CESifo Working Papers
www.cesifo.org/wp

# Relaxing Hukou - Increased Labor Mobility and China's Economic Geography

Maarten Bosker
Steven Brakman
Harry Garretsen
Marc Schramm

CESIFO WORKING PAPER NO. 3271
CATEGORY 12: EMPIRICAL AND THEORETICAL METHODS
DECEMBER 2010

CESifo *Working Paper No. 3271*

# Relaxing Hukou - Increased Labor Mobility and China's Economic Geography


## Abstract

China's *Hukou* system poses severe restrictions on labor mobility. This paper assesses the consequences of relaxing these restrictions for China's internal economic geography. We base our analysis on a new economic geography model. First, we obtain estimates of the important model parameters on the basis of information on 264 of China's prefecture cities over the period 1999-2005. Second, and by using our estimation results as input, we simulate various long-run scenarios of China's internal economic geography that differ in their degree of interregional labor mobility. We find that increased labor mobility will lead to more pronounced core-periphery outcomes. Interestingly, these agglomerations are not necessarily along the coastal regions. Given the increased importance of China's internal market, firms agglomerate in the populous heartland of China. China's internal demand will be the most important determinant of its future economic geography.


JEL-Code: O18, R11, R12.


*Maarten Bosker*
*Faculty of Economics and Business*
*University of Groningen*
*The Netherlands*
*e.m.bosker@rug.nl*

*Steven Brakman*
*Faculty of Economics and Business*
*University of Groningen*
*The Netherlands*
*s.brakman@rug.nl*

*Harry Garretsen*
*Faculty of Economics and Business*
*University of Groningen*
*The Netherlands*
*j.h.garretsen@rug.nl*

*Marc Schramm*
*Utrecht University*
*School of Economics*
*The Netherlands*
*m.schramm@econ.uu.nl*


This version:October 2010
Please send all correspondence to: j.h.garretsen@rug.nl
We thank seminar participants at the Spatial Economics Research Centre at the London School of Economics; the Dept. of Land Economy at the Univ. of Cambridge; the 2009 ETSG Conference in Rome; Utrecht University; University of Groningen, the 2009 NARSC conference in San Francisco, the EBRD in London and especially Bart van Ark, Bernard Fingleton, Vernon Henderson, Laura Hering, Ron Martin, Ana Moreno Monroy, Charles van Marrewijk, Henry Overman, Pierre Picard, Sandra Poncet, Mark Roberts and Tuo Shi for their comments and suggestions.

## 1. Introduction

According to the *World Development Report 2009*, economic development is to a large extent driven by economic geography (World Bank, 2008). Economic geography refers not only to physical characteristics of the landscape, but also – and most importantly – to man-made economic geography in the sense of a region's access to upstream and downstream markets or the degree of urbanization. The analytical underpinnings for this (re)discovery of the relevance of economic geography come in no small part from the new economic geography literature initiated by Krugman (1991) in which the balance between agglomerating and spreading forces determines the location of economic activity. The relevance of this approach has been shown for the cases of the US, and Europe, and only recently attention has shifted towards developing countries like China.

China is one of the fastest growing economies in the world. Also, it is home to a huge (potential) internal market and a large number of very large and fast-growing cities. Arguably, China is a textbook case to analyze how agglomerating and spreading forces can shape the economic landscape. To cite Krugman (2010, p.14): "*Chinese economic geography is highly reminiscent of the economic geography of the advanced nations circa 1900 and it fits gratifyingly well into the new economic geography framework.*"

The recent phenomenal growth record of China has been accompanied by increased migration from the western and central provinces to China's eastern (coastal) provinces (World Bank, 2008). Recent studies like Au and Henderson (2006a,b), however, argue that Chinese cities are undersized due to severe restrictions on labor mobility that are still imposed through the so called *Hukou* system (see also Chan, 2008 and 2009). As a result, China is thought not to reap the full benefits of agglomeration. A simple illustration of this fact is that the Gini-coefficient for the Chinese city size distribution was 0.43 in 2000. Compared to a world-average Gini-coefficient of 0.564, this indicates a more *evenly* sized city size distribution in China (Fujita et al, 2004, p. 2955).

In this paper, we consider the impact of the *Hukou* system on the spatial distribution of economic activity in China. In particular, we use a new economic geography (NEG) model (Puga, 1999) to analyse the agglomeration-labor mobility nexus for China. We argue that NEG models are particularly suited to analyse the case of rapidly industrializing economies like China. NEG models analyze agglomeration in terms of (changes) in the balance between spreading and agglomeration forces.[2] A feature of the NEG model in Puga (1999) that makes

[2] See also Krugman (2010, p.15):" *...to return briefly to the issue of general, abstract models versus historical specificity: surely the strong resemblance between China's industrial clusters today and the industrial clusters of*

it particularly well-suited for our case is that in his model, the balance between these forces and hence the equilibrium spatial distribution of economic activity depends explicitly on the degree of interregional labor mobility. This makes this model a natural starting point to assess the relaxation of the *Hukou* restrictions on China's internal economic geography.

Based on a sample of 264 Chinese Prefecture cities, we first estimate the equilibrium wage equation that is central in NEG models. The estimated wage equation not only empirically establishes the link between a city's wages and its market access (see also Hering and Poncet, 2010); it also provides two key model parameters that are central in our subsequent simulation analysis of the relationship between agglomeration and labor mobility.

Our paper crucially differs from related recent NEG studies for China (e.g. Ma, 2006, Lin, 2003, Hering and Poncet, 2007, 2010, De Sousa and Poncet, 2007, or Amiti and Javorcik, 2008).[3] First, we do not take the spatial allocation of labor (and firms) as given. Instead, we make use of the *complete* NEG model and not just the equilibrium (nominal) wage equation. This enables us to analyse the possible effects of relaxing China's *Hukou* system on its internal economic geography. Second, in doing this we go beyond the simple migration dynamics that underlie most theoretical NEG models and consider more intricate and realistic migration dynamics.

Our main findings show that a relaxation of the restrictions on labor mobility posed by the Hukou system will lead to more pronounced core-periphery outcomes. Interestingly, the economic geography of China that would result from such increased labor mobility is first and foremost determined by within China differences in market access, population and (arable) land. The geography of China's own internal demand is the most important determinant of its future economic geography. Given the increasing importance of China's own huge domestic market, international market access is of less significance in determining the spatial distribution of economic activity. Our simulations show that increased labor mobility primarily benefits non-coastal Prefecture cities like Zhoukou or Liuan in the large and populous Honan and Anhwei provinces. Prefecture cities like Shanghai or Guangzhou that have currently a very high market access do not necessarily end up being the main centers of economic activity once inter-city firm and labor mobility are taken into account.

---

*the 19th-century – especially the export-oriented clusters of late-Victorian Britain – strikes a blow in favor of the argument for simple, common principles*." In particular the focus in most NEG models on the agricultural and manufacturing sector as the 2 relevant sectors may seem outdated for western countries but it is (still) relevant for a country like China.

[3] Roberts et al. (2010) also use the full NEG model in their analyis of Chinese economic geogaphy but their focus is not on labor mobility but on the impact of changes in the Chinese infrastructure on the spatial distribution of economic activity.

## 2. Labor mobility in China and the implications from a NEG perspective

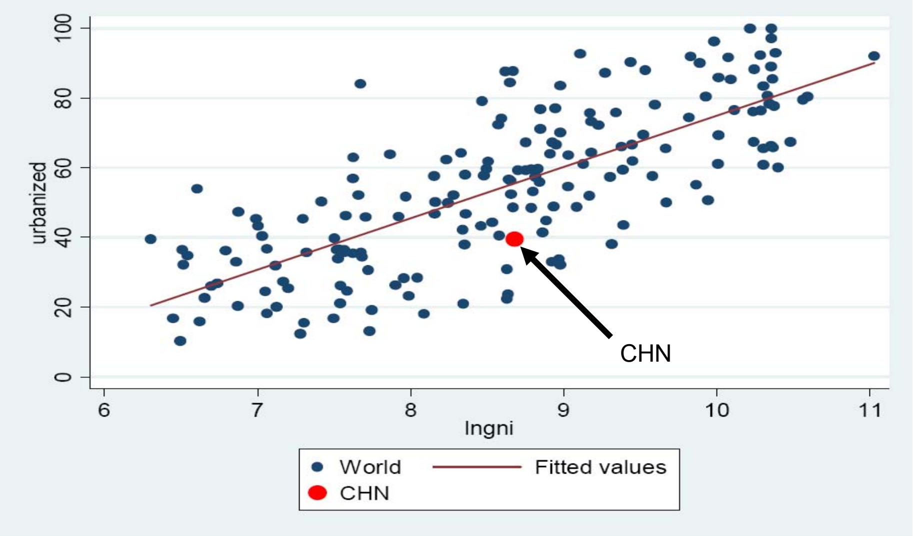
Since the 1950s the Chinese authorities have been much concerned with internal labor migration flows and rural-urban labor migration in particular. The Chinese government alternated between periods of more and less restrictive migration policies (Zhao, 2004, Fujita et al, 2004, Chan and Buckingham, 2008; World Bank, 2008, ch 5, and Chan, 2009) but ever since the 1950s the so called *Hukou* system has been a main feature of its internal migration policies. Although the system might have become somewhat less restrictive over time it is still very much a prominent characteristic of the Chinese labor market (see f.i. the detailed description of the *Hukou* system in *The Economis* of May 6th, 2010). The *Hukou* system is equivalent to an internal visa arrangement that is meant to regulate migration. In recent decades the system has been quite restrictive, not only by limiting (official) migration flows from rural to urban areas, but also by putting a brake on inter-urban migration flows (Chan and Buckinham, 2008, Au and Henderson, 2006a,b; Poncet, 2006, Henderson, 2009). Without a visa for a particular location, a Chinese citizen has no or only limited rights to housing, sell property, education, food or social security in that location. Those rights are tied to one's official place of residence and a change in residency (if a citizen for instance would try to move from a rural area to a city) will only be matched with a transfer of these rights if the (local) authorities hand out a visa or permit for the new place of residence.

Until recently (see in particular Chan, 2009 or Chan and Buckingham, 2008, pp. 13-14), migration under the *Hukou* system had two equally important dimensions. The first concerns the restrictions regarding preferred locations alluded to above. To be granted permission to live and work in a city or any other location, and to be entitled with the aforementioned rights to public provisions, households need a local *Hukou.* The second dimension is functional and refers to the distinction between agricultural and non-agricultural workers. Within each location, a further distinction is made between agricultural and non-agricultural workers. Non-agricultural workers with a local *Hukou* are traditionally entitled to more rights than agricultural workers. The local and functional classification combined thus define four possible categories of residents (Chan, 2009, p.202).

In recent years, the *Hukou* system has become less of a constraint (Chan and Buckingham, 2008). With urban wages outstripping rural wages, the result has been an increase in official as well as illegal (temporary) migration into the booming cities. One important change is that currently, in many cases, the local instead of the central government decides upon the permits. This gives local governments some degree of freedom with respect to the leniency of granting these permits. A second change is that the distinction between a non-agricultural and an agricultural *Hukou* has become relatively less important (Chan, 2009). The key issue for a migrant is now foremost whether he or she has a local *Hukou,* which is primarily concerned with the location of a worker.

Although the system has become less stringent and more differentiated, the fact that temporary and illegal migration flows to the main cities have been surging shows that the *Hukou* system is still very restrictive. "*all these restrictions sharply reduce the benefits and raise the costs of migration, particularly into large cities. Migration is limited and most migration is short-term, or "return" migration. (…). Overall the hukou system holds hundreds of millions of people in locations where they are not exploiting their earnings potential."* (Fujita et al., 2004, p. 2957). Chan (2008) estimates that between 1982 and 2006, the annual volume of Hukou migrants amounted to 17- 20 million people. The impact of the labor mobility restrictions is clear from the stability of this annual official migration flows (Fujita et al, 2004, p. 2957). This stability also suggests that neither the scrapping of the functional *Hukou* nor the decentralization of the migration policy has had a substantial impact on the flow of permanent (official) migrants to the cities.[4]

**Figure 1. Urbanization and gross national income (GNI) per capita for countries**




Source: Henderson (2009) based on World Bank Development indicators, 2006; urbanized = degree of urbanization; lngni = PPP corrected GNI per capita (in logs)

[4] We do not have information on illegal migration/workers. In China there are very substantial temporary migration flows to cities, part of which is illegal. As a consequence, many workers do not have a local Hukou of that city. In our data set, we can only deal with official (legal) city population data. For an instructive lucid summary of the current Hukou system see http://chunzhu.wordpress.com/2008/04/22/understanding-the-hukou-system/. See also Chan (2009) for a very good introduction to and assessment of the Hukou system.

