Hukou and Highways

The Impact of China's Spatial Development Policies on Urbanization and Regional Inequality

Maarten Bosker

Uwe Deichmann

Mark Roberts

WORLD BANK GROUP

Development Research Group
Environment and Energy Team
June 2015

Abstract

China has used two main spatial policies to shape its geographic patterns of development: restricted labor mobility through the Hukou residential registration system and massive infrastructure investment, notably a 96,000 kilometer national expressway network. This paper develops a structural new economic geography model to examine the impacts of these policies. Fitting the model to available data allows simulating counterfactual scenarios comparing each policy's respective impact on regional economic development and urbanization patterns across China. The results suggest large overall economic benefits from constructing the national expressway network and abolishing the Hukou system. Yet,

the spatial impacts of the two policies are very different. The construction of the national expressway network reinforced existing urbanization patterns. The initially lagging regions not connected to the network have not benefitted much from its construction. By contrast, removal of the Hukou restrictions, which Chinese policy makers are considering, would result in much more widespread welfare gains, allowing everyone to gain by moving to where he or she is most productive. Removal of the Hukou restrictions would also promote urbanization in currently lagging (inland) regions, mostly by stimulating rural to urban migration.

This paper is a product of the Environment and Energy Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at bosker@ese.eur.nl, udeichmann@worldbank.org, and mroberts1@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Hukou and Highways

The Impact of China's Spatial Development Policies on Urbanization and Regional Inequality

Maarten Bosker¹, Uwe Deichmann² and Mark Roberts³

JEL Codes: R11, R23, R42

Key words: China, transport investments, highways, migration, new economic geography

Sectors: Transport, Urban Development

¹ Department of Economics, Erasmus University Rotterdam, The Netherlands and CEPR <bosker@ese.eur.nl>

² Development Research Group, World Bank, Washington, D.C., USA <udeichmann@worldbank.org>;

³ Social, Urban, Rural and Resilience Global Practice, World Bank, Washington, D.C., USA
<mroberts1@worldbank.org>.

The authors thank Harry Garretsen, Vernon Henderson, Bert Hofman, Laura Hering, Diego Puga, Karlis Smits, Maisy Wong, and participants at the 2nd Urbanization and Poverty Reduction Research Conference and the 2014 North American Regional Science Conference for helpful comments and suggestions. Financial support from the World Bank's Knowledge for Change Program is gratefully acknowledged.

1. Introduction

In recent decades, China's economy has been characterized by rapid economic development accompanied by equally rapid urbanization. The resulting benefits have not, however, been shared equally across the country. Many rural areas saw considerable improvements in living standards and large reductions in poverty rates, but gains in urban areas and those in China's coastal regions in particular have been considerably larger. The Chinese government has always had a keen interest in the country's spatial economic development. Its main policies have been aimed at restricting the flow of migrants to the big cities, while at the same time trying to develop interior regions by better connecting them to the booming coastal regions.

Two policies stand out in this regard.¹ The first is the Hukou, or household registration, system. Every Chinese citizen's Hukou status is determined by that of their parents at the time they were born. A Hukou is tied to a particular urban or rural location and represents an entitlement to welfare benefits and public services (such as education, health care, etc.) in that place. When migrating to a different city, a person's Hukou status does not change, so that migrants are unable to claim many welfare benefits or public services in their destination city. Despite the Hukou system, many rural migrants have still migrated to the cities. A 2013 survey by China's National Bureau of Statistics estimated a "population that is separated from their household registration" of 289 million (NBS 2014, also Chan, 2013). There is growing consensus that if China's continuing urbanization process is to be economically productive and socially inclusive, the Hukou system—put in place to restrict migrant flows into the cities—will have to be loosened or even abolished (World Bank and DRC 2014). Chinese policy makers seem to agree, and have recently issued significant reform proposals (State Council, 2014)

The second policy is the huge investment in large-scale infrastructure. In recent decades China has constructed an extensive 96,000 km network of highways connecting the largest cities in the country, vast intra-city transport infrastructure (ring roads, metros), and is currently building the longest high speed rail network in the world connecting its main population centers (see Baum-Snow and Turner, 2012; Roberts *et al.*, 2012, Baum-Snow *et al.* 2012; Zheng and Kahn, 2013; Faber, 2014; among others). The explicit aim of these projects, and that of the construction of the highway network in particular, is to spread development from the more developed coast to the cities in the interior. They make it cheaper for these cities to import materials and intermediate goods, and to ship their own products to the rest of China and the world. Indirectly, this policy is also aimed at alleviating migration pressures. By

¹ Another is land reform. The risk of losing one's plot of land keeps many rural residents from moving to the city, even if they were able to obtain an urban Hukou. Land reform, whereby people would be able to sell their land would be a way to alleviate this problem.

contributing to the economic development of China's interior, workers' incentives to migrate to the coastal cities are expected to diminish as they can now find (well-paid) jobs in their own region.

In this paper, we quantify the effects of these two main spatial development policies on China's economic geography. Building on earlier work by Roberts *et al.* (2012), Behrens *et al.* (2013), and Tabuchi and Thisse (2002), we incorporate labor mobility into a structural new economic geography (NEG) model that allows us to assess how the rapid construction of the intercity national expressway network (NEN) and the Hukou system have jointly shaped China's spatial economy. First, we fit this model to the data using information on the rural and urban part of each of China's 331 prefectures. Next, with our estimates of the important model parameters in hand, we simulate various counterfactual scenarios that allow us to pinpoint the impact of the NEN and the Hukou system on the distribution of both people and economic activity across China's prefectural areas, as well as on rural-urban income inequality, and urbanization rates in different parts of the country.

We are not the first to look at the aggregate and/or spatial impacts of either the Hukou system or the NEN. However, earlier papers have looked at these issues separately. Banerjee *et al.* (2012), Roberts *et al.* (2012), and Faber (2014)² focus on the effect of the NEN. Roberts *et al.* (2012), using a similar combination of estimation and calibration as in this paper, find that the NEN did increase aggregate Chinese welfare. However, its construction did not decrease real income inequality between prefectures, nor reduce urban-rural wage inequality. Faber (2014) also finds that the NEN has reinforced the concentration of economic activity in the largest cities.³ But the peripheral regions, while losing economic activity, have also gained better access to the products produced in the industrial centers thanks to the NEN. Taking both into account, he finds positive welfare impacts of the NEN in all Chinese regions.

Whalley and Zhang (2007) and Bosker *et al.* (2012) instead focus on the spatial economic consequences of the Hukou system. Given that a counterfactual China without Hukou restrictions is fundamentally unobserved, these papers establish the effect of the Hukou restrictions by comparing China's regional economy today to various simulated "no Hukou-scenarios".⁴ Whalley and Zhang's (2007) results

² Baum-Snow, *et al.* (2012) is another prominent example. That paper focusses on the effect of the extent and configuration of intra-city infrastructure on urban form. The effect of intra-city infrastructure lies beyond the scope of our paper. Coşar and Fajgelbaum (2013) provide a model and some evidence that China's internal economic geography is affected by the accessibility of the interior cities to the internationally well-connected coastal regions. An improvement in domestic infrastructure in their setup would result in a further migration to the coast. They do not, however, explicitly consider the effects of the NEN.

³ Faber (2014) combines reduced form estimates of the causal effect of the NEN on economic activity and population growth, with the calibration of a spatial economic model that allows him to generalize his reduced form estimates. Banerjee *et al.* (2012) use a very similar empirical strategy as employed by Faber (2014). Unlike Faber (2014), however, they do not calibrate the model they propose, only showing reduced form estimates.

⁴ Desmet and Rossi-Hansberg (2013) can also be considered to simulate spatial economic outcomes in China under a "no-Hukou" scenario. Their model assumes free labor mobility across prefecture cities throughout, but they argue that the migration restrictions can be captured by their city-specific estimates of amenities. Their "equal

suggest that a removal of the Hukou-restrictions would increase overall Chinese welfare, while at the same time substantially increasing real income inequality between provinces. The latter is the result of people moving from the currently underdeveloped interior to the richer provinces on the coast. Bosker *et al.* (2012) come to a similar conclusion when analyzing the effect of the Hukou system at a spatially more disaggregated level (294 prefecture cities). They focus exclusively on the spatial distribution of people across China's prefecture cities, and find that a removal of the Hukou restrictions would strongly reinforce the core-periphery pattern that already exists today.

