes	yes	yes
	Observations	7899	9747	9747	9747
	Number of groups	139	140	140	140
	Within R-squared	0.24	0.24	0.24	0.24

Heteroskedastic consistent standard errors in parentheses, with ***, ** and *
denoting the significance at 1, 5 and 10% level.
Moulton correction for province-industry cluster correlation.

Table 5: Education-contingent market access impact on wage

		Explained variable: individual wage		
		1	2	3
			Skilled	Unskilled
Market access	ln(MA)		0.076 (0.026)***	0.035 (0.046)
	ln(MA)* high skilled workers	0.066 (0.027)**		
	ln(MA)* low skilled workers	0.044 (0.034)		
Indiv variables	low skilled workers	-0.153 (0.060)**		
	Female Gender	-0.118 (0.014)***	-0.093 (0.013)***	-0.164 (0.021)***
	School years	-0.007 (0.011)	-0.013 (0.022)	0.011 (0.034)
	School years^2	0.001 (0.000)**	0.002 (0.001)*	0.000 (0.002)
	Experience	0.019 (0.006)***	0.025 (0.006)***	0.008 (0.010)
	Experience^2	-0.000 (0.000)	-0.000 (0.000)**	0.000 (0.000)
	Age	0.061 (0.009)***	0.050 (0.010)***	0.083 (0.016)***
	Age^2	-0.001 (0.000)***	-0.001 (0.000)***	-0.001 (0.000)***
	Provincial Capital	0.055 (0.024)**	0.069 (0.023)***	0.017 (0.037)
	Constant	7.309 (0.188)***	7.499 (0.217)***	6.786 (0.340)***
	Fixed effects	Sector and province		
Dummies for firm ownership		yes	yes	yes
	Observations	9747	6391	3356
	Number of groups	140	140	129
	Within R-squared	0.24	0.25	0.22

Heteroskedastic consistent standard errors in parentheses, with ***,** and *
denoting the significance at 1, 5 and 10% level.
Moulton correction for province-industry cluster correlation.

Table 6: Firm type-contingent impact of market access

		Explained variable: individual wage		
		1	2	3
Market access	ln(MA)	0.075	0.075	0.071
		(0.030)**	(0.030)**	(0.029)**
	Add for SOE firms	-0.075	-0.075	-0.072
		(0.029)**	(0.029)**	(0.027)***
	Add for private firms	0.358	0.358	0.361
		(0.127)***	(0.127)***	(0.126)***
	Add for urban coll. firms			0.017
				(0.027)
	Add for foreign firms	0.046		
		(0.038)		
	Add for 100% foreign firms		0.179	0.0182
			(0.261)	(0.261)
	Add for Sino-foreign firms		0.040	0.044
			(0.038)	(0.040)
Indiv variables	Female Gender	-0.117	-0.117	-0.117
		(0.014)***	(0.014)***	(0.014)***
	School years	0.026	0.026	0.026
		(0.010)**	(0.010)**	(0.011)**
	$School^2$	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
	Experience	0.020	0.020	0.020
		(0.005)***	(0.005)***	(0.005)***
	Exp^2	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
	Age	0.060	0.060	0.060
		(0.009)***	(0.009)***	(0.009)***
	Age^2	-0.001	-0.001	-0.001
		(0.000)***	(0.000)***	(0.000)***
	Provincial Capital	0.051	0.051	0.050
		(0.023)**	(0.023)**	(0.023)**
	Constant	7.108	7.109	7.093
		(0.187)***	(0.187)***	(0.188)***
	Fixed effects	Sector and province		
Dummies for firm ownership		yes	yes	yes
	Observations	9747	9747	9747
	Number of groups	140	140	140
	Within R-squared	0.24	0.24	0.24
Test	MA SOE=0	0.98	0.99	0.975
	MA SOE=MA private	0.01***	0.02**	0.00***

Heteroskedastic consistent standard errors in parentheses, with ***,** and * denoting the significance at 1, 5 and 10% level.
Moulton correction for province-industry cluster correlation.