What are the consequences of the Hukou system for China? Au and Henderson (2006a,b) forcefully argue that Chinese cities are too small as a consequence of the migration restrictions and that the agglomeration rents associated with urbanization are therefore underutilized (see also Fujita et al, 2004).[5] The associated welfare losses of the lack of labor mobility are considerable (Whalley and Zhang, 2007). Figure 1 illustrates this line of argument in a stylized way (Henderson, 2009): China's income per capita is associated with a below average degree of urbanisation.

However, apart from the general prediction that relaxing *Hukou*'s migration restrictions will most likely result in higher urbanization rates and larger and/or more uneven-sized Chinese cities (Au and Henderson, 2006a), little work has been done on assessing the impact of increased labor mobility on the internal economic geography of China. Will it result in very strong core-periphery patterns, with most economic activity concentrated in a few very large cities? And if so, which cities or regions will be able to attract economic activity and which cities will thereby become more peripheral? Or will core periphery patterns become less pronounced? From an economic development perspective this is important given the well established positive relationship between agglomeration (urbanization) and economic development (see World Bank (2008) and Figure 1 above).

The main aim of this paper is to shed light on this subject by investigating the implications of different labor mobility regimes for the spatial distribution of people and economic activity across Chinese cities. Based on a general NEG model, we analyze theory-based predictions regarding the effect of changes in interregional labor mobility for China's internal economic geography. We argue that the NEG approach is useful because labor mobility plays a key role in NEG models.

In NEG models a less than perfect interregional labor mobility results in wage differences between regions: large agglomerations have higher wages than more peripheral areas. Economic centers offer firms (and workers) better access to upstream and downstream markets. With perfect labor mobility instead, i.e. perfectly elastic interregional labor supply, economic centers will simply attract more workers (and firms) until (real) wages are equalized across locations (or all footloose workers have migrated to the centre before complete real wage equality is reached).[6] The Chinese *Hukou* system poses considerable restrictions on the degree of interregional labor immobility. According to NEG theory this has the effect that

[5] Or as *The Economist* of September 18th 2010 states: "*China's small cties exploded in number. But its biggest metropolises conspiciously failed to explode in size (...). This partly reflects a conscicious policy. Although China's rulers have embraced urbamisation, theyare still wary of mega-cities*" (p. 84).

[6] See also our discussion of the NEG model in section 3 and in particular of the wage equation (2).

Chinese cities with a high market access are expected to have higher wages than cities with a low market access.[7]

In our empirical sections we will confirm this *wages-market access* hypothesis using information on a sample of 264 Chinese prefecture cities. Subsequently, and based on this result, we adopt a simulation strategy to reveal the possible consequences of relaxing the *Hukou* restrictions for China's economic geography.

.

## 3. The NEG model

We start by briefly introducing the NEG model that we use as the basis of our empirical investigation. We adopt the NEG model by Puga (1999) as our benchmark model. This model encompasses a variety of NEG models as special cases, notably the core models by Krugman (1991), Krugman and Venables (1995) and Venables (1996). More importantly for our purposes (see also the discussion in Hu, 2002), it explicitly distinguishes between different labor mobility regimes. It easily facilitates comparing the consequences of different degrees of labor mobility on the resulting spatial distribution of economic activity.

The $R$ $(i=1,\ldots,R)$ region version of the model can be understood as follows (see Appendix A1 for a more detailed exposition of the model). Each of the $R$ regions is populated by $L_i$ workers and endowed with $K_i$ units of arable land. Each region's economy consists if two sectors: agriculture and industry. Labor is used by both sectors and is mobile between sectors within a region and it is either mobile or immobile between regions. Land on the other hand is used only by the agricultural sector and is immobile between regions.

This stylized depiction of a region's economy resembles the definition of a Chinese Prefecture city with its non-urban part (producing agricultural goods) and urban part (producing manufacturing goods) quite well. In addition, see Chan and Buckingham (2008) and our discussion in section 2, the current *Hukou* system allows for intra-Prefecture city labor mobility between the urban and non-urban part of a Prefecture city. In particular, the declined importance of the distinction between the agricultural and the non-agricultural *Hukou* within each Prefecture implies that, in terms of the model, labor can move relatively easy between the two sectors. Empirical evidence also backs up this idea that inter-sector

[7] Indeed, the link between imperfect labor mobility and the wages-market access relationship is the starting point for the recent NEG empirical studies that try to establish if Chinese wages do depend positively on market access, see Hering and Poncet (2006, 2007), Au and Henderson (2006a), De Sousa and Poncet (2007), Moreno Monroy (2010), Lin (2003) and our own estimations in section 4. Market access is also a determinant of the location decisions of firms. Building on Head and Mayer (2004) and also invoking imperfect labor mobility, Amiti and Javorcik (2008) show for instance that FDI flows to Chinese regions depends positively on the market access of those regions, for a similar result using an NEG model see Ma (2006).

labor mobility within Chinese Prefecture cities is important. Fujita et al (2004, p. 2958, Table 2) e.g. document for instance that in the 1990s the non-agricultural employment growth clearly outstripped population growth which can be interpreted as a sign of inter-sector labor mobility.

The agricultural good is produced using land and labor (combined in a Cobb-Douglas production function, with $0 < \theta < 1$ denoting labor's share in agricultural production)[8]. Its market is assumed to be perfectly competitive with free entry and exit. Moreover it is freely tradable between regions. The assumed production structure in agriculture implies diminishing returns in this sector, so that any attempt of manufacturing firms to lure workers away from the agricultural sector implies wage increases (see Chan, 2009, p. 208 on the wage elasticity of intra-regional manufacturing labor supply). In this sense $K_i$ acts as a spreading force in the model with similar consequences as the housing sector in Helpman (1998). It captures in a stylized way the costs of congestion associated with larger agglomerations. This presence of this additional congestion force sets Puga (1999) apart from earlier contributions by Krugman (1991) or Venables (1996).

The industrial sector produces heterogeneous varieties of a single good under monopolistic competition and free entry and exit, incurring so-called 'iceberg' trade costs when shipped between regions ($\tau_{ij} \geq 1$ goods have to be shipped from region *i* to let one good arrive in region *j*). Industrial production technology is characterized by increasing returns to scale. The production input is a Cobb-Douglas composite of labor and intermediates, with $0 \leq \mu \leq 1$ the Cobb-Douglas share of intermediates. Intermediates enter the production function as a composite manufacturing good that is specified as a CES-aggregate (with $\sigma > 1$ the elasticity of substitution across varieties) of all manufacturing varieties produced. Finally, regions are allowed to differ in their production efficiency, yet firms within the same region are assumed to be similar in this[9].

Consumers in turn have Cobb-Douglas preferences over the agricultural good and a CES-composite (also with $\sigma > 1$ the elasticity of substitution across varieties) of manufacturing varieties, with $0 \leq \gamma \leq 1$ the Cobb-Douglas share of the composite manufacturing good. Specifying the composite manufacturing good in this way ensures

[8] Puga (1999) defines the agricultural sector somewhat more general. However, when deriving analytical results, a Cobb-Douglas production function in agriculture is used, see Puga (1999, p. 318).

[9] This is our only deviation from the standard Puga (1999) model. It is introduced to be able to take account of other reasons (human capital, sectoral structure, comparative advantage) why some regions may be able to offer higher wages than interregional differences in market access (see our discussion of equation (5) in section 5.1).

demand for each manufacturing variety from each region, implying that trade takes place between Chinese Prefectures.

Next, equilibrium (factor) prices and demand follow from profit and utility maximization on behalf of firms and consumers respectively (again see Appendix A1 for the details). These in turn determine the spatial distribution of economic activity. Puga (1999) shows that this distribution depends critically on the assumptions made regarding interregional labor mobility.

In case there is *no labor mobility* between regions, the model reduces to the following three equilibrium conditions for each region $i$ (for the (manufacturing) price index $q_i$, the (manufacturing) wage $w_i$, and total expenditures on manufactures $e_i$ respectively):

$$q_i = \left( \frac{1}{1-\mu} \sum_j \left( \varsigma_j L_j q_j^{-\mu\sigma} c_j^{-\sigma} w_j^{1-\sigma(1-\mu)} \tau_{ij}^{1-\sigma} \right) \right)^{1/(1-\sigma)} \tag{1}$$

$$w_i = q_i^{\mu/(\mu-1)} c_i^{1/(\mu-1)} \left( \sum_j e_j q_j^{\sigma-1} \tau_{ij}^{1-\sigma} \right)^{1/(\sigma(1-\mu))} \tag{2}$$

$$e_i = \gamma(w_i L_i + K_i r(w_i)) + \mu/(1-\mu) w_i \varsigma_i L_i \tag{3}$$

where (in addition to the variables and parameters already introduced) $\varsigma_i$ denotes the share of region $i$'s labor force in manufacturing, $c_i$ is a region specific indicator of production efficiency, and $r(w_i)$ the rent earned per unit of land in region $i$. These conditions (partly) follow from the requirement that in equilibrium no worker has to have an incentive to move from the agricultural to the manufacturing sector or vice versa. That is, equilibrium is characterized by *nominal wage equality between sectors within each region*.

When labor is instead allowed to move between regions, nominal wage equality between sectors within each region is no longer enough to ensure equilibrium. Workers now also move in response to *real wage differences between regions*. As a result, with interregional labor mobility, equilibrium is characterized not only by (1) – (3) but also by interregional real wage equality (i.e. labor moves until real wages $\omega_i$ are equalized across all regions):

$$\omega_i = q_i^{-\gamma} w_i = \omega \qquad \forall i \tag{4}$$

From a model perspective, the absence or presence of (some degree of) interregional labor mobility is important because it changes the mix of agglomeration and spreading forces. Hereby the two different scenarios can result in radically different long run spatial equilibria. Without any interregional labor mobility, and given the (crucial) assumption that agricultural production takes place under decreasing returns to labour, the agglomeration of manufacturing firms to particular core regions has a price tag in the form of higher wages. The additional workers needed in the manufacturing sector can only come from the local agricultural sector. The excess labor demand from the manufacturing sector drives up wages in this region, creating interregional wage differences between core and peripheral regions, that persist in equilibrium. The rising wages also create, *ceteris paribus*, a spreading force that makes manufacturing firms return to regions with lower wages. With perfect interregional labor mobility, this upward pressure on wages is instead countered by the additional supply of labor that moves from peripheral to core regions in response to the higher wages offered in the core regions. As such, a higher degree of labor mobility typically results in an economic geography characterized by stronger core-periphery patterns than in case of imperfect labor mobility[10].

The empirical sections in this paper are closely related to the NEG model set out above. In section 5, we estimate the equilibrium wage equation (2) in order to obtain estimates of some of the important model parameters. In section 6 we go one step further than earlier empirical NEG studies on China (or in fact most empirical NEG studies in general[11]), and simulate the full NEG model to assess what the long run spatial distribution of economic in China would look like under different assumptions regarding the degree of interreginal labor mobility.

## 4. Data set and a first look at China's economic geography

At the highest level of aggregation, China is composed of 33 administrative units (22 provinces, 5 autonomous regions, 4 large municipalities (Beijing, Shanghai, Tianjin, and Chongqing) and 2 special regions (Hong Kong and Macau). The 2nd tier of regional division, the so called *Prefecture level*, in China consists of 333 regions. Of these 333 regions at the Prefecture level, 283 regions are Prefecture-*cities*[12]. Our data set consists of data for a large subset of these Prefecture cities for the period 1995-2002: 264 of the 283 cities are included. Even though the data do not cover the whole of China, most of its population and economic activity is included. Our 264 prefecture cities cover 86% of total population in China and 96%

[10] Although this need not necessarily be the case.