Our paper's contribution to this existing body of literature is twofold. First, by explicitly considering both of China's two main spatial development policies within the same framework, we are able to assess their effects without abstracting from the other spatial policy.⁵ The earlier papers considering the impact of the NEN (Faber, 2014 and Roberts *et al.*, 2012) all assume that people are completely immobile. This, in effect, represents an extreme Hukou scenario⁶: in reality many people did decide to migrate even in the presence of the Hukou restrictions. Abstracting from any migration response to changes in real income as a result of the construction of the NEN, may lead to seriously over- or understating the (spatial) economic impact of the NEN. Some of the NEN's immediate impact on the agglomeration of economic activity may e.g. have been reinforced as people, despite the Hukou restrictions, decided to move to places that, because of the NEN, now offer higher expected real incomes.

The earlier papers considering the impact of the Hukou system instead abstract from China's enormous investments in transport infrastructure. Whalley and Zhang (2007) base their analysis on a regional economic model that abstracts from the large differences in accessibility across China. Bosker *et al.* (2012) do take trade costs into account, but approximate them by using the great-circle distance between prefectures. As such, they still abstract from the large differences in actual trade costs across China that are the result of the unequal investment in both the quality and quantity of infrastructure (and that have been shown by Roberts *et al.* (2012) and Faber (2014) to significantly affect regional economic outcomes).

In our model instead, we use detailed information on the travel times between prefectures before and after the construction of the NEN as a more direct measure of trade costs between prefectures⁷.

amenities across cities" could then be viewed as a tentative no-Hukou scenario: they find an increase of overall welfare in that scenario accompanied by a more unequal city size distribution (large cities become larger, and small cities become smaller). Their model however also abstracts from differences in cities' accessibility, i.e. costs involved in shipping goods into or out of the city, focusing instead on within-city frictions (congestion).

⁵ Moreover, we can for example also assess whether the construction of the NEN has changed the expected effect of abolishing the Hukou restrictions.

⁶ Both papers actually justify their assumption of complete labor immobility by arguing that the Hukou system effectively keeps (almost) everyone from migrating.

⁷ Note that using these travel times can still be argued not to be the ideal measure of trade costs. It tells something about how long it takes to get something from one place to another, which need not necessarily be reflected in the costs paid to ship something along that route.

Moreover, we explicitly allow for migration both within (from rural to urban areas) and between prefectures. Here we follow Behrens *et al.* (2013) and Tabuchi and Thisse (2002), and model people's migration decisions as depending on real wage- and amenity-differences between locations, as well as on individual-specific idiosyncratic preferences for living in a particular location.⁸ Only relatively few studies estimate the relative importance of these different factors for people's migration decisions in the Chinese context. One important reason for this is the absence of comprehensive data on bilateral migration flows (either urban-rural migration flows or inter-prefectural migration flows).⁹ Relying on the equilibrium conditions of our model, shows that we can estimate their importance using readily available data on the stock of migrants in the rural and urban part of each prefecture only. Our results confirm the notion that migration in China is predominantly driven by people in search of higher real wages, and better provision of public amenities. Migration within the same province responds more strongly to these factors than migration to prefectures in other provinces (consistent with findings in Zhang and Zhao, 2013).

The second contribution of our paper is that our model and accompanying data set allow us to provide a much more complete picture of the impact of the NEN and/or Hukou system on China's spatial development. Given that earlier papers considering the impact of the NEN all assume regional labor immobility, they only focus on the NEN's effect on the spatial distribution of real income - either between prefectures in Roberts *et al.*, (2012) and Faber (2014), or between the urban and rural areas within prefectures (Roberts *et al.*, 2012). Bosker *et al.*, (2012), instead focuses exclusively on the effect of the Hukou on the spatial distribution of people and firms across China's prefectural cities. They do not detail the effects on real income inequality or urbanization rates.¹⁰ In this paper we consider the effect of the NEN and the Hukou system on *i*) real income inequality between and within (urban vs. rural areas) prefectures, *ii*) evaluate how they affect the spatial distribution of people across China's prefectures, and *iii*) assess their impact on the urbanization rates in different parts of China.¹¹

⁸ Modelling and estimating people's migration dynamics in this way improves upon Whalley and Zhang (2007) and Bosker *et al.* (2012). Whalley and Zhang model people as moving in response to regional (provincial) wage differences only, whilst Bosker *et al.* incorporate migration dynamics based on evidence relating to interprovincial migration flows and ad hoc assumptions on migration costs.

⁹ Poncet (2006) also relies on migrant flows at the provincial level, focusing on rural to urban migration only. This hides a lot of variation at the inter-prefectural and intra-prefectural level. The bulk of migration in China is within provinces. Other studies (notably Rozelle *et al.*, 1999; Zhao, 1999a,b and more recently Zhang and Zhao, 2013 or Giuletti *et al.*, 2014) provide evidence on the determinants of people's migration by relying on surveys in particular areas of China. Although very interesting, and despite the fact that these latter two studies use surveys covering the 15 cities that are the largest migration destinations, it is not obvious that their results generalize to other parts of China.

¹⁰ An exception is Whalley and Zhang (2007). They do consider the effect of the Hukou restrictions on income inequality, as well as on the spatial distribution of people across China. However the most detailed spatial level at which they conduct their analysis is that of China's 31 provinces. These provinces are still very large. As a consequence, their analysis hides a lot of variation at a more detailed geographical level.

¹¹ In both Faber (2014) and Roberts *et al.* (2012), the NEN implicitly has no impact on the rate of urbanization because of the assumed absence of migration.

Our main findings show substantial overall economic benefits of the construction of the NEN and of abolishing the Hukou restrictions. The spatial impact of these two policies is however very different. We find that the construction of the NEN has only reinforced existing urbanization patterns. The initially largest and most urbanized prefectures have gained most from its construction. Although not necessarily the initially richest places, these cities saw the largest improvements in their connectivity due to the construction of the NEN. The initially lagging regions not connected to the NEN have not benefitted much from its construction. Hukou reform would instead benefit the initially richest prefectures along China's South-Eastern coastline. These cities are initially not necessarily the most urbanized nor those with the largest urban populations, but their economic success will attract migrants from across the country when the Hukou restrictions are relaxed. However, many places in China's interior also benefit more from Hukou reform than from the construction of the NEN, be it mostly because of rural out-migration. By allowing everyone to gain by moving to where he/she is most productive, the benefits of easing the Hukou restriction are more widespread.

The structure of the remainder of the paper is as follows. Section 2 briefly reviews China's two main spatial development policies— transport investments and migration restrictions. We present our modeling strategy in Section 3. Section 4 presents the data, estimation and modelling strategy, and Section 5 discusses the estimated impacts of highway construction and migration restrictions on both overall and spatial economic outcomes. Section 6 concludes.

2. A brief history of the NEN and the Hukou system¹²

China's national expressway network, also known as the National Trunk Highway System, was conceived in 1988 with the goal of establishing seven highways radiating from Beijing, nine North-South connections, and 18 East-West connections, giving it the unofficial name "7918 network". The first phase of the network—connecting all cities with population above 200,000 people—was completed by 2007 with a length of about 40,000 km (see Roberts *et al.*, 2012; World Bank, 2007). Since then, the network has further expanded, reaching more than 96,000 km in 2012 (China Statistical Yearbook, 2013)—larger than the U.S. Interstate Highway System. In this paper, we restrict analysis to the impact of the first phase of construction, which was largely concentrated between 1997 and 2007. The main reason for this is that in 2008 China introduced the first high speed railway lines, and today China has the largest high speed rail network in the world. Although these two networks have different user profiles (more goods transported on highways, more business travel on high speed rail), separating the impact

¹² We keep the discussion in this section brief given that both of these policies have been described in much greater detail elsewhere in the literature, see in particular the references cited in this section.

of the NEN from those of railway lines would be difficult (especially since there is a substantial amount of overlap in the cities they connect).

While large scale transport infrastructure investments have been a fairly recent feature of China's spatial policies, China's population registration system, or Hukou, was introduced well over 50 years ago, primarily as a way to control population movements and the allocation of labor to state-controlled production (Chan and Buckingham, 2008; Chan, 2009; Bosker *et al.*, 2012). The main distinction was between agricultural (rural) versus non-agricultural (urban) residence status. It historically considered the rural population as self-sufficient, while providing food rations, housing and educational and health services to the urban population. This general distinction persisted even as urban dwellers became far better off during economic liberalization.