Appendix B

Table B1: Descriptive statistics of wages (in yuan) per province

	All	Beijing	Shanxi	Liaoning	Jiangsu	Anhui	Skilled workers
Obs	10316	856	932	1093	1307	823	6741
Mean	5899	7424	4854	5450	6241	4573	6192
Standard Dev.	3477	3562	2418	2544	3011	2100	3581
Median	5309	5309	7140	5094	5964	4366	5560
	Henan	Hubei	Guangdong	Sichuan	Yunnan	Gansu	Nonskilled workers
Obs	810	1044	975	1196	624	656	3575
Mean	4462	5446	10147	5504	5588	4323	5346
Standard Dev.	2250	2551	5883	2891	2056	1958	3201
Median	4223.5	5127.5	9300	5151.5	5427	4203	4763

Table B2: Trade equation estimations

Dependent Variable: Ln (Exports) in 1997	
Ln Distance	-1.528 (0.024)***
Contiguity	1.162 (0.123)***
Inter-foreign country Border effect	-1.731 (0.320)***
China-foreign country Border effect	-3.681 (0.353)***
Intra China Border effect	-2.766 (0.780)***
Constant	19.488 (0.440)***
Fixed effects by exporter Fixed effects by importer	
Observations	22290
Number of importers	270
Within R-squared	0.59

Heteroskedastic consistent standard errors in parentheses, with *** denoting the significance at 1% level.

Table B3: Statistics on Market Access

		Decomposition of total market access in %		Ratio of city MA to highest
Province	City	National	Foreign	in %
Guangdong	Foshan	79.4	20.6	100
Beijing	Beijing	64.1	35.9	71
Jiangsu	Suqian	93.7	6.3	50
Guangdong	Guangzhou	43.2	56.8	41
Jiangsu	Nantong	92.8	7.2	38
Jiangsu	Changzhou	90.1	9.9	38
Henan	Xinxiang	91.6	8.4	38
Guangdong	Shunde	46.7	53.3	35
Guangdong	Huizhou	65.0	35.0	32
Anhui	Wuhu	96.0	4.0	31
Hubei	Huangshi	96.3	3.7	30
Jiangsu	Wuxi	88.4	11.6	29
Guangdong	Zhaoqing	52.2	47.8	28
Guangdong	Shenzhen	53.2	46.8	25
Henan	Kaifeng	86.8	13.2	24
Jiangsu	Nanjing	72.7	27.3	24
Hubei	Xiangfan	95.8	4.2	23
Shanxi	Changzhi	97.6	2.4	22
Henan	Pingdingshan	85.8	14.2	21
Liaoning	Jinzhou	94.9	5.1	20
Anhui	Hefei	92.4	7.6	20
Anhui	Bengbu	93.6	6.4	19
Jiangsu	Taixing	80.4	19.6	19
Jiangsu	Xuzhou	84.7	15.3	15
Henan	Zhengzhou	74.4	25.6	14
Shanxi	Yangquan	94.1	5.9	13
Guangdong	Puning	58.0	42.0	13
Jiangsu	Yixing	78.9	21.1	13
Guangdong	Zhanjiang	64.0	36.0	13
Sichuan	Zigong	95.2	4.8	12
Henan	Xiangcheng	77.5	22.5	12
Gansu	Lanzhou	47.3	52.7	11
Gansu	Pingliang	55.9	44.1	11
Anhui	Huainan	87.9	12.1	11
Jiangsu	Dafeng	76.8	23.2	10
Liaoning	Shenyang	76.5	23.5	9
Henan	Huixian	66.3	33.7	9
Shanxi	Fenyang	91.7	8.3	9
Sichuan	Chengdu	94.2	5.8	9
Yunnan	Dali	79.8	20.2	9
Yunnan	Gejiu	78.4	21.6	9
Liaoning	Dalian	91.4	8.6	9
Gansu	Wuwei	39.0	61.0	8
Anhui	Tongcheng	85.7	14.3	8
Shanxi	Taiyuan	87.8	12.2	7
Yunnan	Kunming	68.4	31.6	7
Anhui	Bozhou	84.0	16.0	7
Sichuan	Luzhou	90.3	9.7	6
Hubei	Tianmen	81.7	18.3	6
Hubei	Honghu	81.8	18.2	6
Shanxi	Datong	88.8	11.2	6
Sichuan	Leshan	90.4	9.6	6
Hubei	Wuhan	74.7	25.3	5
Hubei	Macheng	77.5	22.5	5
Yunnan	Xuanwei	63.4	36.6	4
Sichuan	Guangyuan	86.2	13.8	4
Average		78.5	21.5	