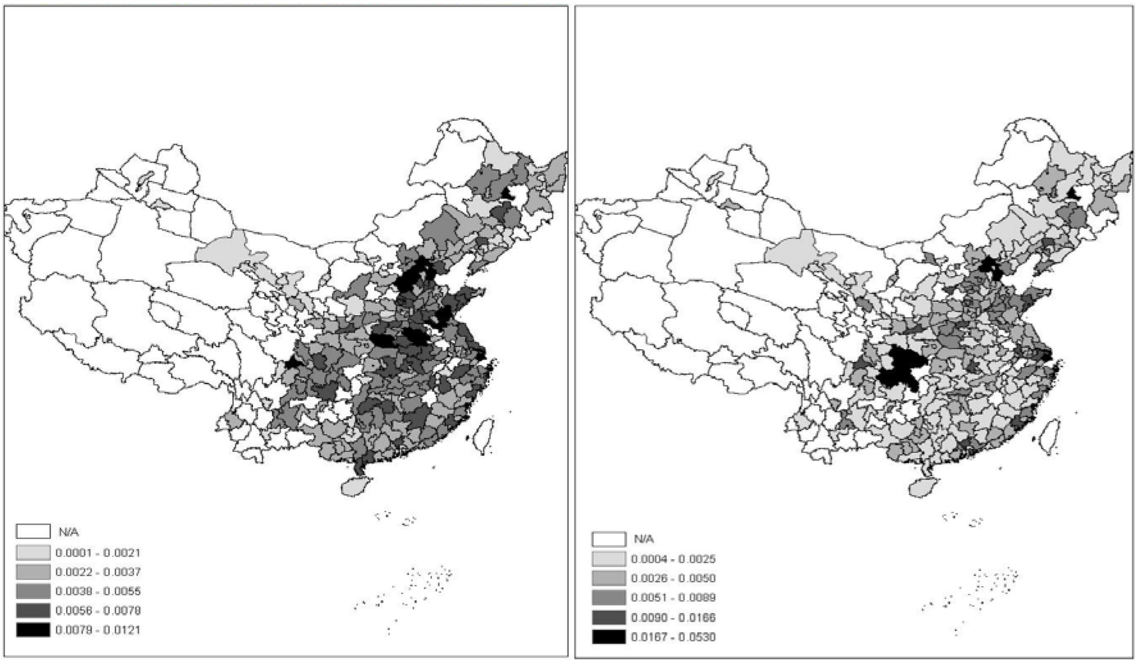
[11] Virtually all empirical contributions in the NEG literature are based on the wage equation. See e.g. Redding and Venables (2004), Hanson (2005), Knaap (2006) or Hering and Poncet (2010).

[12] Note that the prefectures not only include the urban population of a city, but also the rural area surrounding a particular city, in many Prefecture cities in our data set the majority of the population as well as the bulk of the land area is classified as non-urban. Apart from the 283 Prefecture cities the 2nd tier of regional administration also consists of 17 Prefectures (mainly in Xinjiang and Tibet), 30 Autonomous Prefectures (in western China, regions with a large share of ethnic minorities) and 3 Leagues (regions in Inner Mongolia), see http://en.wikipedia.org/wiki/Administrative_divisions_of_China#Prefecture_level_subdivisions.

of total GDP. In Appendix A2 we list for every regional unit at the highest level of administration (we have data on 30 of the total of 33 regions)[13] the corresponding Prefecture cities that are in our data set.

Our data on these 264 Chinese prefecture cities come from the Chinese Data Center at the University of Michigan (see http://chinadatacenter.org/newcdc/). The original data source is the *National Bureau of Statistics of China*. From this database we collect data on each prefecture city's expenditures (income), total population, employment, its available arable land (in $km^2$), as well as distance to the nearest major ports, secondary education enrolment, the share of employment in banking and finance, and its total area. Moreover we obtained the (great-circle) distances between each prefecture city-pair.

**Figure 2. Distribution of people and industry employment across Prefecture cities**




**a. Distribution of population** **b. Distribution of sec. industry employment**

Source: our data set; N/A = not available and/or not a prefecture city; data shown are for 2002.

To give an idea of China's current economic geography, Figure 2a and 2b show the (2002) distribution of population, and secondary industry employment (a proxy for the distribution of manufacturing activity) respectively across our sample of 264 Prefecture cities.

Figure 2a shows that the largest concentration of population is found in an area that can be loosely defined as the part of China that lies within the Shanghai-Chongqing-Beijing

[13] We have no data on the following 3 regions out of the group of 33 regions: Hong Kong, Macau, and Tibet. The same underlying data set is also used by for instance Au and Henderson (2006a) or Moreno Monroy (2010).

triangle. This triangle mainly encompasses the provinces of Honan, Hupeh, Anhwei, Kiangsu and Sjantung (it also covers the populous part of Hebei province). Taken together these 5 provinces alone are home to 34% of the Chinese population and to almost 28% of our 264 Prefecture cities (these numbers are 40% and 32% resp. if Hebei province would be included). The presence of these many, relatively populous, Prefecture cities in China's heartland will turn out to be very important in the long-run equilibrium analysis of our NEG model in section 5.

Figure 2b furthermore shows that the distribution of manufacturing activity does not correspond one-to-one with the distribution of people. Industry employment (panel b) is more spatially concentrated (Herfindahl index (HI) = 0.01) than population (HI=0.006), with a few notable Prefecture cities like Beijing, Tianjin, Shanghai and Chongqing (the large black area in panel b) each home to 3-5% of total Chinese industry employment.[14] By contrast, the population share of either of these four cities amounts to no more than 1%. The five earlier mentioned provinces in China's populous heartland together contain about 30% of total industry employment.

To assess how China's internal economic geography, depicted by Figure 2, might change due to a loosening of the *Hukou* restrictions, we now turn to our empirical strategy. We start by estimating NEG's wage equation in order to obtain empirically grounded estimates of important model parameters. We then subsequently use these estimates, and simulate the full NEG model under various different labor mobility scenarios.

## 5 Estimating the wage equation for Chinese cities

In order to be able to provide theory-based predictions of the effects of different labor mobility regimes for China's internal economic geography, we need estimates of all the important model parameters of our NEG model. Estimates of the share of intermediate inputs ($\mu$) in manufacturing production, the share of income spent on manufactures ($\gamma$) and the Cobb-Douglas share of labor in agriculture ($\theta$) are relatively easy to obtain. Based on a regional input-output table for China for 2000 we can infer $\mu$ and $\gamma$ for China as a whole. To be specific, we take $\mu = 0.51$ (I-O table: Chinese intermediate demand for Chinese manufacturing by Chinese manufacturing firms) and $\gamma = 0.34$ (I-O table: Chinese final

[14] A perfectly even spread of population or secondary employment would give a HI of 0.0038. As we will explain in section 5 the Prefecture "city" of Chongqing is a rather special case.

demand for Chinese manufacturing as share of total final demand for Chinese output).[15] Moreover, we obtain $\theta = 0.87$ from the *Statistical Yearbook* of the Chinese Bureau of Statistics as the share of wages in total agricultural value added (indicating that agricultural production is (still) relatively labor intensive in China).

However, besides these three important parameters, we also need an estimate of $\sigma$, the elasticity of substitution between manufacturing varieties. To get at this, we estimate the following log-linearized version of the equilibrium wage equation (2):

$$\ln\left(w_{it}\right) = \frac{1}{\sigma}\ln\underbrace{\left(\sum_{j=1}^{30} e_j \tau_{ij}^{1-\sigma} q_j^{\sigma-1}\right)}_{market\ access} + \alpha_i + \alpha_t + \alpha_2 \ln X_{it} + \varepsilon_{it} \qquad (5)$$

where the productivity differences between regions, $c_i$ in (2), are captured by several observed control variables ($X_{it}$) on the one hand (see below for more on our specific choice of controls) and a random error term $\varepsilon_{it}$ that is assumed to be uncorrelated with the included regressors on the other hand. In addition, we allow $c_i$ to differ between cities for unobserved reasons by including prefecture city dummies (captured by the $\alpha_i$'s in (5)), and we include year dummies to control for general improvements in production efficiency equally affecting all cities in the sample (captured by the $\alpha_t$'s in (5)).

Estimating the NEG wage equation lies at the heart of most earlier empirical NEG studies. It has been used to verify whether market access (the term between brackets) plays an important role in explaining interregional wage differences. In case of China, Hering and Poncet (2010) for example estimate a similar wage equation as (5). They provide strong evidence in support of NEG's prediction that workers living in regions with better market access than others c.p. earn higher wages.

Note that (5) is actually a simplified version of (2). Besides the market access term, (5) should in effect also contain an additional term involving $q_i$. This other term is frequently referred to as a region's supplier access. We exclude this term from (5), effectively assuming that $\mu = 0$ (see also Hering and Poncet, 2010). As shown and emphasized by Redding and Venables, 2004), including both supplier and market access in the wage equation leads to multi-collinearity problems.

Besides this issue of multicollinearity, a few other estimation issues remain when estimating (5). First, in order to arguably limit problems with reverse causality (see Hanson, 2005), we include a market access measure in (5) that is based on data at one level of

[15] Source: regional I/O table from Institute of Developing Economies (2003), "Multi-Regional Input-Output Table for China 2000".

aggregation higher than our prefecture data. These province variables *(j=1, ...30)* are constructed by aggregating the corresponding prefecture variables in the sample to the province level (Hanson (2005) uses a similar aggregation for his sample of US counties).

Second, in the absence of actual trade costs data, we, as all other existing empirical NEG studies, proxy these trade costs by specifying a trade cost function (see Bosker and Garretsen, 2010 for a recent discussion on the use of a trade cost function in empirical NEG). In particular, we proxy trade costs by a simple power distance function: $\tau_{ij}=(D_{ij})^{\delta}$, $D_{ij}$ is defined as the distance between city $i$ and the capital of province $j$. Trade costs within one's own province ($\tau_{ii}$) are proxied similarly using a, by now standard, measure of internal distance: $D_{ii} = \frac{2}{3}\sqrt{\frac{area_j}{\pi}}$ . The use of such a distance function introduces the distance decay parameter $\delta$ as an additional important model parameter.

Thirdly, we do not directly observe the manufacturing price index, $q_i$, that is present in the market access term. To overcome this problem we follow Brakman, Garretsen and Schramm (2004) and approximate $q_j$ as follows[16]. Under the earlier assumption of $\mu = 0$, wages are the most important determinant of prices. As such, we proxy a province's manufacturing price index by a weighted combination of information on a province's own wage level with the average wage outside that province (weighted by distance to correct for the transport costs involved in importing goods from other provinces):

$$q_j = \left\{ \lambda_j W_j^{1-\sigma} + \left(1-\lambda_j\right)\left(\overline{w} D_{j,centre}^{\delta}\right)^{1-\sigma} \right\}^{\frac{1}{1-\sigma}} \tag{6}$$

where $\overline{w}$ is the average wage of all other provinces in the sample, and $D_{j,centre}$ the distance between the capital of province $j$ and the nearest economic centre; Beijing, Shanghai or Quandung. $\lambda_j$ is the share of employees of province $j$. When estimating (5) we substitute (6) for $q_j$ in equation (5).

Two basic empirical strategies have been used to estimate (5). The first, introduced by Redding and Venables (2004), is to use the information contained in bilateral trade data to construct a theory-based measure of each region's market access. Subsequently this constructed measure of market access is included in the estimation of the wage equation. In case of China, Hering and Poncet (2006, 2007), Lin (2003) and Ma (2006) have for example followed this strategy using data on inter-*provincial* trade. The second strategy, following

[16] Two other options to deal with unavailability of data on city priced indices are; to go for nominal instead of real market access (see Au and Henderson 2006a) which takes us out of the world of NEG, or to opt for a solution whereby other equilibrium conditions of the underlying NEG model can be used (Hanson, 2005). The 2nd route followed by Hanson (2005) has the drawback that real wage equalization must be assumed, which is unrealistic for China.

Hanson (2005), is to estimate (5) directly using nonlinear estimation techniques. We opt for this second strategy here because there is, to our knowledge, no bilateral trade data available at the Prefecture-city level that (sufficiently) cover our sample. Other related market access studies for China, adopting such a more direct approach are e.g. Amiti and Javorcik (2008) or Au and Henderson (2006a).

Table 1 shows the results of estimating (5). In all our regressions, we use prefecture city specific gdp per capita as our measure of wages $w_i$. Data on the average wage of workers in the urban part of each prefecture city, and data on the total wage of employees are available but following Hering and Poncet (2007, p. 13) who argue that the available prefectural city wage data do not sufficiently reflect wages in the *private* sector, we did not use either of these two wage measures as our main dependent variable, opting for gdp per capita instead. Expenditures $e_j$ are proxied by a prefecture city's income.