After policy changes in the 2000s, preventing population mobility is no longer the dominant motivation for maintaining the Hukou system. The reason for its persistence, despite frequent expectations that it will be abolished since at least the mid-1990s, is a concern that giving migrants equal rights will exceed the fiscal capacity of cities to provide public services and welfare benefits to everyone. This problem is made more severe because China lacks a mechanism to transfer fiscal resources from rural to urban areas in proportion to a changing population distribution (World Bank and DRC 2014). Further reforms of the Hukou system are currently debated (State Council 2014). These reforms would essentially abolish the Hukou system for smaller cities below 1 million population. Registration in larger cities would be increasingly more difficult, but the reforms would give cities flexibility, including using "point systems" to determine eligibility based on job qualifications, length of residence, and so on. The reforms would also strengthen the rights—including land use rights—of rural residents, ensure more equitable service delivery across rural and urban areas, and improve public finance transfer mechanisms between regions of outmigration and those whose population is growing.

Despite their restrictive nature, China's registration policies have not prevented large scale migration to urban areas since the beginning of the reform era. Between 1995 and 2000, for example, an estimated 50 million people moved from rural to urban areas (Chan, 2013). In 2013, China's National Bureau of Statistics estimated the total non-Hukou population (people living in a place without having a Hukou for that location) at 289 million (NBS, 2014). Nevertheless, by reducing or preventing access to benefits in the destination cities, and by limiting the portability of accrued benefits and monetization of assets in the rural areas, migration under the Hukou system has likely still been considerably lower than it would otherwise have been.

3. The model

Our model introduces labor mobility into the NEG model developed in Roberts et al. (2012). We explicitly allow for people to migrate between prefectures as well as between the urban and rural part

of each prefecture. To do this, we base ourselves on Tabuchi and Thisse (2007) and Behrens *et al.* (2014), and assume that each person j 's likelihood to choose to live in location i (the urban or rural part of any of the 331 Chinese prefectures in our sample), is based on the utility he or she derives from living in that place. This utility depends linearly on real wages, W_i , earned in location i , location i 's amenities, A_i , and an individual-specific idiosyncratic preference for living in location i , ε_{ij} .¹³

$$U_{ij} = W_i + A_i + \varepsilon_{ij} \quad (1)$$

People choose to live in the location which provides them with the highest utility¹⁴, so that the probability that individual j chooses to live in location i is:

$$P(U_{ij} > \max_{k \neq i} U_{kj}) \quad (2)$$

Assuming that the ε_{ij} are drawn from a double exponential function with mean $\pi^2 \mu^2 / 6$, this probability can be written in the following logit form (see McFadden, 1974):

$$P(U_{ij} > \max_{k \neq i} U_{kj}) = \exp((W_i + A_i)/\mu) / \sum_k [\exp((W_k + A_k)/\mu)] \quad (3)$$

Here, μ determines the importance of a person's idiosyncratic preferences in determining his/her location choice. If μ is very small, people basically choose their location only based on $W_i + A_i$. That is, everybody chooses that location offering them the best combination of real wages and amenities. In this scenario, multiple cities can only exist if possible real wage differences between locations are perfectly offset by differences in amenities. By contrast if μ is very large, people basically choose each location with equal probability $1/K$, where K denotes the number of possible locations. In this case people's idiosyncratic preferences for each location are very heterogeneous and, as a result, real wages and amenities do not matter for each individual's location choice.

Spatial equilibrium in the model is reached when the probability in (3) corresponds to the actual observed share of people living in location i :

$$L_i / (\sum_k L_k) = \exp((W_i + A_i)/\mu) / \sum_k [\exp((W_k + A_k)/\mu)] \quad (4)$$

¹³ The idiosyncratic preferences could also reflect differences between people in the cost of migration cost, in the value they put on living close to their friends or relatives, or in the information they have about the potential benefits/cost of migration.

¹⁴ Many NEG-models (notably Krugman's original 1991 contribution) assume that people only move in response to real wage differences. This would mean that in a spatial equilibrium real wages are equalized, both across the urban and rural sectors, as well as across all prefectures in China. Introducing other motives for migration (here amenities and other unobserved location-specific preferences) means that real wage differences between locations can persist in equilibrium. Moreover, the location-specific idiosyncratic preferences ensure that, in equilibrium, at least some people live in each and every location. Note that in Desmet and Rossi-Hansberg (2013) people are also perfectly mobile and decide where to live based on more than just real wages. However, they do not consider individual specific heterogeneous preferences for living in each city. As a result, some cities disappear in their counterfactual scenarios (the same holds for Bosker *et al.* (2012), where in the equilibrium with labor mobility only a few cities remain in existence).

In our model we take each location's amenities A_i as exogenously given. Real wages in each and every location are instead endogenously determined in our model.

3.1 Determining real wages – the role of transportation infrastructure

Real wages in each and every location are determined in the exact same way as in Roberts *et al.* (2012). Here we briefly set out the main features of their model, referring to their original paper for the full details. We focus in particular on the way that transport costs determine real wages in the urban and rural sectors of each prefecture. By influencing market access they are the main conduit through which the NEN affects spatial economic outcomes in our model.

The model in Roberts *et al.* (2012) is an elaborated n -region version of the original NEG model of Krugman (1991a, 1991b). Each location consists of an urban and a rural part. The urban part of each prefecture houses the urban sector which is characterized by internal economies of scale and monopolistic competition. Its rural part is instead home to a constant returns rural sector in which perfect competition prevails. Perfectly tying each of the two sectors in Roberts *et al.* (2012) to the rural and urban part of each prefecture allows us to explicitly model migration between the urban and rural part of each prefecture. In effect, in our model, the decision to move from the rural part of a prefecture to its urban part necessarily involves a change of sector.¹⁵

Both sectors face “love of variety” preferences and (iceberg) transport costs to the rural sector, whereas, in the original Krugman model, these are confined to just the urban sector. It implies that the construction of the NEN also has an impact on the prices and wages in the rural sector. Finally, following Südekum (2005), the model allows for variations in labor efficiency both across regions and between the urban and rural sectors.

As set out in Appendix A, equilibrium in the model is characterized, for each prefecture, by a system of five simultaneous non-linear equations. For each prefecture, these equations determine wages and the price indices in both its urban and rural part (sector). Importantly, these depend on two factors – (i) the region's (exogenous) level of labor efficiency in its urban/rural sector, and (ii) the region's level of real market access (RMA) in that sector. The RMA in each sector provides the main channel through which transport costs – and, hence, the construction of the NEN – affects the overall spatial equilibrium. In particular, a decline in transport costs associated with e.g. the construction of a new highway network link between two prefectures has two opposing effects on wages working through RMA. To see this, take the example of the urban sector. A reduction in the costs for urban firms in prefecture i of

¹⁵ We think this is a realistic assumption because migrants in urban areas of China predominantly work in low-skilled manufacturing or service sector jobs. See also World Bank (2013).

transporting their output to another prefecture j , increases demand for prefecture i 's output. This puts upward pressure on urban wages. However, the reduction in transport costs also exposes urban firms in region i to greater competition from urban firms located in prefecture j . This resulting increase in competition results in countervailing downward pressure on real wages. The overall impact on the urban wage depends on which of these opposing forces dominates.¹⁶

3.2 Spatial equilibrium – with and without the Hukou restrictions

A spatial equilibrium in our model is defined by (4) and [A1]-[A5] in Appendix A¹⁷. The spatial equilibrium that these six equations tie down is however a spatial equilibrium that assumes perfect interregional labor mobility. Everyone freely chooses his/her location. In other words, this represents a “no Hukou equilibrium”: an equilibrium where every Chinese citizen can choose where to work and live without facing any restrictions.

However, we know that the Hukou restrictions do still pose significant barriers for many people, keeping them from migrating to their preferred place of living. This makes it unlikely that (4) holds in reality. This is especially important for the papers simulating the impact of the NEN in a China where the Hukou restrictions are in place (Roberts et al, 2012; and Faber, 2014).¹⁸ These papers basically deal with this issue by assuming that the observed distribution of people and economic activity can be proxied by a spatial equilibrium without any interregional labor mobility, i.e. an extreme Hukou scenario. In the context of our model, this means that equilibrium is defined by equations [A1]-[A5] only (as in Roberts et al., 2012).