Table B4: Correlation matrix

	ln (wage)	ln (MA)	Female gender	School years	School years^2	Exp.	Exp.^{2}	Age	Age^2	Prov. Capital	ln (GDP)
ln(wage)	1.0000										
ln(MA)	0.2014	1.0000									
F. gender	-0.1671	-0.0010	1.0000								
Sch. years	0.1725	0.0345	-0.1064	1.0000							
Sch. years^2	0.1785	0.0337	-0.1129	0.9886	1.0000						
Exp.	0.3475	0.0368	-0.1393	-0.0913	-0.0696	1.0000					
Exp.^{2}	0.3115	0.0439	-0.1689	-0.0597	-0.0371	0.9627	1.0000				
Age	0.3423	0.0311	-0.1410	-0.0581	-0.0286	0.9420	0.9040	1.0000			
Age^2	0.3246	0.0355	-0.1582	-0.0390	-0.0086	0.9299	0.9263	0.9896	1.0000		
P. Capital	0.1350	0.1643	-0.0004	0.1493	0.1472	0.0776	0.0783	0.0745	0.0740	1.0000	
ln(GDP)	0.2636	0.3134	-0.0027	0.1478	0.1442	0.1114	0.1100	0.1150	0.1134	0.7405	1.0000

Table B5: Summary statistics

Variables	Mean	Median	Std. Dev.
Market Access	0.050	0.029	0.047
School years	10.85	11	2.74
School years^2	125.23	121	61.42
Experience	19.52	20	9.56
Experience^2	472.58	400	384.48
Age	38.45	39	9.46
Age^2	1568.44	1521	731.11
GDP (yuan)	3465424	1415811	4001181

Table B6: Market Access Instrumentation first step procedure (Table 2, Col. 3)

		Explained variable: City-level Market Access
		market access
Indiv variables	Female gender	0.013
		(0.007)*
	School years	-0.021
		(-0.009)**
	School years^2	0.000
		(0.000)**
	Experience	0.002
		(0.002)
	Experience^2	0.000
		(0.000)
	Age	-0.000
		(0.001)
	Age^2	-0.000
		(0.000)**
	Capital city	-0.291 (0.01)***
Excluded Instruments	Instrument1 (centrality China)	0.084 (0.017)***
	Instrument2 (centrality ROW)	-6.590 (0.215)***
	Constant	0.020 (0.058)
	Fixed effects	Sector and province
Dummies for firm ownership		yes
	Observations	9747
	Number of groups	140
	Within R-squared	0.29
	R-squared excluded	0.27

Heteroskedastic consistent standard errors in parentheses, with ***,** and * denoting the significance at 1, 5 and 10% level.

Appendix C: Description of Data

C-1 International Data

International trade flows are in current USD. They are from the IMF Direction of Trade Statistics (DOTS).

Internal trade flows are in current USD and are calculated as the difference between domestic primary and secondary sector production minus exports. Production data for OECD countries come from the OECD STAN database. For other countries, ratios of industry and agriculture output in percentage of GDP are extracted from Datastream. They are then multiplied by the countries' GDP (in current USD) coming from the World Development Indicators 2005.

C-2 Chinese Data

C-2-1 Trade flows data

Intra-provincial trade flows

Total production for Chinese provinces is computed as the sum of industry and agriculture output. Output in yuan is converted into current USD using the annual exchange rate. All statistics come from the China Statistical Yearbooks.

Inter-provincial trade flows

Provincial input output tables[40] provide the decomposition of provincial output, international and domestic trade of tradable goods. Domestic trade flows, that is trade between each

[40] Most Chinese provinces produced square input output tables for 1997. A few of them are published in provincial statistical yearbooks. We obtained access to final-demand columns of these matrices from the input output division in China's National Bureau of Statistics.

province and the rest of China, were obtained for the year 1997[41].

The rest of China, denoted by roC, differs for each province considered and can be thought of as a distinct country whose characteristics (distance to partners d_{i-roC}) are generated on the basis of the characteristics of the provinces that make it up. See Poncet (2005) for more details.

Chinese international trade flows

The provincial foreign trade data are obtained from the Customs General Administration database, which recorded the value of all the import or export transactions through the customs. Provincial imports and exports are decomposed into up to 230 international partners. We rely on data for 1997. The background of this database is discussed in Lin (2005) and Feenstra et al. (1998).