Finally, to capture possible productivity difference between prefecture cities ($c_i$ in (2)) we include the following control variables ($X_{it}$) in (5). In urban economics a city's density itself is typically related to higher productivity levels because of increasing returns associated with the well-known Marshallian externalities of e.g. labor market pooling and knowledge or input sharing. To control for this, and in line with for instance Hering and Poncet (2010), Breinlich (2006) and Bosker and Garretsen (2008), we include *population density* to (5). Second, we also follow these authors and include a proxy for *human capital* as a higher educated workforce is likely to be more productive[17]. To control for the possible relevance of the economic structure of a city, we add the *share of banking and finance in total Prefecture-city employment* to our regression.

Finally, we allow for unobserved (time-invariant) differences in productivity between cities by including a full *set of prefecture city dummies*, at the same time also controlling for possible common shocks to overall Chinese productivity by including a full *set of year dummies* as well. Including prefecture fixed effects allows some cities to have a comparative advantage in attracting workers and/or industry. In particular it helps to control for the cities in our sample that are located along China's coast and that are often specialized in exporting to world markets (Hu, 2002, Au and Henderson, 2006a) and able to pay higher wages. As Krugman (2009, 2010) points out China's overall comparative advantage in the international economy (e.g in low cost manufacturing goods) goes along with a strong and very uneven

[17] For their sample of 51 Chinese cities, Hering and Poncet (2010) use micro data which allows for a much better control of various wage determinants besides market access.

spatial concentration of economic activity in China itself.[18] We try to control for this supply driven part of China's economic geography in our estimations by including city fixed effects. The fixed effects may also to some extent capture the effect(s) of Chinese industrial policy where specific regions or cities are favored over others. Recall that the basic term in the wage equation that we are interested in, the market acces term, is about the spatial allocation of demand for a prefecture city's manufacured goods.

**Table 1. Results of estimating the NEG wage equation**

| Dependent variable: ln(urban gdp per cap.) | | |
|---|---|---|
| | (a) | (b) |
| σ | 5.886***<br>(.70) | 5.776***<br>(.70) |
| δ | .632***<br>(.04) | .636***<br>(.04) |
| secondary education | 1.664***<br>(.51) | 1.657***<br>(.51) |
| % Banking & finance in total employment | .795*<br>(.47) | .791*<br>(.40) |
| Ln(Distance to nearest seaport city)<br>* ln(distance-weighted foreign GDP) | --- | .0182<br>(.03) |
| Ln(Population density) | .101***<br>(0.02) | .101***<br>(.02) |
| | | |
| Sample period | 1999-2005 | 1999-2005 |
| Observations | 1779 | 1779 |
| Adj. $R^2$ | 0.968 | 0.968 |
| Foreign GDP included in market access | Yes | No |
| Fixed effects? | prefecture & year | prefecture & year |

*Notes:* White heteroskedasticity-consistent standard errors between brackets. ***, **, * denotes significance at the 1%, 5%, 10% level respectively. *NLS*, estimation. *secondary education* is the student enrolment in regular secondary schools as a fraction of city *i's* population. *population density* is measured as city *i's* population (in 10,000 persons) per $km^2$. *Distance-weighted foreign GDP* is defined here as $\sum_i (Y_{it}/Distance_{kj})$, where *i* is USA, Japan or EU 15, *k* is the major seaport of USA, Japan or EU 15 (Los Angeles, Yokohama and Rotterdam, respectively), *j* is the major seaport in China (Shanghai). Distance is measured as great-circle distance. The distance between Shanghai and Rotterdam is calculated as the sum of the distance between Shanghai-Djibouti, Djibouti-Suez, Suez-Gibraltar, Gibraltar-Rotterdam.

Table 1 shows our results. Its two columns differ in the way we deal with *foreign market access*. Our market access measure as depicted in (5) only captures access to China's internal market. Yet, in fact China is not a closed economy and market access with respect to the rest of the world may also be relevant. It is easy to incorporate this in the market access term (see also Hering and Poncet, 2010). We make the basic assumption that China trades with the rest of the world mostly by sea and that, for simplicity, the *rest of the world* consists of the three economic blocs: the USA, Japan and the EU. Column (a) and (b) then incorporate this in (5) in a different way. In our preferred specification (a), we directly include this in the market

[18] This spatial allocation is of course not given once and for all. Recently, there have been reports that production is moving away from coastal regions to more westward, non coastal regions (like the province Anhwei) where wage costs are lower (see "The Next China"in *The Economist* of July 31st 2010, pp. 46-48).

access term by adding (before taking logs) the economic mass outside of China (measured as the sum of distance-weighted GDP of the USA, Japan and the EU 15) to the market access term in (5). In column (b) we instead include the economic "mass" outside of China as a separate regressor to (5). To take account of differences between prefecture cities in their ease of access to the coast we interact this measure with a city's distance to the nearest seaport[19].

The differences between columns (a) and (b) our however minimal. The reason for this appears to be that the variation in foreign market access does not contribute much in explaining the difference in gdp per capita between Chinese Prefecture cities (it is insignificant in column (b)). This, arguably somewhat surprising, finding can be explained by the fact that the variation between cities in their access to China's internal markets dwarfs that in their market access to the rest of the world. It makes it hard to attribute the substantial variation in gdp per capita between prefecture cities to their foreign market access. This is not to deny that coastal cities like Shanghai are not special, but this is already to a large extent picked up by our set of Prefecture dummies[20].

Most importantly, our findings confirm those by earlier studies: we find strong evidence that market access plays an important role in explaining the observed wage differences between Chinese prefecture cities. The most relevant finding for our purposes is however that both the substitution elasticity $\sigma$ and the transport cost parameter $\delta$ are significant. A substitution elasticity of 5.88 is well within the range found in similar NEG studies (for e.g. Europe, the US or Sub-Saharan Africa) and it fulfils the theoretical requirement that $\sigma > 1$. It is also very similar to earlier empirical NEG studies on China. It implies a market access (MA) coefficient [$1/\sigma$, see (5)] of about 0.17. This is in line with Hering and Poncet (2010, MA-coeff. ≈ 0.1), Hering and Poncet (2007, MA-coeff. ≈ 0.07) and Moreno Monroy (2010, MA-coeff. ≈ 0.25). This size of the market access coefficient implies that a 1% increase in a city's market access boosts its wages by about 0.17%. Second, our estimated distance decay coefficient $\delta$ equals 0.63. Again this is very similar to earlier studies. Au and Henderson (2006a) e.g., following Poncet (2006), put the distance coefficient at 0.87.

The results on our control variables largely confirm earlier findings by e.g. Hering and Poncet (2010), Breinlich (2006) or Bosker and Garretsen (2010). Human capital is an important determinant of wages. Moreover, population density is also significantly positively related to wages, suggesting that the benefits of agglomeration still outweigh any negative

[19] Note that distance to the nearest seaport itself is captured by the prefecture fixed effects. Similarly the distance weighted sum of foreign GDP is captured by the year fixed effects.
[20] For many coastal cities the fixed effect coefficient is indeed significantly positive (not shown here).

congestion effects for the average Chinese city. The share of banking and finance employment is not significant at the 5% level.

As an important robustness check to our findings, we also ran separate *cross-section* estimations of wage equation (5) for 1999 and for 2000 separately. The reason being that it is believed that the reliability of the city-population data may be less from 2000 onwards as the Chinese authorities allegedly manipulated these data to a much larger extent than before.[21] Doing this (since we are unable to include prefecture dummies in these cross-section regressions, we include province dummies instead), our estimation results for the substitution elasticity and the trade cost parameters are very close to the full-sample estimation results shown in Table 1. For 1999 we find that these corresponding coefficients are (standard errors between brackets, and *, **, *** denotes significance at the 10%, 5%, 1% respectively): 5.658* (2.95) and 0.478*** (0.013), and for 2000 we find: 5.750** (2.74) and 0.491*** (0.012) respectively. Full results are available upon request.

Contrary to earlier studies that estimated the wage equation for China, here it is only *a means to an end* in our case. Our results in Table 1 confirm the findings in these earlier papers that economic geography matters also for the case of China. However, e.g. Hering and Poncet (2007, 2010) or Au and Henderson (2006a) stop at this point. As such, they verify whether there is evidence for the type of spatial wage pattern suggested by NEG theory given the current spatial allocation of workers and firms in China. Our main objective is, however, to take the analysis one step further. In the long-run, the location of workers and production is not given: market access is endogenous (the hallmark of NEG). In the complete NEG model, as discusssed in section 3, workers and firms move in response to the observed differences in wages and profits respectively. Although our estimation results in Table 1 confirm the existence of the type of spatial wage pattern predicted by NEG theory, it does not give any information as to what the future economic geography of China may look like or even whether the current economic geogeaphy can be seen as a long-run equilbrium outcome. It is not *a priori* clear why (if at all) an initially large or centrally located (coastal) city, like Shanghai, should also be a (even larger) economic centre in the long run equilibrium.

Here, the full NEG model of section 3 comes into play and not just wage equation (2). As such, we view our wage equation estimates as (one of the) inputs to a simulation analysis based on the complete NEG model. Doing this allows us to find out what the equilibrium

[21] We thank Vernon Henderson for pointing out this problem to us and for the suggestion to use cross-section estimations for the initial years of our sample as a robustness check. For the years 1999 and 2000 we only have data for 233 of our 264 cities.

allocation of economic activity across our sample of Chinese cities would look like under various "regimes" of interregional labor mobility. In a stylized way, and within the well-defined boundaries of our general NEG-model, this gives us predictions on what the impact of a loosening of the *Hukou* restrictions can be on China's internal economic geography.

This is what we turn to in the next section. We will use our estimates in column (a) as our benchmark simulation parameters. These estimation results are based on a specification of (5) that includes foreign market access in the theoretically most consistent way. We thus set the substitution elasticity at 5.88 and the distance parameter at 0.63 in our simulations of the full-blown NEG model.

## 6 Different labor mobility regimes and China's internal economic geography .

We are now in the position to analyse how the long run spatial equilibrium allocation of economic activity across Chinese cities may respond to various assumptions with respect to interregional labour mobility, i.e. changes in the *Hukou* regime. As useful analytical solutions for the case of $R > 2$ regions are non-existent in NEG models[22], one has to rely on simulation analyses (Behrens and Thisse, 2007, Fujita and Mori, 2005 and Bosker et al. 2010) to do this.

### *6.1. Simulation set-up*

In our long run analysis of the Puga (1999) NEG model, additional variables will be taken into account compared to those used in the estimation of the wage equation in the previous section. Most importantly, see our discussion in section 3, the allocation of the production factors $K_i$ (land) and $L_i$ (labor) is taken into account. For $K_i$ we take the cultivated arable land share of each Prefecture city $i$ in total Chinese arable land in 2002. And, similarly, for $L_i$ we take the population share of each Prefecture city $i$ in total Chinese population in 2002.[23] We

[22] Only special cases provide analytical solutions for multi region settings, such as the assumption that all regions are at equidistance.

[23] We decided to consider the Prefecture city of Chongqing as an outlier in the following way. In 1990, the then province of Chongqing was awarded the same status as Shanghai, Beijing and Tianjin that is of (effectively) a province that is also a prefecture city. This administrative change creates a problem for our present purposes because the "city" Chongqing is not only much larger than all other prefecture cities it is also huge in terms of (arable) land area. In 2002 Chongqing had a population of 31 million people and an arable land area of 2283 (x 1000 hectare). In population terms this about 3 times as large as the 2nd and 3rd largest cities, Shanghai and Beijing, and in arable land terms this is more than 4 times as large as these 2 other main cities with the same administrative status. These differences really matter in the long run analysis. Especially the land area difference matters in the sense that because of the decreasing returns to food production, see section 3, labor is very scarce in Chongqing and this puts an upward pressure on local wages to the effect that all Chinese footloose workers invariably will end up in Chongqing in the long run. A similar effect would occur if the Chinese authorities would suddenly decide to "upgrade" the large and populous provinces of Honan or Anhwei to the prefecture city status. To correct for this artificial size of the Prefecture city of Chongqing, we took as its total (2002) population only the *urban* population (9.9 million) of Chongqing and 456 (x1000 hectare) as the arable land area. The latter

combine this information with our estimates of the main model parameters (see the previous subsection). They are summarized in Table 2 below:

**Table 2. Main parameter settings in our simulations.**

| | |
|---|---|
| σ | 5.886 |
| δ | 0.632 |
| γ | 0.343 |
| μ | 0.511 |
| δ | 0.879 |

*Notes:* for the details on how we obtained these parameter settings, see section 5.