Here we use a different strategy to adapt the model to a situation with restricted, but not prohibitively restricted, labor mobility. Given that many people in China did choose to migrate despite the Hukou restrictions, assuming that no one responds with his/her feet to changes in economic opportunities across locations brought about by the construction of the NEN, might result in over- or underestimating the impact of the NEN on regional economic outcomes. To take restricted labor mobility more seriously, we use the information that we have on each location's total migrant and non-migrant population, where $L_i = L_i^{migrant} + L_i^{non-migrant}$. In particular, we assume that, under the current Hukou system, only

¹⁶ Also, the impact of a decline in transport costs in the urban sector tends to positively spill over to the rural sector because these sectors are linked through income. However, this positive relationship can break down when the construction of a network link reduces both urban and rural transport costs (as will be the case in our NEN-scenarios). This, in particular, can occur when there is a large degree of asymmetry between prefectures i and j in the sense that one is heavily specialized in the urban sector and the other in the rural sector. It generates the possibility that a prefecture's urban and rural wages may move in opposite directions in response to the construction of a new network link.

¹⁷ Equations [A1]-[A5] correspond to equations [3] – [7] in Roberts et al. (2012).

¹⁸ For the papers considering the impact of the Hukou system (Bosker et al., 2012; Whalley and Zhang, 2007), this is not so much an issue. They simply compare today's spatial distribution of people or economic activity to the simulated distribution under completely unrestricted labor mobility. This means that they do not have to simulate any model outcomes under a scenario of restricted labor mobility.

those people that have actually been observed to migrate with the system in place, are potentially willing to migrate (i.e. the “revealed Hukou migrants” in our dataset, $\sum_k L_k^{migrant} = L_{China}^{migrant}$). These revealed Hukou migrants are assumed to choose their location according to (1) – (3). For the non-migrants in our data, we instead assume that they will never migrate as long as the current Hukou restrictions remain in place. This means that a spatial equilibrium with the current Hukou restrictions in place is formally characterized in our model by equations [A1] – [A5] in Appendix A in combination with:

$$\begin{aligned}
L_i / (\sum_k L_k) &= \left(\frac{L_i^{migrant}}{L_{China}^{migrant}} \right) \left(\frac{L_{China}^{migrant}}{L_{China}} \right) + \left(\frac{L_i^{non-migrant}}{L_{China}} \right) \\
&= P(U_{ij} > \max_{k \neq i} U_{kj}) \left(\frac{L_{China}^{migrant}}{L_{China}} \right) + \left(\frac{L_i^{non-migrant}}{L_{China}} \right) \\
&= [\exp((W_i + A_i)/\mu) / \sum_k [\exp((W_k + A_k)/\mu)]] \left(\frac{L_{China}^{migrant}}{L_{China}} \right) + \left(\frac{L_i^{non-migrant}}{L_{China}} \right)
\end{aligned} \tag{5}$$

(5) clearly shows that when the construction of the NEN results in changes in real wages, W_i , across prefectures, this will trigger a restricted migration response by those people that have been shown to be willing to migrate despite the Hukou restrictions. The rest of the Chinese population stays where they are. Rewriting (5) shows that it is basically an equilibrium condition on the (relative) size of a location’s migrant population only:

$$\left(\frac{L_i^{migrant}}{L_{China}^{migrant}} \right) = [\exp((W_i + A_i)/\mu) / \sum_k [\exp((W_k + A_k)/\mu)]] \tag{5b}$$

4. Estimating the main model parameters

Having set out our model, detailing how we allow for different types labor mobility, we now turn to how we estimate the main model parameters. These estimates are the crucial inputs into the counterfactual exercises that we perform in the next section to assess the spatial impacts of the NEN and/or the Hukou restrictions on labor mobility. We briefly discuss our data and review the estimates that we take from Roberts *et al.* (2012). The main focus in this section, however, is on how we estimate the importance of real wages and (natural) amenities in people’s migration decision.

4.1 Data

We build on Roberts *et al.* (2012, Appendix A) for data on the NEN and on each prefecture’s urban and rural wages, income, investment per worker, human capital (years of education), and land area. This data is available for 331 prefectural cities (three prefectural cities were excluded due to missing data)

and similar administrative units.¹⁹ Travel times as a proxy for trade costs between each pair of provinces are derived from two detailed GIS road network data sets—one without and one with the NEN—using a standard shortest-path algorithm. Travel times in the “before” network (i.e. without the NEN in place) will always be no larger (usually substantially lower) than in the “after” (i.e. with the NEN in place) network, because highways allow greater speeds and more direct connections. One simplification is that we assume that travel times within a prefecture are zero ($t_{ii} = 0$).²⁰

The novel data that we use in this paper concern information on migration and some of its alleged determinants. Chinese migration statistics are notoriously difficult to interpret (Chan, 2013). This is especially the case for bilateral migration flows between locations. A virtue of our structural modeling approach is that we can infer the importance of different migration determinants by looking at migration stock only. This information is available²¹. Our data on the stock of migrants comes from a comprehensive data set for 2000 and 2010 derived from official Chinese census publications and other statistical data sets such as the provincial statistical yearbooks (Chreod Ltd., 2013). The data set includes information on “total population of migrants from the same county”, “total population of migrants from other counties in the same province” and “total population of migrants from other provinces”, each for 2000 and 2010. It also provides data on “total population with household registration”, i.e. with a local Hukou, for both years.

We complement it with data on geographic characteristics/amenities possibly influencing people’s migration decisions. Information on terrain ruggedness comes from Nunn and Puga (2012). Data on cooling and heating degree days are from NASA (2009). Annual cooling degree days measure the total number of degrees by which daily temperatures exceeded 18°C. It is often used as a measure of the need for air conditioning. Similarly, heating degree days indicate the accumulation of degrees when the daily mean temperature is below 18°. Information on amenities (the share of households with access to tap water, toilets and natural gas supply) come from the same data set as our migration stocks.²² Finally, we have collected the dominant language for each prefectural region from a GIS data set of linguistic regions called the World Language Mapping System (Ethnologue, 2004).

¹⁹ The 331 prefectures covers 280 of China’s 283 prefectural cities, 11 of its 17 prefectures, all 30 of its ethnic minority autonomous prefectures, and all three of its leagues. Also included are the municipalities of Beijing, Shanghai and Tianjin. Three areas that together comprise Chongqing are also included in the sample – the “One Hour Circle”, the Northeast wing and Southeast wing (see Roberts *et al*, p 590, for further details).

²⁰ This means that we abstract from any difference in the quality of infrastructure within prefectures. See Baum-Snow and Turner (2013) and Baum-Snow *et al.* (2013) for detailed studies on inter-prefectural infrastructure.

²¹ Its quality can still be debated. The official statistics that we use are likely to underreport actual migration stocks in some cities given the many unofficial migrants.

²² We use the 2010 figures in most of our estimations. Results are very similar when using estimates for 2007 instead (the same year for which we have our NEN-travel time-information), obtained by linearly interpolating the figures from 2000 and 2010.

4.2 Estimation

Since our only adaptation of the Roberts *et al.* (2012) model is the introduction of labor mobility, the only parameters that we need to assign values to here are those determining the importance of real wages and (natural) amenities in people's migration decisions. We discuss the estimation of these parameters below. Appendix B sets-out the estimates of all the other main model parameters that we adopt directly from Roberts *et al.* (2012).