C-2-2 Other data

Statistics on GDP, land area and population at the city and province level comes mainly from two sources: (1) Urban Statistical Yearbook, various issues, published by China's State Statistical Bureau, and (2) China City Statistics, data for County cities. This data is collected by China's State Statistical Bureau, but is downloadable (for a fee) at http://chinadataonline.org.

[41] IO tables are available for 28 provinces as data are missing for Tibet, Hainan and Chongqing). Four provinces (Anhui, Heilongjiang, Shandong and Guizhou) list only net outflows and are thus not useful for studying inter-provincial trade. Nine provinces separate inflows and outflows into domestic and foreign sectors. Poncet (2005) deduced domestic trade flows for the other provinces using industry-level provincial import and export data from the General Administration of Customs. These data match the data reported as international trade by provinces that separate international and domestic transactions in their input-output tables. This finding gives some confidence in the method used as input-output tables and customs data appear to use consistent methodology.

Map 1: Chinese provinces

Figure 1 : Market Access and average wage for 56 prefectures

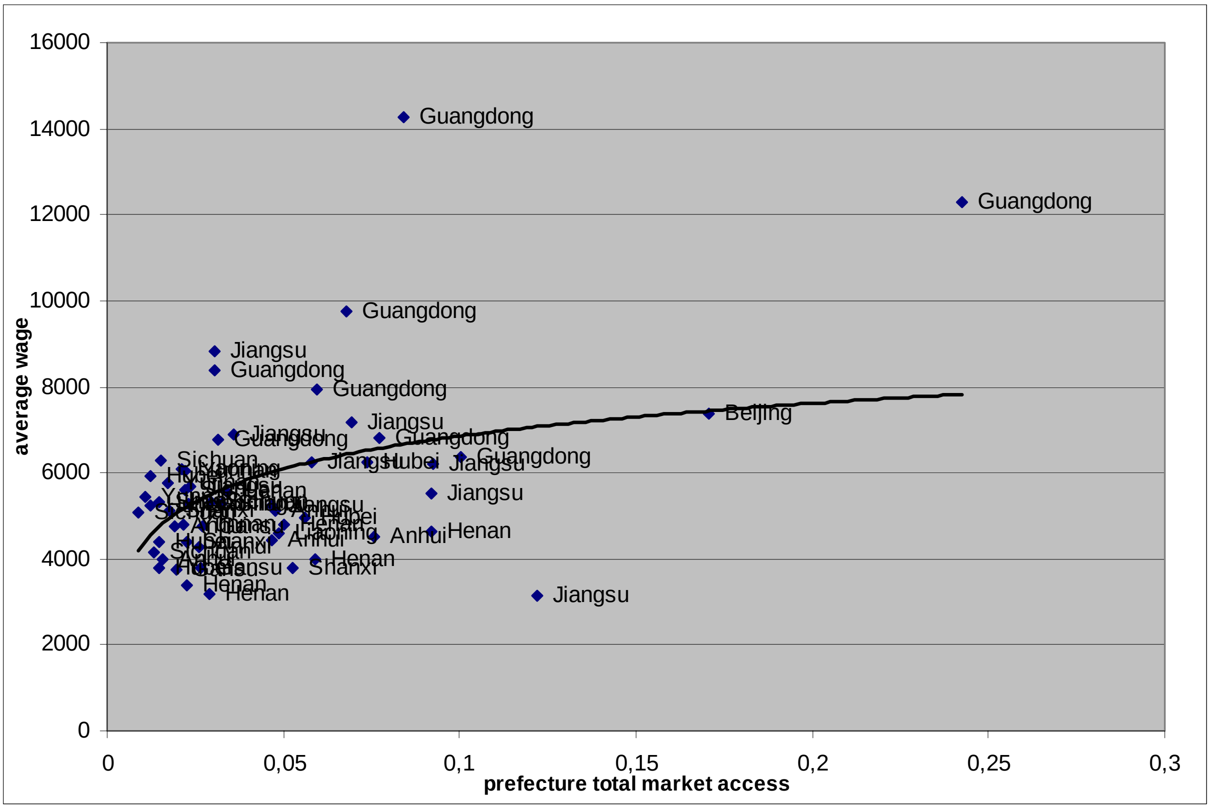

Figure 2 : Market Access and average distance to Beijing and Guangzhou for 56 prefectures