Furthermore, in some simulations we also take explicit account of the productivity differences between prefecture cities, $c_i$, that are due to their differences in human capital or are of an unobserved time-invariant nature (most notably controlling for coastal location)[24].

*6.2 Allowing for different interregional labor mobility regimes*

Puga (1999) distinguishes between two different cases of labor mobility[25]. He considers either the case of *complete labor immobility* or allows for *perfect interregional labor mobility*. In the latter case, workers move in response to the smallest wage differential between cities. The typical migration dynamics assumed in standard NEG models are very simple, only taking account of this simple notion that workers move in response to any real wage differential between locations[26]:

$$\frac{d\lambda_i}{\lambda_i} = \psi\left(\omega_i - \bar{\omega}\right) \tag{7}$$

where $\bar{\omega} = \sum \lambda_i \omega_i$ is the average real wage across all locations and $\lambda_i$ is the share of manufacturing firms in region $i$; $\psi$ is a parameter determining people's speed of adjustment.

These two extreme cases, although a useful starting point, can be argued to be too stylized for our purposes. Neither is the *Hukou* system so restrictive as to prevent any

is arrived at by taking the average ratio of poputlation and arable land area for Shanghai, Beijing and Tianjin and applying this to urban population of 9.9. million of Chonqging.

24 Simulating the model becomes increasingly complex when taking account of these differences in productivity. Especially so because they are assumed to be exogenous. It is however quite likely that, due to firm or worker reallocation a city's human capital base or sectoral structure will also change. The NEG model we use is silent on these issues. Therefore we do not consider our results on sectoral structure nor population density in Table 1 in our simulation exercises (both are almost by construction endogenous to migration of firms or workers resp.).

25 Note that we do not simulate different scenarios regarding intersectoral labor mobility. We assume throughout that within each Prefecture city inter-sectoral labor mobility is perfect.

26 Note that the typical NEG model is not a truly dynamic model. Talking about dynamics is therefore somewhat artificial.

interregional migration, nor does the assumption of workers moving in response to any, however small, wage differential realistically depict a world without *Hukou's* restrictions. In this section we will therefore also look at what happens to China's long run spatial distribution of economic activity in case of other, arguably more realistic migration dynamics. In particular, to analyse the relevance of changes in the degree of inter-city labor mobility we look at two basic types of scenarios for labor mobility:

1) *No - or restricted - labor mobility across China*

a. interregional labor mobility is completely restricted, only firms are allowed to move between regions.
b. interregional labor mobility is only allowed *within provinces only*, i.e. people can move freely, but only between prefecture cities located in the same province.

2) *Unrestricted labor mobility across China*

a. perfect interregional labor mobility, workers respond to the smallest wage differentials; in equilibrium wages are fully equalized as in (4)
b. interregional labor mobility, yet workers only respond to a certain wage differential, i.e. in equilibrium real wages are not fully equalized as in (4), instead equilibrium is characterized by the following condition:

$$\max(\left|\omega_i - \omega_j\right| / \omega_j) = x\% \qquad \forall i, j \tag{8}$$

c. interregional labor mobility, but the migration dynamics take into account that workers tend to move more easily within provinces and migrate more easily to closer-by locations. This takes results in Poncet (2006) seriously who shows empirically that Chinese migration patterns depend negatively on distance, and have a strong provincial bias. Instead of (7) we use (denoting by $m_{ij}$ the change a city's share in total population due to immigration from other lower wage cities or due to emigration to other higher wage cities):

$$\frac{d\lambda_i}{\lambda_i} = \lambda_i(1 + \sum_j m_{ij}) \text{ where } m_{ij} = \begin{cases} \left(\left|\omega_i - \omega_j\right|\right)^{2.35} D_{ij}^{-0.87} e^{0.96 NB_{ij}} & if \quad \omega_i > \omega_j \\ -\left(\left|\omega_i - \omega_j\right|\right)^{2.35} D_{ij}^{-0.87} e^{0.96 NB_{ij}} & if \quad \omega_i < \omega_j \end{cases} \tag{9}$$

where $NB_{ij}$ is a dummy variable indicating whether cities $i$ and $j$ are located in the same province (1=yes; 0=no).

Besides using either of these five different scenarios, we also consider possible combinations of these scenarios. For example, we combine migration dynamics as in (9) with incomplete real wage equalization as in (8). Or we combine restricted labor mobility between provinces (case 1b above), with migration dynamics within provinces defined by (9).

Before we present the results of our model simulations under different labor mobility scenarios, we would like to stress that these model simulations are certainly not aimed at giving a pin-point prediction as to what will happen to Chinese agglomeration patterns once more labor mobility is introduced. The model is too stylized to be able to perfectly predict which cities will "gain or lose" or by how much (in cases of some degree of labor mobility it generally tends to overpredicts the degree of agglomeration – see e.g. Bosker et al. (2010) for more on this). But, using the "no labor mobility" case as a benchmark *Hukou* case, we believe that at least in a qualitative sense the model can be used to better understand the implications of increased labor mobility for long-run agglomeration outcomes. It is here that our model simulations have value added compared to previous NEG-market access studies for China that take the spatial allocation of firms and workers as given.

### *6.3 China's internal economic geography under different labor mobility regimes*

Given the parameter values discussed above and the initial shares of city population $L_i$ and arable land $K_i$, we are now in a position to simulate NEG-based predictions regarding China's internal economic geography under different labor mobility regimes. Referring to Bosker et al. (2010) for the full details of the procedure used to simulate the model, we start with presenting the results under the assumption of no labor mobility (an extreme *Hukou* case) and subsequently relax the restrictions on inter-Prefecture city labor mobility.

#### *Case 1a: no interregional labor mobility – an extreme Hukou scenario*

As explained in section 3, the long run equilibrium under the assumption of labor immobility is such that wages are equalized between the manufacturing and agricultural sectors in each location. This simulation refers to an extreme *Hukou* scenario which is described by equations (1)-(3). Under such a scenario the spatial distribution of population remains as in Figure 2a, yet firms are able to move between cities in response to differences in profit opportunities. Our simulation results for the equilibrium spatial allocation of *firms* differ slightly from the actual distribution of firms across Prefecture cities as shown in Figure 2b. The correlation between the simulated distribution of manufacturing firms and the actual distribution in Figure 2b is 0.54. The overall degree of spatial concentration is smaller (Herfindahl index (HF) is 0.006 compared to the current HF of 0.010). The main difference with the actual distribution is that cities that are centrally located in provinces in the populous heartland of

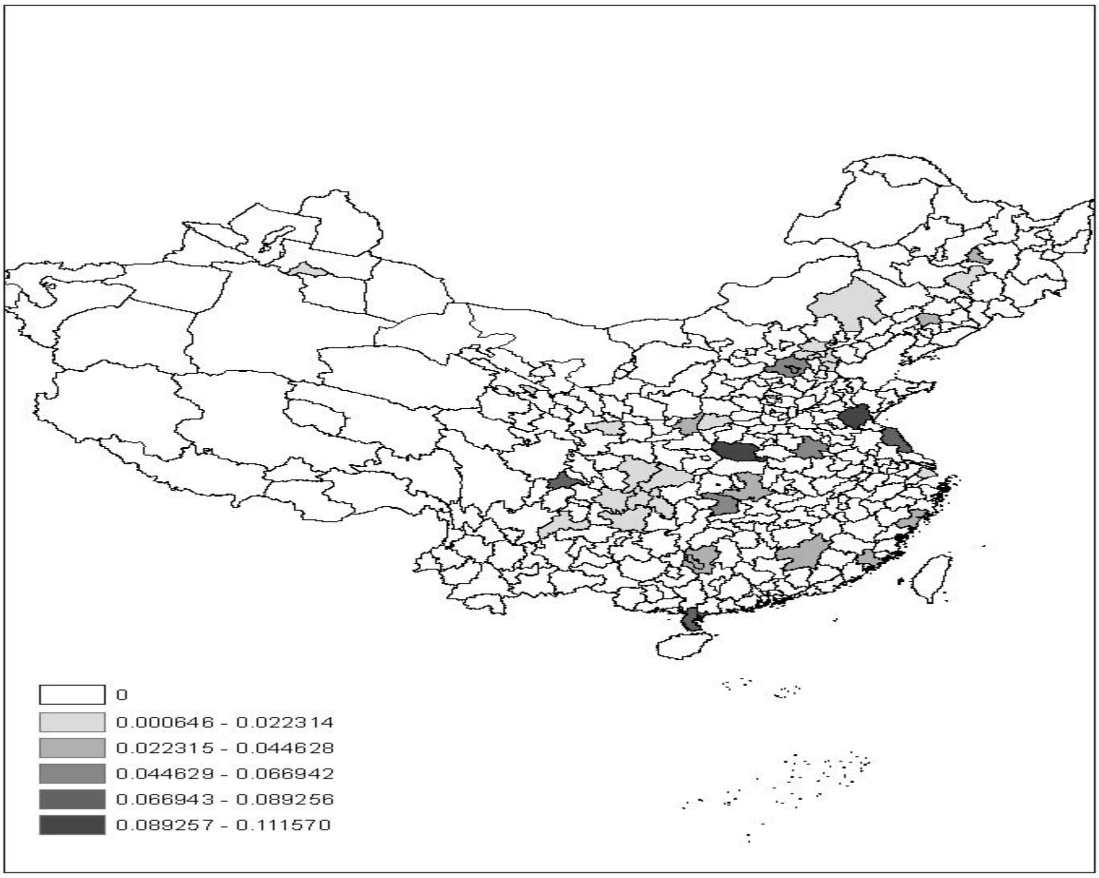
China (recall Figure 2a) slightly “gain” in terms of their firm share at the expense of the currently very large cities like Beijing, Tianjin or Shanghai, and also at the expense of more peripheral cities. The prefecture cities in the populous provinces of Hebei, Ahnwei, Honan and Hupeh gain the most compared to the initial distribution of industrial employment (their share of total industry activity increases by 7.5%). As Chinese internal demand continues to become more important, firms will locate close(r) to the largest sources of this demand in China’s populous heartland.

*Case 1b: only intraprovincial labor mobility*

Next, we loosen the restriction on interregional labor mobility somewhat and look at what happens if labor migration is allowed, but only within provinces. In this case, real wages are equalized within provinces, but not necessarily between provinces (i.e. we assume migration dynamics as in (4) to determine migration patterns between prefecture cities within the same province). This follows for instance Fujita et al. (2004) or Poncet (2006) who argue that provincial borders matter most for migration and that migration is restricted to nearby cities. The case of intra-Provincial labor mobility is also interesting because it has been suggested to offer a feasible policy option for Chinese policy makers for a (gradual) loosening of the Hukou system (Henderson et al, 2007, p. 9).

Figure 3 illustrates the resulting long run spatial distribution of firms under this scenario (the 0 category means no manufacturing firms at all left in equilbrium)

**Figure 3. Intra-provincial Migration – firm distribution**




All footloose activity now basically ends up in a single city within each Province, see the grey and black dots in Figure 3. The cities with largest firm (and population) shares are Linyi (Sjantung province) and Nanyang (Honan province). These cities gain over 7% points each in their share of industrial activity compared to their current share. This is a direct consequence of the fact that Sjantung and Honan are the two most populous provinces. Moreover, they are more centrally located within China (compared to e.g. the also populous provinces of Kwantung and Szetsjwan) so that firms producing there can not only take advantage of the large internal market of these provinces, but also face lower transport costs to consumers outside these provinces. In other words: firms locate there in response to the fact that these centrally located cities offer superior market access. This finding is consistent with the previous case of no labor mobility when firms already displayed a tendency to move to the cities in the centrally located populous provinces. Increasing interregional labor mobility only reinforces this pattern.