4.2.1 The determinants of migration in China

To identify the relative importance of real wages and amenities in people's migration decisions, we follow Behrens *et al.* (2013) and make use of the spatial equilibrium condition under the current Hukou system (5b). The strong assumption underlying our identification is that China's spatial economy is in equilibrium in 2007.²³ Under this assumption, we can first obtain real wages in each prefecture's urban and rural part in the exact same way as calibrated in Roberts *et al.* (2012). This step relies solely on parameters already identified in that paper (see also Appendix B). Next, we plug these calibrated real wages, W_i , into equilibrium condition (5b). This allows us to back out the indirect utility levels that correspond to the 2007 Hukou spatial equilibrium: First normalize $W_1 + A_1 \equiv 0$ (this basically means that we measure utility *relative* to that in a baseline location, which we take to be the urban part of Shijiazhuang prefecture²⁴). Next divide each location i 's migrant share in total Chinese population by that in the baseline location, use the normalization $W_1 + A_1 \equiv 0$, and take logs on both sides of the equation. From this it immediately follows that each location's log migrant population relative to that in the baseline location is directly related to the indirect utility derived from the real wages and amenity levels in that location:

$$\ln\left(\frac{L_i^{migrant}}{L_1^{migrant}}\right) = (W_i + A_i)/\mu \quad (6)$$

Now, assume each location's amenities consist partly of observed amenities (i.e. those that we have data on), A_i^{obs} , and partly of unobserved amenities, A_i^{unobs} . We can then estimate the importance of real wages and observed amenities in driving people's migration decision by running the following regression of each location's log migrant population relative to that in the baseline location on its calibrated real wage W_i and its observed amenities, A_i^{obs} :

²³ In our defense, the exact same strong assumption is made in all other studies using calibration exercises to gauge the impacts of either the NEN or Hukou restrictions (e.g. Bosker *et al.*, 2012; Whalley and Zhang, 2007; Roberts *et al.*, 2012; Faber, 2014; and Desmet and Rossi-Hansberg, 2013).

²⁴ The choice of reference location is arbitrary, and does not influence any of the results.

$$\ln\left(\frac{L_i^{migrant}}{L_1^{migrant}}\right) = \beta_0 + \beta_1 W_i + \beta_2 A_i^{obs} + \varepsilon_i \quad (7)$$

where β_0 captures the reference location's share in China's migrant population. Moreover, the residuals can be interpreted as the unobserved part of each location's amenities, $A_i^{unobs} = \hat{\varepsilon}_i$.²⁵

We use the above outlined strategy to estimate the relative importance of real wages, various geographical amenities, and different public amenities in determining individual location choices. The geographical characteristics of each location that we are able to include in our analysis are its ruggedness, temperature (number of heating and cooling days), rainfall, a dummy variable for location on the Yangtze River, and availability of natural resources. The public amenities we are able to include focus on the provision of tap water, sewage, and natural gas to households. Tables 1a and 1b below show the results of estimating different versions of equation (7). Table A1 in Appendix C provides some additional extensions to these results that show that people are more sensitive to real wage and/or amenity differences in nearby locations (i.e. within the same province).

We need to mention three details before discussing our findings. First, all regressions control for unobserved determinants of migration that do not vary between the rural parts of prefectures in the same province as well as unobserved factors that do not vary between the urban parts of prefectures in the same province (i.e. we include province-rural and province-urban fixed effects). In addition, each regression also includes longitude and latitude and a location's area as controls. Including area as a control takes account of the fact that the same population on less land means a larger population density, and thus more congestion, whereas including longitude and latitude aims to control for unobservables that are related to absolute (and arguably also relative) location. Finally, we include dummies for four of the major languages spoken in China, to (roughly) capture possible language barriers to internal migration.²⁶

Second, in almost all specifications we allow the geographical amenities to differentially affect migration decisions in urban and rural areas. We do this as most of our geographical data reports the same value for each prefecture's urban and rural part.

Third, and most importantly, by regressing our real wage and amenity data on log *migrant* population relative to a baseline location, we effectively identify people's willingness to migrate in response to real wage and/or amenity differences from the variation in location choices of people who did decide to migrate. This implies that in our counterfactual "no Hukou" scenario we make the implicit assumption

²⁵ Using the residuals as a measure of each location's unobserved amenities does assume: (i) that our linear utility model is well-specified (i.e. with real wages, observed and unobserved amenities entering utility in an additively separable way), and (ii) that these unobserved amenities are uncorrelated with both a location's real wages as well as its observed amenities.

²⁶ Full results are available upon request.

that current migrants' preferences are representative of those of the entire Chinese population.²⁷ An alternative could be to estimate (7) using log total population as the dependent variable. We do show these results as a robustness check, but prefer our estimates using log migrant population. In fact, a complication that arises when estimating (7) using total population shares, is that the residual does not only capture unobserved amenity differences, but also the Hukou restrictions' capability to keep people from moving out of each location.²⁸

Table 1a builds up to the baseline results that we use to simulate our counterfactual scenarios in the next section. Column 1 shows results when only including real wages in the regression. Subsequently, in columns 2 and 3 we consecutively add our geographic and public amenity variables. Column 4 complements the other columns by showing results when restricting the impact of the geographical variables to be the same in urban and rural areas.

Irrespective of the amenities included, we always find a large and significant positive effect of real wages on the size of a location's migrant population.²⁹ In our preferred specification in column 3, a 1 percent higher real wage corresponds to attracting a 0.75 percent larger migrant population. Of the geographical characteristics we find that only rainfall, heating days and ruggedness are significantly associated with a location's ability to attract migrants. However, with the exception of ruggedness, this is only so in rural areas. Finally, we find that the provision of public services, as captured by the three variables we were able to collect comprehensive data on, is another important determinant of people's migration decision. Especially the availability of sewage facilities, and a direct natural gas connection for heating and cooking are important. Some care is warranted in taking these results regarding public amenity provision too literally, however, as they may suffer from endogeneity issues (induced by either reverse causality or omitted variables). For example, other amenities (that we do not observe) correlated with both our included public amenities and migrant stocks are likely to exist. Overall, however, we take our findings as indicative of the importance of both real wages and public amenity provision in shaping migration patterns in China.

²⁷ This is a strong assumption. Whether or not it is a reasonable assumption is hard to test given the available data. One way to infer this would be to ask non-migrants what factors keep them from migrating, as well as what factors would determine their migration decision in the absence of the current Hukou restrictions. This is beyond the scope of the current paper.

²⁸ See also Desmet and Rossi-Hansberg (2013)

²⁹ This corroborates earlier findings by Poncet (2006), Rozelle et al. (1999) and Zhao (1999).

Table 1a. Determinants of the stock of migrants

	(1)	(2)		(3)		(4)
Dep. Variable:	ln migrants	ln migrants		ln migrants		ln migrants
ln real wage	0.986 [0.00]***	0.923 [0.00]***		0.749 [0.00]***		0.608 [0.008]***
		<u>urban</u>	<u>rural</u>	<u>urban</u>	<u>rural</u>	<u>no split Geo</u>
ln ruggedness	-	-0.137	-0.151	-0.109	-0.138	-0.139
		[0.023]**	[0.001]***	[0.044]**	[0.003]***	[0.003]***
ln cooling days	-	0.336	0.165	0.156	0.022	0.200
		[0.254]	[0.537]	[0.622]	[0.933]	[0.403]
ln heating days	-	0.365	0.392	0.194	0.291	0.172
		[0.056]*	[0.001]***	[0.486]	[0.045]**	[0.394]
ln rainfall	-	-0.117	0.679	-0.072	0.679	0.237
		[0.564]	[0.003]***	[0.735]	[0.004]***	[0.23]
D yangtze	-	0.235	-0.196	0.051	-0.323	-0.135
		[0.104]	[0.273]	[0.768]	[0.12]	[0.489]
nat.res.index	-	-0.014	0.005	-0.009	-0.006	-0.006
		[0.37]	[0.871]	[0.55]	[0.839]	[0.71]
% hh water	-	-		0.117		0.584
		-		[0.702]		[0.098]*
% hh toilet	-	-		1.197		1.401
		-		[0.009]***		[0.004]***
% pop gas	-	-		0.515		0.548
		-		[0.022]**		[0.031]**
nr.obs	662	662		662		662
R2	0.589	0.630		0.670		0.620

Notes: all regressions include province-urban and province-rural fixed effects, as well as controls for a prefecture's $\ln(x\text{-coordinate})$, $\ln(y\text{-coordinate})$, $\ln(\text{area})$ and four dummy variables denoting whether the dominant language spoken in each prefecture corresponds to one of four of China's main languages spoken (Mandarin, Yue, Wu and Jinyu (often considered a dialect of Mandarin)). p -values, based on standard errors clustered at the province level, in brackets. ***, **, * denotes significance at the 1%, 5%, 10% respectively. In all columns we allow the coefficients for all geography-related variables to differ between the urban and rural parts of prefectures respectively (effectively this means that we include each geography variable interacted with our rural dummy as well as interacted with our dummy indicating the urban part of each prefecture).