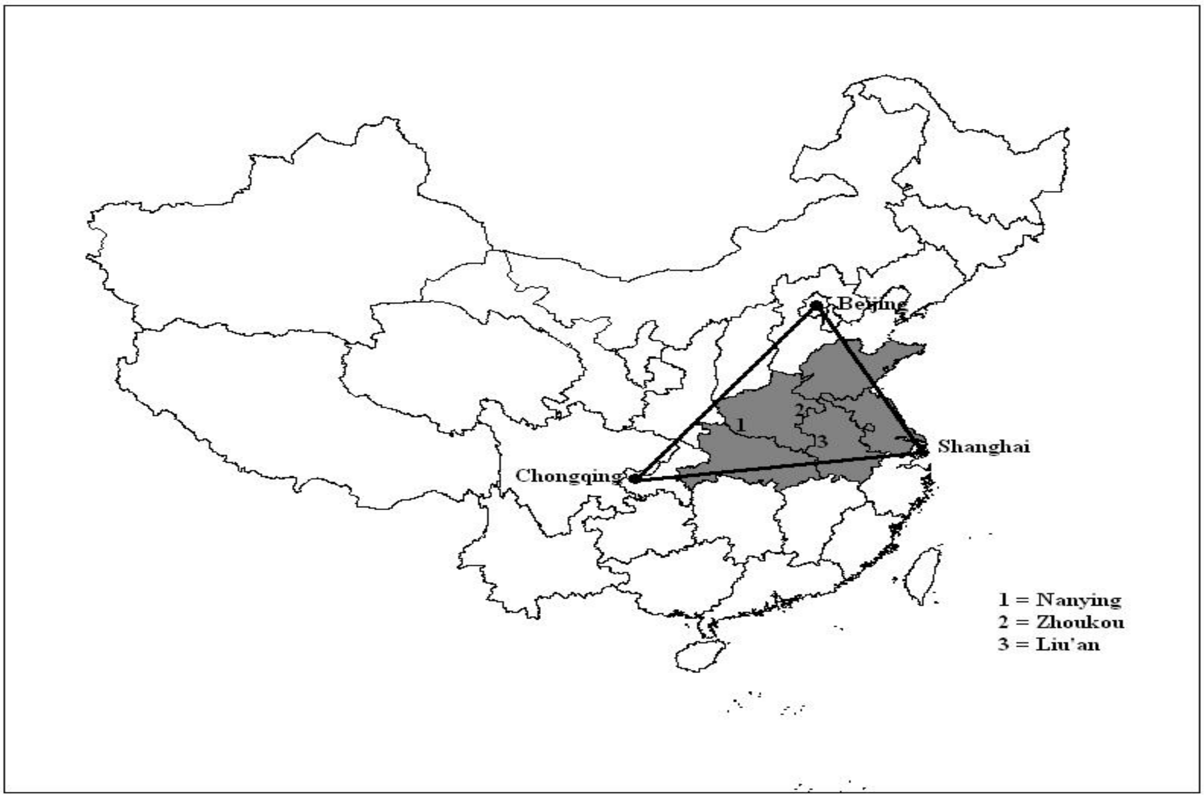
*6.4 A relaxation of the Hukou restrictions.*

Our next four scenarios all completely abandon the *Hukou* system and allow people to freely move between prefecture cities. They differ in the migration dynamics assumed, (7) or (9), and/or the assumption about people's reservation wage needed to make them decide to move to a different city, (4) or (8).

Though the details differ (and will be briefly discussed below), the upshot of the model simulations for all these 4 cases of unrestricted inter-city labor mobility across China is that one ends up with very different agglomeration outcomes compared to either the extreme *Hukou* case (1a above) or the provincially restricted *Hukou* system scenario (1b above). They indicate much stronger core-periphery patterns that these two other cases. Although suffering from the general NEG tendency to overstate the degree of agglomeration (in two of the four scenarios all footloose activity agglomerates in a single city), our simulation results lead to two key qualitative findings

First, all four scenarios strongly reinforce the tendency of firms to locate in the currently populous heartland of China (roughly within the Shanghai-Beijng-Chongqing triangle) that already showed (albeit much less pronounced) in scenario 1a and 1b. Firms locate close to large sources of demand that are further more centrally located within China so that other cities can be reached at low cost. However, in case of unrestricted labor mobility (especially when people move in response to the smallest wage differential) this pattern is further reinforced by people following firms to these places in search of higher wages. This in turn draws in more firms, etc. In the end production becomes concentrated in a few cities only. They are located in the (non-coastal) provinces of Honan and Anhwei that gain most in importance compared to their current share in overall economic activity.

**Figure 4. Increased agglomeration within the Shanghai-Beijng-Chongqing triangle**




This general prediction follows from each of our four different scenarios. They however differ in their exact prediction of where footloose activity ends up.

*Case 2a: migration dynamics as in (7) and migration in response to the smallest wage differentials (4)*

This is the opposite scenario to the extreme *Hukou* case. People are completely free as to where they locate. Moreover people's preferences are such that they will move in response to the smallest difference in wages between cities. This scenario basically takes the Puga (1999) NEG model at face value and assumes that labor migration is determined by real wage equalization condition (4) and the associated standard NEG migration dynamics (7). In this case the Prefecture city of Nanyang (Honan province) becomes the only industrial center of China. This tendency of NEG models to display (near) full agglomerations in case of labor mobility is well-known and too stark to be true (see Bosker et al., 2010). It is a result of the underlying equilibrium condition (4) that people move in response to the smallest wage differential.

*Case 2b: migration dynamics as in (9) and migration in response to the smallest wage differential as in (4)*

Again due to assumption (4) we end up with one city becoming China's sole industrial center. Yet in this case, it is the city of Zhoukou that attracts all footloose activity. This is the result of the different migration dynamics assumed. The distance penalty on migration imposed by (9), takes account of the empirical finding by Poncet (2006) that people prefer moves over shorter distances. This makes the adjustment process to the long run spatial equilibrium behaves more smoothly over space. As such the currently large population concentrations in China's coastal provinces move more slowly towards its populous heartland – so that a more easterly located city, i.e. Zhoukou, becomes the industrial centre compared to the case when assuming the much simpler, a-spatial, migration dynamics in (4).

*Cases 2c and 2d: migration dynamics as in (7) or (9) respectively and migration only in response to a sufficiently large wage differential (8)*

The extreme agglomeration outcomes in case 2a and 2b arise mainly from the (unrealistic) assumption of people being willing to move in response to any wage differential, however small. Case 2c and 2d relax this assumption and take account of the fact that migration is costly: we assume that a threshold of 10% in wage differences needs to be crossed before workers are willing to migrate, see (8). Imposing such a threshold basically increases the number of possible long-run equilibria because real wages do not have to be equalized. Any spatial allocation for which all inter-city real wage differentials are smaller than 10% will now be a long-run spatial equilibrium.

The result of doing this, is that we end up with less extreme agglomeration outcomes. We no longer find that all footloose industrial activity agglomerates in a single city. Still, most industrial activity is located in prefecture cities in the Honan province (most notably Nanyang and Zhoukou) or in the other cities in the triangle depicted in Figure 2. Yet, we now also find some, be it mostly only little, footloose economic activity in the currently very large agglomerations of Tianjin, Harbin, Shanghai, Chongqing, or Chengdu.

### *6.5 Sensitivity of our model simulations*

We performed *three sensitivity analyses* to verify the robustness of our findings. First, we allowed for the fact that market access is not the only determinant of Chinese wages and incorporated the estimation results for the city fixed effects and human capital into our model simulations as well. Second, we lowered the distance decay parameter in our distance function to capture in a simple way what would happen if transport costs will fall, due to for instance the large scale improvements in the inter-city infrastructure that are currently underway in many parts of China (see also Faber, 2009 and Roberts et al, 2010). In particular, we lowered our baseline distance decay parameter of 0.63 (see Table 2) to 0.36 or 0.1 respectively. Third, given the importance of the (arable) land variable in the Puga (1999) NEG model and the considerable (arable) land differences between the Prefecture cities in our sample, we also re-ran all simulations with all cities having the same (arable) land size (effectively removing differences between prefecture cities in the congestion force posed by differences in their land supply). In general, all these sensitivity analyses do not change the resulting equilibria in a qualitative sense.

Taking account of other (exogenously driven) productivity differences between cities as captured by our city fixed effects and human capital, still results in cities in China's populous heartland attracting the bulk of footloose economic activity. However, now the city of Liu'an in the Ahnwei province attracts the footloose activity in case of perfect labor mobility in response to the smallest real wage differential (that is, equations (4) and (7) hold). This city (see Figure 4 above) is located somewhat closer to the coast, reflecting the generally higher wages coastal locations can pay (captured by our city fixed effects) given their preferential location for exporting to world markets.

Lowering trade costs (lowering δ) generally reinforces the tendency of firms to agglomerate. This is in accordance with NEG's prediction that ceteris paribus lower trade costs tend to foster agglomeration. With lower trade costs firms can take advantage of co-locating in big agglomerations while incurring fewer costs of shipping their goods to consumers in the periphery.

Things also marginally change when we abstract from differences in arable land endowment between cities. In the presence of these differences, the same amount of workers poses stronger congestion problems to cities that are smaller in terms of their land endowment. Assuming each city to be of similar size, essentially equalizes the strength of congestion forces between cities.[27] Again most of our earlier results come through. However, we now find one notable exception. In the case of perfect labor mobility *Shanghai* attracts all footloose activity when migrants move in response to the smallest real wage differential (4).

[27] As explained in section 3 arable land differences are important in the NEG model of Puga (1999) because with decreasing returns in the production of the agricultural good, less arable land implies lower agricultural (rural) wages, implying that for manufacturing firms, *ceteris paribus*, it is relatively easy to lure workers away from the agricultural sector. Without these arable land differences the (relative) agglomeration force of regions with more arable land is absent.

However this is only the case when people move according to the simple migration rule in (9). When taking account of the empirical finding that internal migration in China is also determined by distance and has a strong provincial bias (using migration dynamics as in (9) based on Poncet, 2006), we no longer find that Shanghai attracts all footloose economic activity. Instead, and similar to our baseline case, the economic centre shifts back to the cities in the Anhwei or Honan province (the Prefecture cities of Liu'an or Zhoukou respectively, depending on whether or not we also allow for (exogenous) productivity differences between cities).

### *6.6 Summing up*

Overall, our simulation results show the value added of moving beyond confirming the spatial wage pattern predicted by NEG theory. Taking the full NEG model seriously, leads to interesting insights about the (possible) future of China's internal economic geography. It provides theory-based predictions as to how the spatial distribution of footloose economic activity may respond to different labor mobility regimes.

First, and similar to e.g. Au and Henderson (2006a) that use a different empirical strategy, we find that relaxing *Hukou*'s restrictions will lead to more pronounced core-periphery patterns. More interesting perhaps, our simulations show that the initial differences in market size do not need to be conclusive in determining the long-run equilibrium location of firms and workers once labor mobility is allowed for. We do not find that increased labor mobility will only strengthen the current agglomeration pattern in China. Instead, and although the NEG model probably overpredicts actual future agglomeration trends, an important insight from our results is that the spatial distribution of Chinese internal demand will be one of the important determinants of its future economic geography. Our results show a clear tendency of firms to agglomerate in currently medium-sized (from a Chinese perspective) cities, like Zhoukou, Nanjing or Liu'an in China's *populous heartland* (see Figure 4).

Based on simulations of a multi-region version of the core Krugman (1991) NEG model, Krugman (1993) already showed that initial size need not be decisive in determining which location(s) attract most (or all) footloose economic activity in equilibrium.[28] In case of China this explains why a Prefecture city like Liuan with "only" 6.3 million people can

[28] For a further illustration of the feature of NEG models that initial size need not be decisive for the final or equilibrium allocation of footloose activity, see for instance Brakman, Garretsen, and Van Marrewijk (2009), pp. 296-299.

become an agglomeration in the long-run. Liuan thrives upon its centrality. It offers firms and workers a superior location in terms of access to China's vast (and increasingly important) internal market.

## 7. Conclusions

In this paper we use a new economic geography (NEG) model to analyse the relationship between market access, labor mobility and agglomeration for China. Using this model we provide theory-based answers to the question how China's internal economic geography might be affected by a relaxation of its current *Hukou* system that puts severe restrictions on workers' interregional mobility.

Our paper crucially differs from related recent NEG studies for China (e.g. Ma, 2006, Lin, 2003, Hering and Poncet, 2010, 2007, De Sousa and Poncet, 2007, Amiti and Javorcik, 2008). We do not take the spatial allocation of labor (and firms) as given. Instead, we make use of the complete NEG model and not just the equilibrium (nominal) wage equation. Based on our estimates of the key NEG model parameters, we go one step further and simulate the full NEG model. This enables us to analyse the possible effects of relaxing China's *Hukou* system on its internal economic geography. In doing so, we also go beyond the simple migration dynamics that underlie most theoretical NEG models, and consider more realistic migration dynamics.

Our main findings show that a relaxation of the restrictions on labor mobility posed by the Hukou system will lead to more pronounced core-periphery outcomes. Interestingly, the economic geography of China that would result from such increased labor mobility is first and foremost determined by within China differences in market access, population and (arable) land. The geography of China's own internal demand is the most important determinant of its future economic geography. Given the increasing importance of China's own domestic market, international market access is of less significance in determining the spatial distribution of economic activity within China itself. Our simulations show that increased labor mobility would primarily benefit non-coastal Prefecture cities like Zhoukou or Liuan in the large and populous Honan and Anhwei provinces. These cities offer firms and workers superior access to China's vast internal market. The currently economically dominant prefecture cities like Shanghai, Tianjin or Guangzhou do not necessarily end up being the (only) main centers of economic activity once inter-city firm and labor mobility is taken into account. As such, our analysis suggests that the (spatial) distribution of China's own internal demand will be an important determinant of its future economic geography, in particular so when interregional labor mobility would become less restricted.

## References


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

Au, C-C. and J.V. Henderson, (2006a),Are Chinese Cities Too Small? *Review of Economic Studies*, Vol. 73, pp. 549-576.

Au, C-C. and J.V. Henderson (2006b),How migration restrictions limit agglomeration and productivity in China, *Journal of Development Economics*, Vol. 80, pp. 350-388.

Behrens, K. and J.F. Thisse (2007), Regional economics: A new economic geography Perspective, *Regional Science and Urban Economics*, 37, 457-465

Bosker, E.M. and H. Garretsen (2008), Economic geography and economic development in Sub-Saharan Africa, *CESifo Working paper,* 2940, CESifo Munich.

Bosker, E. M., S. Brakman, H. Garretsen, and M. Schramm (2010), Adding Geography to the New Economic Geography, *Journal of Economic Geography,* forthcoming {see also *CESifo Working Paper,* 2038, CESifo, Munich, 2007}

Brakman, S., H. Garretsen and M. Schramm (2004), The spatial distribution of wages: estimating the Helpman-Hanson model for Germany, *Journal of Regional Science,* 44(3), 437-466.

Brakman S., H. Garretsen, and M. Schramm (2006), Putting new economic geography to the test: free-ness of trade and agglomeration in the EU regions, *Regional Science and Urban Economics,* 36(5), 613-636.

Brakman, S. H.Garretsen, C.van Marrewijk (2009), *The New Introduction to Geographical Economics*, Cambridge University Press, Cambridge.

Breinlich H. (2006). The spatial income structure in the European Union – what role for Economic Geography? *Journal of Economic Geography*, Vol. 6, p. 593-617.

Chan, K.W. (2003), Urbanization in China in the 1990s: New Definitions, Different Series, and Revised Trends, *The China Review,* 3(2), pp. 49-71.

Chan, K. W. (2008), Internal Labor Migration in China: Trends, Geographical Distribution and Policies, Proceedings of *United Nations Expert Group Meeting on Population Distribution, Urbanization, Internal Migration and Development,* ESA/P/WP.206, United Nations, pp.93-122.

Chan, K.W. and W.Buckinham (2008), Is China Abolishing the Hukou System, Forthcoming, *The China Quarterly*, No.195.

Chan, K.W. (2009), The Chinese *Hukou* System at 50, *Eurasian Geography and Economics,* 50(2), pp. 197-221.

Crozet, M., 2004. Do migrants follow market potentials? An estimation of a new economic geography model. Journal of Economic Geography, Vol. 4 , 439– 458.

De Sousa, J and S. Poncet (2007), How are Wages Set in Beijing?, *CEPII Working Paper* 2007-13, CEPII, Paris.

Faber, B. (2009), Integration and the Periphery: The Unintended Effects of New Highways in a Developing Country, paper presented at Urban Economics Sessions at the NARSC 2009 conference 18-21 November 2009 San Francisco.

Fujita M., T. Mori, J.V. Henderson and Y. Kanemoto (2004), Spatial Distribution of Economic Activities in Japan and China in J. V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics, volume IV,* North Holland, Amsterdam, 2911-2980.

Fujita, M. and T. Mori (2005), Frontiers of the new economic geography, *Papers in Regional Science,* 84(3), 377-407.

Hanson, G.H. (2005), Market potential, increasing returns, and geographic concentration, *Journal of International Economics,* 67(1), 1-24.

Helpman, E. (1998), The size of Regions, in: D.Pines, E.Sadka, and I. Zilcha (eds.), Topics in Public Economics, Cambridge University Press, pp. 33-54.

Henderson, J.V, with J. Quigley and E Lim, (2007), *Urbanization in China: Policy Issues and Options,* China Economic Research and Advisory Programme, available at http://www.econ.brown.edu/faculty/henderson/FinalFinalReport-2007050221.pdf

Henderson, J.V. (2009), Cities and Development, 50(1)), *Journal of Regional Science,* pp. 515-540.

Hering, L. and S. Poncet (2010), Market Access and Individual Wages: Evidence from China, Vol. 92( 1), *Review of Economics and Statistics,* 145-159.

Hering, L. and S. Poncet (2007), Economic Geography, Spatial Dependence, and Income Inequality in China. *mmeo*. Université Paris I Panthéon-Sorbonne.

Hu, D. (2002), Trade, rural-urban Migration, and Regional Income Disparity in Developing countries: a spatial general Equilibrium model inspired by the case of China, *Regional Science and Urban Economics*, Vol.32, pp. 311-338.

Knaap. T., 2006. Trade, location, and wages in the United States. *Regional Science and Urban Economics,* 36(5), p. 595-612.

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

Krugman, P.R. (1993), On the Number and Location of Cities, *European Economic Review,* 37, 293-298.

Krugman, P.R. and A.J. Venables (1995), Globalization and the inequality of nations, *The Quarterly Journal of Economics*, vol 110, 857-880.

Krugman, P.R. (2009), Increasing Returns in a Comparative Advantage World, November 2009, available at http://www.princeton.edu/~pkrugman/deardorff.pdf

Krugman, P.R. (2010), The New Economic Geography, Now Middle-Aged, April 2010, available at http://www.princeton.edu/~pkrugman/aag.pdf

Lin, S. (2003), International Trade, Location, and Wage Inequality in China, *WIDER Discussion Paper,* no. 2003/61, WIDER United Nations University.

Ma, A.C. (2006), Geographical Location of Foreign Direct Investment and Wage Inequality in China, *The World Economy,* pp. 1031-1055.

Moreno Monroy, A.I. (2010), Market access and the heterogeneous effect of shocks on wages: Evidence from Chinese cities, forthcoming in *Papers in Regional Science*.

Poncet, S. (2006), Provincial migration dynamics in China: Borders, costs and economic motivations, *Regional Science and Urban Economics*, 36, pp. 385– 398.

Puga, D. (1999), The rise and fall of regional inequalities, *European Economic Review,* Vol. 43, 303-334.

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

Roberts, M., U. Deichmann, B. Fingleton, and T. Shi, 2010, The Impact of Infrastructure on Economic Development: Using the New Economic Geography to Evaluate China's Highway Expansion, *mimeo,* Univ of Cambridge/World Bank.

Venables, A.J. (1996), Equilibrium Locations of Vertically Linked Industries, *International Economic Review*, 37, 341-359.

Whalley, J. and S.Zhang (2007), A numerical simulation analysis of (Hukou) labor mobility restrictions in China, *Journal of Development Economics*, Vol. 83, pp. 392-410.

World Bank, 2008, *World Development Report 2009*, Washington.

## Appendix A1. Basic set up of the M-region version of the Puga (1999) model

Consider a world consisting of R regions. Each of region $i = 1,\ldots,R$ is populated by $L_i$ workers and endowed with $K_i$ units of *arable* land. Each region's economy consists if two sectors: agriculture and industry. Labor is used by both sectors and is mobile between sectors within a region and it is either mobile or immobile between regions. Land on the other hand is used only by the agricultural sector and is immobile between regions.

*Production*

The agricultural good is produced under perfect competition and free entry and exit using Cobb-Douglas technology. Moreover it is freely tradable between regions.

The industrial sector produces heterogeneous varieties of a single good under monopolistic competition and free entry and exit. Industrial production technology is characterized by increasing returns to scale, i.e. production of a quantity *x(h)* of any variety *h* requires fixed costs $c_i\,\alpha$ and variable costs $c_i\beta$ that are both assumed to be the same in each region, but can differ between regions due to differences in e.g. production efficiency, $c_i$. This, together with free entry and exit and profit maximization, ensures that in equilibrium each variety is produced by a single firm in a single region. The production input is a Cobb-Douglas composite of labor and intermediates, with $0 \leq \mu \leq 1$ the Cobb-Douglas share of intermediates. Intermediates enter the production function as a composite manufacturing good that is specified as a CES-aggregate (with $\sigma > 1$ the elasticity of substitution across varieties) of all manufacturing varieties produced.

Firms in principle sell their goods to all regions. But, shipping their goods to foreign markets incurs so-called 'iceberg' trade costs ($\tau_{ij} \geq 1$ goods have to be shipped from region $i$ to let one good arrive in region $j$). Taking these costs into account, gives the following profit function that is similar for each firm in region $i$:

$$\pi_i = \sum_j^R p_{ij}(h)x_{ij}(h)/T_{ij} - w_i^{M\,1-\mu} q_i^{\mu} c_i[\alpha + \beta\sum_j^R x_{ij}(h)] \tag{1}$$

where $p_{ij}(h)$ is the price of a variety produced in country $i$, $q_i$ is the price index of the composite manufacturing good, $w_i^M$ the manufacturing wage in region $i$.

*Preferences*

Consumers have Cobb-Douglas preferences over the agricultural good and a CES-composite (also with $\sigma > 1$ the elasticity of substitution across varieties) of manufacturing varieties, with $0 \leq \gamma \leq 1$ the Cobb-Douglas share of the composite manufacturing good. Specifying the composite manufacturing good this way ensures demand from each region for each manufacturing variety, which, together with the fact that each variety is produced by a single firm in a single region, implies that trade takes place between regions.