Table 1b shows results when using different dependent variables instead of migrant stocks when estimating (7). They provide crucial perspective on our choice of identifying the (relative) importance of different migration determinants looking at each location's migrant population only. Column 1 and 2 show results when considering the share of non-migrants or the share in total Chinese population in each location as the dependent variable respectively.³⁰ The most striking difference is the fact that we find a much weaker association between real wages and these shares.

³⁰ We also performed a regression with the share of Hukou-holders in each location as the dependent variable. The results from this regression were almost identical to those when using non-migrant shares. This is not that surprising as both measures are highly correlated (with a correlation coefficient of 0.996). These results are available upon request.

Table 1b. Non-migrants, Hukou holders, and total population (growth)

	1		2		3	
Dep. Variable:	ln non-migrants		ln tot. pop		D tot. Pop 2000-10	
ln real wage	0.249 [0.100]*		0.371 [0.046]**		0.051 [0.022]**	
	<u>Urban</u>	<u>rural</u>	<u>urban</u>	<u>rural</u>	<u>urban</u>	<u>rural</u>
ln ruggedness	-0.122 [0.005]***	-0.174 [0.001]***	-0.128 [0.006]***	-0.166 [0.001]***	-0.012 [0.278]	-0.009 [0.172]
ln cooling days	-0.072 [0.814]	-0.031 [0.911]	0.012 [0.968]	-0.028 [0.919]	0.032 [0.628]	-0.093 [0.266]
ln heating days	0.263 [0.043]**	0.439 [0.001]***	0.271 [0.073]*	0.419 [0.002]***	-0.028 [0.626]	-0.070 [0.003]***
ln rainfall	-0.027 [0.847]	0.677 [0.012]**	-0.011 [0.943]	0.670 [0.012]**	-0.079 [0.178]	-0.017 [0.555]
D yangtze	0.045 [0.763]	-0.194 [0.331]	0.041 [0.788]	-0.247 [0.241]	0.022 [0.52]	0.007 [0.842]
nat.res.index	-0.001 [0.932]	-0.039 [0.062]*	-0.002 [0.87]	-0.036 [0.112]	-0.002 [0.465]	0.009 [0.181]
% hh water	-0.756 [0.012]**		-0.560 [0.066]*		0.104 [0.008]***	
% hh toilet	1.121 [0.023]**		1.259 [0.012]**		0.109 [0.098]*	
% pop gas	0.397 [0.029]**		0.337 [0.081]*		-0.182 [0.004]***	
nr.obs	662		662		662	
R2	0.735		0.709		0.544	

Notes: all regressions include province-urban and province-rural fixed effects, as well as controls for a prefecture's $\ln(x\text{-coordinate})$, $\ln(y\text{-coordinate})$, $\ln(\text{area})$ and four dummy variables denoting whether the dominant language spoken in each prefecture corresponds to one of four of China's main languages spoken (Mandarin, Yue, Wu and Jinyu). p -values, based on standard errors clustered at the province level, in brackets. ***, **, * denotes significance at the 1%, 5%, 10% respectively. In all columns we allow the coefficients for all geography-related variables to differ between the urban and rural parts of prefectures respectively (effectively this means that we include each geography variable interacted with our rural dummy as well as interacted with our dummy indicating the urban part of each prefecture).

Both the significance, as well as the size, of the estimated coefficient on real wages falls substantially. This is very different for the geographical characteristics, where we find similar results as when considering migrant stocks. The same holds for our three public amenity variables, except for a surprising negative effect of the percentage of households that have tap water in their home. This most likely reflects the fact that providing tap water is easier when serving a smaller population. The results when considering total population shares (migrants + non-migrants) in column 2 are more similar to those using non-migrant shares, which is not that surprising given that migrants typically make up only a small part of a prefecture's population. Finally, when considering population growth instead (column 3), we again find a positive effect of real wages. It is not as significant as in our "migrant-regressions",

but one has to keep in mind that population growth is only partly driven by migration.³¹ Combined with our earlier “migrants-results” in Table 1a (as well as the extensions in Table A1), and the survey and provincial evidence in earlier studies (Poncet, 2006; Rozelle et al., 1999; and Zhao, 1999a,b), all our evidence supports the idea that real wages are indeed a very important driver of migration decisions in present day China.³²

In the counterfactual scenarios that allow for labor mobility, we always use the results shown in Table 1a, column 3, as shaping people’s migration decisions. We do not only use the estimated parameters of (7), but also the corresponding estimates of the unobserved amenities in each prefecture’s urban and rural parts (the sum of each location’s residual and its urban-rural/province fixed effect).

5. The spatial impacts of the construction of the NEN and the Hukou system

With all the structural model parameters in hand, we are now in a position to assess the impact of the NEN and/or Hukou restrictions on China’s spatial economy. We do this by simulating our full structural NEG model under three different scenarios. We discuss in detail how we generate the counterfactual spatial equilibria that each of these scenarios rely on. In all three scenarios our analysis focuses on the distribution of four main variables across China’s prefectures: real income per worker, the urban-rural real wage gap, urbanization rates, and overall population.³³ For sake of comparison, we also sometimes report the Roberts *et al.* (2012) results that consider the impact of the NEN under an extreme Hukou scenario (i.e. no labor mobility), and/or compare our “no Hukou”-findings to those found by Bosker *et al.* (2012).

5.1 Simulating the impact of the NEN and/or Hukou system

We observe the current 2007 distribution of people, urbanization and real incomes across the urban

³¹ On average, only about 40% of Chinese city growth is driven by migration. Another 40% is driven by urban expansion into rural areas, leaving 20% as being determined by natural population growth [World Bank and DRC, 2014].

³² From the perspective of our later simulation exercises, the coefficient on real wages in these regressions is most important. In our setup, the (construction of the) NEN affects people’s location decision solely through its effect on real wages in the rural and urban part of each prefecture. We are aware that our main results regarding the importance of real wages in people’s migration decision might still be plagued by reverse causality or omitted variable bias. We lack a convincing source of exogenous variation to deal with this issue. But, in light of all findings in our robustness checks as well as in earlier papers that have shown the importance of wage/income considerations when deciding to migrate or not, we find it very unlikely that addressing these endogeneity concerns will entirely undo the positive effect of real wages that we identify in our regressions. It could result in a smaller positive effect however. Our results in section 5 are qualitatively robust to assigning a smaller role for wages in people’s migration decision.

³³ Our simulations also allow us to look at nominal incomes, price levels, etc. in the urban and rural part of each prefecture. Results are available upon request.

and rural parts of each Chinese prefecture (in total we consider 662 locations of which 331 are urban and 331 rural). We take this as our baseline spatial equilibrium with the NEN in place and with restricted labor mobility because of the Hukou system, i.e. a spatial equilibrium satisfying equations [A1]-[A5] and (5).

Scenario 1: The impact of the NEN under the current Hukou system

In this scenario, we compare the observed 2007 “NEN + Hukou” spatial equilibrium to the counterfactual “no NEN + Hukou” spatial equilibrium. We simulate this counterfactual by changing the travel times between locations back to the situation before the NEN was built.

To find the spatial equilibrium corresponding to this “new” situation, we first numerically solve for the (short run) equilibrium values for w_i^U , w_i^R , G_i^U , and G_i^R , in each prefecture i , using equations (A1) – (A5) in Appendix A1, the “no-NEN” travel time matrix, and the iterative procedure detailed in Roberts *et al.* (2012). In Roberts *et al.* (2012) this would already constitute the counterfactual spatial equilibrium. This is not the case when allowing for (restricted) labor mobility: now the new equilibrium wage and price levels in each location may induce people to change their location according to the migration dynamics we estimated in Section 4. In particular, we use the simulated (short run) equilibrium values for w_i^U , w_i^R , G_i^U , and G_i^R to calculate real wages in the urban and rural part of each prefecture: $\omega_i = w_i^U/P_i$ if i is the urban part of a prefecture and $\omega_i = w_i^R/P_i$ if i is the rural part of the prefecture respectively, where $P_i = (G_i^U)^\theta (G_i^R)^{1-\theta}$ and G^U and G^R are the urban and rural price indices respectively (see Appendix A).