*Equilibrium*

Having specified preferences over, and the production technologies of, the manufacturing and agricultural good, the equilibrium conditions of the model can be calculated. Profit maximization and free entry and exit determine the share of labor employed, $L_i^A$, and the wage level, $w_i^A$, in agriculture, as well as the rent earned per unit of land $r(w_i^A)$. The former two in turn pin down the share of workers in manufacturing, $\varsigma_i$. Given the assumed Cobb-Douglas production function in agriculture, with labor share $\theta$, we have that:

$$\varsigma_i = \frac{L_i^M}{L_i} = 1 - \frac{L_i^A}{L_i} = 1 - \frac{K_i}{L_i}\left(\frac{\theta}{w_i^A}\right)^{\frac{1}{1-\theta}} \tag{2}$$

where $0 \leq \theta \leq 1$ denotes the Cobb-Douglas share of labor in agriculture, and $L_i^M$ and $L_i^A$ the number of workers in manufacturing and agriculture respectively. Equation (2) shows that, in contrast to Krugman (1991), where agriculture uses only land ($\theta = 0$), or to Venables (1996), where agriculture employs only labor ($\theta = 1$), the share of a region's labor employed in manufacturing is endogenously determined in this model. It increases with a region's labor endowment and agricultural wage level and decreases with a region's land endowment and with the Cobb-Douglas share of labor in agricultural production. Consumer preferences in turn determine total demand for agricultural products in region $i$ as:

$$x_i^A = (1-\gamma)Y_i \tag{3}$$

In the industrial sector, profit maximization and free entry and exit, gives the familiar result that all firms in region $i$ set the same price for their produced manufacturing variety as being a constant markup over marginal costs:

$$p_i = \frac{\sigma\beta}{\sigma - 1} c_i q_i^{\mu} w_i^{M(1-\mu)} \tag{4}$$

where $q_i$ is the price index of the composite manufacturing good in region $i$ defined by:

$$q_i = \left(\int_j \tau_{ij}^{1-\sigma} n_j p_j^{(1-\sigma)}\right)^{\frac{1}{1-\sigma}} \tag{5}$$

where $n_i$ denotes the number of firms in region $i$ and

$$w_i^M = \left[(1-\mu)n_i p_i\left(\frac{(\sigma-1)}{\sigma\beta}(\alpha+\beta x_i)\right)\right](\varsigma_i L_i)^{-1} \quad (6)$$

is the manufacturing wage level in region *i.*

Utility maximization on behalf of the consumers in turn gives total demand for each manufacturing variety produced (coming from both the home region *i* as well as foreign regions *j*) which is the same for each variety in the same region due to the way consumer preferences are specified:

$$x_i = \int_j p_i^{-\sigma} e_j q_j^{(\sigma-1)} \tau_{ij}^{1-\sigma} \quad (7)$$

where in (7) demand from each foreign region *j* is multiplied by $\tau_{ij}$ because ($\tau_{ij}$ –1) of the amount of the product ordered from region *i* melts away in transit (the iceberg assumption).

$$e_i = \gamma Y_i + \mu n_i p_i\left(\frac{(\sigma-1)}{\sigma\beta}(\alpha+\beta x_i)\right) \quad (8)$$

is total expenditure on manufacturing varieties in region i (the first term representing consumer expenditure and the second term producer expenditure on intermediates), where

$$Y_i = w_i^A(1-\varsigma_i)L_i + w_i^M \varsigma_i L_i + r(w_i^A)K_i + n_i\pi_i \quad (9)$$

is total consumer income consisting of workers' wage income, landowners' rents and entrepreneurs' profits respectively. Due to free entry and exit these profits are driven to zero, which (after substituting for wages, tthereby uniquely defining a firm's equilibrium output at:

$$x_i = \alpha(\sigma-1)/\beta \quad (10)$$

Finally, to close the model, the labor markets are assumed to clear:

$$L_i = L_i^M + L_i^A = \underbrace{\left[(1-\mu)n_i p_i\left(\frac{(\sigma-1)}{\sigma\beta}(\alpha+\beta x_i)\right)\right]w_i^{M-1}}_{L_i^M} + \underbrace{K_i\left(\frac{\theta}{w_i^A}\right)^{\frac{1}{1-\theta}}}_{L_i^A} \quad (11)$$

where the demand for labor in agriculture, $L_i^A$, follows from the assumption of Cobb-Douglas technology in agriculture and the term between square brackets represents the total manufacturing wage bill. Moreover equating labor supply to labor demand in the industrial sector gives an immediate relationship between the number of firms and the number of workers in industry:

$$n_i = \frac{\varsigma_i L_i}{\alpha\sigma(1-\mu)q_i^{\mu} w_i^{M-\mu}} \quad (12)$$

*Long run equilibrium*

Next, to solve for the long run equilibrium, Puga (1999) distinguishes between the case where labor is both interregionally and intersectorally mobile and the case when it is only intersectorally mobile. Without interregional labor mobility, long run equilibrium is reached when the distribution of labor between the agricultural and the industrial sector in each region is such that wages are equal in both sectors. This is ensured by labor being perfectly mobile between sectors driving intersectoral wage differences to zero. When instead labor is also interregionally mobile, not only intersectoral wage differences are driven to zero in all regions in equilibrium. Workers now also respond to real wage (utility) differences between regions by moving to regions with the higher real wages (utility) until real wages are equalized between all regions, hereby defining the long run equilibrium

*Interregional labor immobility*

The long run equilibrium in case of interregional labor immobility can now be shown to be a solution $\{w_i,q_i\}$ of three equations that have to hold in each region[29]. In our case (when using wage-worker space) these are, using the fact that in equilibrium $w_i^M = w_i^A = w_i$ :

$$q_i = \left( \frac{1}{1-\mu} \sum_j \left( \varsigma_j L_j q_j^{-\mu\sigma} c_j^{-\sigma} w_j^{1-\sigma(1-\mu)} \tau_{ij}^{1-\sigma} \right) \right)^{1/(1-\sigma)} \qquad (13)$$

$$w_i = q_i^{\mu/(\mu-1)} c_i^{1/(\mu-1)} \left( \sum_j e_j q_j^{\sigma-1} \tau_{ij}^{1-\sigma} \right)^{1/(\sigma(1-\mu))} \qquad (14)$$

$$e_i = \gamma(w_i L_i + K_i r(w_i)) + \mu/(1-\mu) w_i \varsigma_i L_i \qquad (15)$$

where (13) is obtained by substituting (4) and (12) into (5), (14) by substituting (4) and (10) into (7), and (15) by substituting (4), (10) and (12) into (8).

*Interregional labor mobility*

In case of interregional labor mobility, a solution to (13)-(15) merely constitutes a short run equilibrium. With interregional labor mobility, workers will move between regions in response to real wage differences until the interregional real wage differences, that are possible to persist when workers are unable (or unwilling) to move between regions, are no longer present. More formally, the LRE solution $\{w_i,q_i\}$ for each region $i$ has to adhere to the additional condition that real wages, $\omega_i$, are equal across all regions:

$$\omega_i = q_i^{-\gamma} w_i = \omega \qquad \forall i \qquad (16)$$

[29] Without loss of generality, α and β are set at 1/σ and (σ-1)/ σ respectively.

## Appendix A2. Chinese provinces and the Prefecture cities in our data set

| Beijing | Tianjin | Hebei | Sjansi | Inner Mongolian Aut. Region | Liaoning | Kirin | Heilunkiang | Shanghai | Kiangsu | Tsekiang | Anhwei |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Beijing | Tianjin | Shijiazhuang | Taiyuan | Hohhot | Shenyang | Changchun | Harbin | Shanghai | Nanjing | Hangzhou | Hefei |
| | | Tangshan | Datong | Baotou | Dalian | Jilin city | Qiqihar | | Wuxi | Ningbo | Wuhu |
| | | Qinhuangdao | Yangquan | Wuhai | Anshan | Siping | Jixi | | Xuzhou | Wenzhou | Bengbu |
| | | Handan | Changzhi | Chifeng | Fushun | Liaoyuan | Hegang | | Changzhou | Jiaxing | Huainan |
| | | Xingtai | Jincheng | Tongliao | Benxi | Tonghua | Shuangyashan | | Suzhou | Huzhou | Maanshan |
| | | Baoding | Shuozhou | | Dandong | Changbaishan | Daqing | | Nantong | Shaoxing | Huaibei |
| | | Zhangjiakou | Yuncheng | | Jinzhou | Songyuan | Yinchun | | Lianyungang | Jinhua | Tongling |
| | | Chengde | Xinzhou | | Yingkou | Baicheng | Jiamusi | | Huaian | Quzhou | Anqing |
| | | Cangzhou | Linfen | | Fuxin | | Qitaihe | | Yancheng | Zhoushan | Huangshan |
| | | Langfang | | | Liaoyang | | Mudanjiang | | Yangzhou | Taizhou | Chuzhou |
| | | Hengshui | | | Panjin | | Heihe | | Zhenjiang | Lishui | Fuyang |
| | | | | | Tieling | | Suihua | | Taizhou | | Suzhou |
| | | | | | Chaoyang | | | | Suqian | | Chaohu |
| | | | | | Huludao | | | | | | Liuan |
| | | | | | | | | | | | Bozhou |

| Fujian | Kiangsi | Sjantung | Honan | Hupeh | Hunan | Kwantung | Kwangsi | Hainan | Szetsjwan | Kweitsjou | Yunnan |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Fuzhou | Nanchang | Ji'nan | Zhengzhou | Wuhan | Changsha | Guangzhou | Nanning | Haikou | Chengdu | Guiyang | Kunming |
| Xiamen | Jingdezhen | Qingdao | Kaifeng | Huangshi | Zhuzhou | Shaoguan | Liuzhou | Sanya | Zigong | Liupanshui | Qujing |
| Putian | Pingxiang | Zibo | Luoyang | Shiyan | Xiangtan | Shenzhen | Guilin | | Panzhihua | Zunyi | Yuxi |
| Sanming | Jiujiang | Zaozhuang | Pingdingshan | Yichang | Hengyang | Zhuhai | Wuzhou | | Luzhou | Anshun | Zhaotong |
| Quanzhou | Xinyu | Dongying | Anyang | Xiangfan | Shaoyang | Shantou | Beihai | | Deyang | | Baoshan |
| Zhangzhou | Yingtan | Yantai | Hebi | Ezhou | Yueyang | Foshan | Fangchenggang | | Mianyang | | |
| Nanping | Ganzhou | Weifang | Xinxiang | Jingmen | Changde | Jiangmen | Qingzhou | | Guangyuan | | |
| Longyan | Ji'an | Jining | Jiaozuo | Xiaogan | Zhangjiajie | Zhanjiang | Guigang | | Suining | | |
| Ningde | Yichun | Taian | Puyang | Jingzhou | Yiyang | Maoming | Yulin | | Neijiang | | |
| | Shangrao | Weihai | Xuchang | Huanggang | Chenzhou | Zhaoqing | Hezhou | | Leshan | | |
| | | Rizhao | Luohe | Xianning | Yongzhou | Huizhou | Baise | | Nanchong | | |
| | | Laiwu | Sanmenxia | Suizhou | Huaihua | Meizhou | Hechi | | Yibin | | |
| | | Linyi | Nanyang | | Loudi | Shanwei | | | Dazhou | | |
| | | Dezhou | Shangqiu | | | Heyuan | | | Yaan | | |
| | | Liaocheng | Xinyang | | | Yangjiang | | | Bazhong | | |

**Table continued**

| Fujian | Kiangsi | Sjantung | Honan | Hupeh | Hunan | Kwantung | Kwangsi | Hainan | Szetsjwan | Kweitsjou | Yunnan |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Binzhou | Zhoukou | | | Qingyuan | | | Ziyang | | |
| | | Heze | Zhumadian | | | Dongguan | | | | | |
| | | | | | | Zhongshan | | | | | |
| | | | | | | Chaozhou | | | | | |
| | | | | | | Jieyang | | | | | |
| | | | | | | Yunfu | | | | | |

| Shaanxi | Kansu | Tsinghai | Ningsia | Xinjiang Uyghur Aut. Region | Chongqing |
|---|---|---|---|---|---|
| Xi'an | Lanzhou | Xining | Yinchuan | Urumqi | Chongqing |
| Tongchuan | Jiayuguan | | Shizuishan | Karamay | |
| Baoji | Jinchang | | Wuzhong | | |
| Xianyang | Baiyin | | | | |
| Weinan | Tianshui | | | | |
| Yan'an | Jiuquan | | | | |
| Hanzhong | Zhangye | | | | |
| Yulin | Wuwei | | | | |
| Ankang | Pingliang | | | | |


Administrative Divisions of the
People's Republic of China (PRC)
XINJIANG UYGHUR
AUTONOMOUS REGION
Ürümqi
Indian claim line
TIBET
AUTONOMOUS
REGION
Lhasa
Chinese claim line
QINGHAI
Xining
GANSU
Lanzhou
NINGXIA
HUI
AUTONOMOUS
REGION
Yinchuan
INNER MONGOLIA
AUTONOMOUS REGION
Hohhot
HEILONGJIANG
Harbin
JILIN
Changchun
LIAONING
Shenyang
BEIJING
TIANJIN
HEBEI
Shijiazhuang
SHANXI
Taiyuan
SHANDONG
Jinan
SHAANXI
Xi'an
HENAN
Zhengzhou
JIANGSU
Nanjing
SHANGHAI
ANHUI
Hefei
HUBEI
Wuhan
SICHUAN
Chengdu
CHONGQING
ZHEJIANG
Hangzhou
JIANGXI
Nanchang
HUNAN
Changsha
GUIZHOU
Guiyang
FUJIAN
Fuzhou
YUNNAN
Kunming
GUANGXI
ZHUANG
AUTONOMOUS
REGION
Nanning
GUANGDONG
Guangzhou
HONG KONG S.A.R.
MACAU S.A.R.
HAINAN
Haikou
TAIWAN
controlled by the
Republic of China;
claimed by the People's
Republic of China.
Province
Autonomous Region
Municipality
Special Administrative Region