Subsequently we allow these counterfactual real wage changes to trigger a migration response. More specifically, they change each location’s migrant population relative to that in our baseline location (defined earlier) according to equation (7).³⁴

$$\ln(M_i^*/M_1^*) = \widehat{\beta}_0 + \widehat{\beta}_1 \omega_i + \widehat{\beta}_2 A_{i,U}^{obs} + \varepsilon_{i,U}^{obs} \quad (8)$$

In a Hukou scenario with restricted labor mobility, we can then calculate each location’s counterfactual population as follows based on (5):

$$L_i^* = L_i^{non-migrant} + L_{CHINA}^{migrants} \exp\left(\frac{M_i^*}{M_1^*}\right) / \sum_k \exp\left(\frac{M_k^*}{M_1^*}\right) \quad (9)$$

³⁴ Note that this assumes that the level of amenities in a location is independent of its population. In particular, we hold amenities fixed at their 2010 levels. In doing so, we effectively abstract from the fact that a population inflow may put a strain on local governments to keep up the same level of public service provision, or may give rise to congestion.

Next, this new spatial configuration of population across the urban and rural parts of all 331 prefectures changes each location's real market access and thus its real wages. To capture these changes, we numerically resolve for the equilibrium urban and rural wages and price levels based on equations [A1] – [A5] and the new population levels in each location. These new real wages in turn may affect migrants' location choice, and so on. We repeat these steps iteratively until convergence is reached (i.e. real wages, and population shares no longer change between iterations). Upon convergence³⁵, we have obtained our counterfactual “no NEN + Hukou” spatial equilibrium.

Comparing it to the 2007 “NEN + Hukou” situation reveals the general equilibrium impact of the NEN on China's spatial economy making the more realistic assumption of restricted labor mobility instead of the “extreme Hukou” assumption of no labor mobility used in earlier papers (Roberts *et al.*, 2012, or Faber, 2014). By allowing for more realistic migration dynamics, we can assess the possible mitigating effect of labor mobility on the increase in regional inequality, associated with the construction of the NEN found in Roberts *et al.* (2012).

Scenario 2: The impact of abolishing the Hukou system

Our second scenario focusses on the effect of the complete abolishment of the Hukou system on China's spatial development.³⁶ It does so by comparing the observed 2007 “NEN + Hukou” spatial equilibrium to the counterfactual “NEN + no Hukou” spatial equilibrium. We simulate a counterfactual scenario under a complete abolishment of the Hukou system as follows. In the 2007 “NEN + Hukou” spatial equilibrium condition (5) or equivalently (5b) holds. In a spatial equilibrium where the Hukou restrictions are abandoned, everyone is now a potential migrant, and equilibrium is defined instead by (4). Using (8) and our calibrated real wages in 2007, we can easily verify whether (4) holds. If it does not, we update each location's population according to the following equation that now takes all Chinese citizens as potential migrants:

$$L_i = L_{CHINA} \exp\left(\frac{M_i^*}{M_1^*}\right) / \sum_k \exp\left(\frac{M_k^*}{M_1^*}\right) \quad (10)$$

As in Scenario 1, this new spatial configuration of population across the urban and rural parts of all 331 prefectures changes each location's real market access and thus its real wages. To capture these changes, we numerically re-solve for the equilibrium urban and rural wages and price levels based on equations

³⁵ To be precise, we define convergence to be reached when the squared difference between iterations in urban, rural and total population is smaller than 0.00001 for each of these three variables respectively.

³⁶ A complete abolishment of the Hukou system is unlikely to happen in the near future. In principle, we could also look at scenarios involving a partial abolishment of the Hukou system (e.g. allowing free migration to all but the already largest cities, or in particular provinces only). However, such results will depend quite heavily on the specific “partial-Hukou system” considered. Moreover, a complete abolishment allows us to most clearly contrast the effects of infrastructure investment vs. migration restrictions on China's economic geography.

[A1] – [A5] and the new population levels in each location. These new real wages in turn may affect migrants' location choice according to (8) and (10), and so on. We repeat these steps iteratively until convergence is reached (i.e. real wages, and population shares no longer change between iterations). Upon convergence, we have obtained our counterfactual “NEN + no Hukou” spatial equilibrium.

Scenario 2 can be directly compared to Scenario 1 to assess whether, and if so how, China's two main spatial policies differently affect its spatial economy. Also, it can be directly compared to the results shown in Bosker *et al.* (2012) that also consider the spatial effects of the abolition of the Hukou restrictions at the prefecture level, but without explicitly taking the large-scale investments in infrastructure into account. Note that the comparison to Bosker *et al.* (2012) is limited insofar as that paper focusses exclusively on the changes in the spatial distribution of people across prefectures, whereas we also consider the changes in, *inter alia*, real wages and urbanization levels.

Scenario 1b: The impact of the NEN under free labor mobility (i.e. no Hukou restrictions)

Scenarios 1 and 2 are our main scenarios of interest. This third scenario complements Scenario 1 (explaining why we call it Scenario 1b). It provides evidence on whether, and if so how, the construction of the NEN would have had different effects on China's spatial economy had there been no Hukou restrictions on labor mobility. It allows us to further assess the postulated mitigating effect of free labor mobility on the increase in regional inequality associated with the construction of the NEN that was found in Roberts *et al.* (2012).

In this scenario, we compare the counterfactual “NEN + no Hukou”- spatial equilibrium that we already used in Scenario 2 to a counterfactual “no NEN + no Hukou”- spatial equilibrium. We simulate this last counterfactual using the no NEN travel time matrix to proxy transport costs between prefectures when solving equations [A1]-[A5], and combine it with (8) and (10) to iteratively solve for this “no NEN + no Hukou” spatial equilibrium.

5.2 Results – the impact of the NEN and/or the Hukou system on China's spatial economy

In this section, we discuss our main findings. Table 2 shows the changes in our main variables of interest in each of the three scenarios defined in Section 5.1. Table 3 complements these results by showing the correlation between these changes, as well as with the initial levels of urbanization, population, and welfare in each prefecture. Finally, Figures 1 – 4 and Appendix D further illustrate some of the most interesting correlations and provide detailed maps of the spatial distribution of these changes across China's prefectures.

5.2.1 *The impact of the NEN*

We start by considering the impact of the NEN under the current Hukou system (Scenario 1). Table 2 shows that our Scenario 1, that makes the more realistic Hukou assumption of restricted labor mobility, delivers results that are very close to those obtained in earlier papers that completely abstract from labor mobility (e.g. Roberts *et al.*, 2012; but also Faber, 2014). We find a slightly larger overall increase in real income per worker of about 6.6 percent, accompanied by an also slightly larger increase in regional inequality (as measured by the standard deviation of real income per worker across prefectures) of about 9 percent. Also the urban-rural wage gap increases more in Scenario 1 than in the Roberts *et al.*'s (2012) “extreme Hukou” case.

We also observe a slight increase in the agglomeration of people when allowing for restricted labor mobility. However, the population changes that we find are small. China's overall urbanization rate e.g. increases by a mere 0.4 percent point. This small increase does however hide some significant spatial differences. Some of the fastest urbanizing places see an increase in their urbanization rate of more than 4 percentage points. As shown in Figure 1b and the maps in Figure 2, these fastest urbanizing places are the initially already most urbanized places. They are also the cities attracting most new migrants from other cities following the construction of the NEN (see the correlations in Table 3).

In sum, the construction of the NEN has only strengthened the existing agglomeration pattern in China. It has reinforced its urban core at the expense of its rural hinterland, both in economic as well as in population terms.

Of course the fact that we only assume that only “revealed Hukou migrants”, about a fifth of the total Chinese population, move in response to changes in real wages is an important part of the explanation for these relatively small differences between Scenario 1 and the earlier “extreme Hukou” results reported in Roberts *et al.* (2012). However, Scenario 1b shows that even in the hypothetical absence of any Hukou restrictions, we still find almost the same overall welfare effect of the construction of the NEN: an overall national increase of 6 percent in real income per worker. However, the spatial consequences of the construction of the NEN do differ substantially in this “no Hukou” Scenario 1b.

Two things stand out in this respect. First, we find a much smaller increase in regional inter-prefectural income inequality as a result of the construction of the NEN in Scenario 1b. The increase in the standard deviation of real income is only about half that in Scenario 1 (see Table 2). Also, the relationship between initial real income per worker and its change due to the construction of the NEN is much weaker (see Table 3). Figure 1a shows that now many initially wealthy places even observe a fall in real income per worker. And, on top of this, we also see a drop in urban-rural income inequality within the average prefecture, as well as a fall in the number of prefectures with rising urban-rural inequality (only 36 percent compared to more than 50 percent in Scenario 1). A higher degree of labor mobility indeed

appears to mitigate the rise in regional income inequality associated with the construction of the NEN, while delivering a similar overall welfare increase for the average Chinese citizen.

Table 2 Counterfactual results

	Roberts et al. (2012)	Scenario 1	Scenario 1b	Scenario 2
Counterfactual scenario:	NEN vs. no NEN (extreme Hukou)	NEN vs. no NEN Hukou	NEN vs. no NEN no Hukou	Hukou vs. no Hukou NEN
change in aggregate Chinese				
real income (pw) (%)	6.0	6.6	6.0	169.7
sd real income pw (%)	8.6	9.0	4.7	50.2
urbanization (ppt)	-	0.4	0.6	35.4
sd urbanization (%)	-	3.2	2.3	9.0
sd population (%)	-	1.3	2.3	160.9
mean (std dev) change in				
real income pw (%)	4.0 (3.4)	3.9 (3.7)	2.7 (5.3)	23.0 (26.2)
urban/rural wage gap (%)	0.4 (7.1)	1.3 (6.3)	-1.1 (6.5)	11.4 (35.3)
total population (%)	-	-0.3 (1.75)	-1.7 (2.8)	3.7 (143.6)
urbanization (ppt)	-	0.1 (0.7)	-0.3 (1.0)	21.8 (15.9)
% prefectures with increasing				
real income pw	96.1	97.3	79.8	98.5
urban/rural wage gap	44.7	52.9	35.6	52.3
population	-	15.4	18.7	32.9
urbanization rate	-	40.2	35.6	89.1

Second, also in Scenario 1b, the places whose connectivity is most improved by the construction of the NEN gain most in terms of real income per worker and population (see the correlations in Table 3). With the Hukou restrictions in place these now better connected places also witnessed an increase in intra-prefecture urban-rural disparities, both in terms of population (urbanization) and welfare (urban-rural wage gap). This is very different in Scenario 1b. In the absence of any restrictions on labor mobility this relationship with intra-prefecture inequality disappears. Figure 3 shows this in more detail for the NEN's impact on urbanization rates across China's prefectures.

These differences between Scenarios 1 and 1b can be explained by the fact that without any Hukou restrictions in place, many more people respond with their feet to the changes in real incomes induced by the construction of the NEN. As a result, more people do move to the now better connected, initially large prefectures than in the Hukou scenario (aptly illustrated by the maps in Figure 2c, and the higher correlation between the change in real income per worker and the change in population in Scenario 1b in Table 3). It results in more spatial inequality in terms of population (the standard deviation in prefecture population increases almost twice as much as in Scenario 1b), reinforcing the existing agglomeration pattern in China.

Figure 3. % reduction in travel time because of the NEN and ppt change in urbanization

Yet, it is also exactly these migration patterns that mitigate the rise in both intra-prefecture and inter-prefecture income inequality associated with the construction of the NEN that we saw in the Hukou Scenario 1. On the one hand, the bigger inflow of people into the initially largest places mitigates the increase in income per worker in these places as a result of fiercer competition for jobs. This explains the much lower association between initial real income per worker and both growth in real income per worker as well as in urban-rural inequality in Scenario 1b (see Table 3). On the other hand, regional inequality is also reduced because especially those people living in prefectures exhibiting the highest levels of urban-rural inequality are the ones moving out of their prefecture in search of the higher real incomes offered elsewhere. This outflow of mostly rural inhabitants in prefectures losing population, results in a relative increase of rural wages that mitigates the urban-rural wage gap in these places.

Table 3: Correlation matrix outcomes different scenarios

<i>Correlations</i>	Scenario 1 NEN vs. no NEN (Hukou)				Scenario 1b NEN vs. no NEN (no Hukou)				Scenario 2 no Hukou vs. Hukou (NEN)			
change in:	[i]	[ii]	[iii]	[iv]	[i]	[ii]	[iii]	[iv]	[i]	[ii]	[iii]	[iv]
travel time NEN (%)	0.54	0.22	0.30	0.26	0.31	-0.01	0.42	-0.11	-0.07	0.02	0.12	-0.04
[i] real income pw (%)	-	-	-	-	-	-	-	-	-	-	-	-
[ii] urban/rural wage gap (%)	0.70	-	-	-	0.66	-	-	-	0.50	-	-	-
[iii] total population (%)	0.61	0.40	-	-	0.80	0.55	-	-	0.69	0.42	-	-
[iv] urbanization (ppt)	0.73	0.76	0.30	-	0.45	0.91	0.30	-	-0.04	-0.68	-0.27	-
initial:												
ln real income pw	0.49	0.39	0.21	0.44	0.27	0.13	0.19	0.01	0.19	0.17	0.44	-0.06
ln real income	0.30	0.10	0.29	0.20	0.30	0.23	0.31	0.07	-0.13	-0.17	0.10	0.08
urban/rural wage gap	-0.54	-0.53	-0.14	-0.48	-0.54	-0.31	-0.44	-0.18	-0.10	-0.42	-0.37	0.25
ln population	0.06	-0.15	0.24	-0.05	0.29	0.26	0.34	0.10	-0.32	-0.36	-0.19	0.16
urbanization rate	0.74	0.80	0.31	0.76	0.66	0.58	0.50	0.42	0.35	0.49	0.57	-0.35

Notes: bold correlations are NOT significant at the 5% level.

Figure 1. Convergence or divergence? Real income per worker and urbanization

a. Real income per worker

b. Urbanization

Figure 2. Spatial impacts in each of the three scenarios

a. Percent change in real income per worker

b. Percent change in urbanization

Figure 2, continued

Summing up, in both Scenarios 1 and 1b, the construction of the NEN results in a similar 6 percent increase in Chinese real income per worker. This average welfare gain is however associated with a substantial rise in regional inequality. The existing urban and economic core regions of China gain most, whereas the initially poorest, rural prefectures tend to lose population as well as economic activity. A higher degree of labor mobility does mitigate this rise in spatial inequality associated with the construction of the NEN. But, even if every Chinese citizen had been free to migrate to the place of his/her preference, the large-scale, spatially targeted, investments, in the NEN would still have resulted in more, not less, spatial inequality.³⁷ This is an important finding, as the main reason to construct the NEN was the exact opposite: mitigating the existing regional inequality in Chinese economic development.³⁸

5.2.2 *The impact of removing the Hukou restrictions*

Now, we turn our attention to the aggregate and spatial development impacts of China's other main spatial development policy: the Hukou system. In Scenario 2, we take the construction of the NEN as given and simulate the (spatial) impact of a complete abandonment of the Hukou system.³⁹

As can be seen in Table 2, allowing every Chinese citizen to freely choose his/her preferred location has a much bigger impact than the construction of the NEN.⁴⁰ The average Chinese worker experiences an increase in real income of 170 percent. However, this much larger increase in real income also comes with a much more unequal spatial distribution of real income and people across prefectures. Despite the fact that real income per worker goes up in basically all prefectures, real income inequality between prefectures rises as a few already rich prefectures experience much faster growth in real income per worker (Figure 2a shows that these are mainly located in China's coastal regions). Some initially poor regions, in e.g. China's South West also benefit a lot (especially when compared to the impact of the

³⁷ The prefectures with the largest gains in accessibility are typically also the most urbanized. This is not surprising given that the aim of the NEN was to connect all cities with 200,000 or more inhabitants. A 1 percent higher "pre-NEN" urbanization rate is associated with a 6.2 percent larger reduction in average travel time as a result of the construction of the NEN. This bias in NEN-investments towards the already existing urban cores is an important explanation for why we do not find an inequality decreasing effect of its construction.

³⁸ Of course our findings are based on a stylized economic model that e.g. does not consider the construction costs of the NEN. Nor does our model take into account that different industrial sectors may respond differently to a change in transportation costs. Moreover, we do not explicitly take congestion costs into account apart from the fact that competitive pressure on wages is higher in larger places.

³⁹ Results are very similar when simulating the impact of the Hukou restrictions in a China without the NEN. We focus on the abandonment of the Hukou system taking the construction of the NEN for granted as this is the most policy relevant scenario given that the NEN is already constructed while the Hukou system is still in place.

⁴⁰ Do note the important caveat made at the end of this section that puts crucial perspective on the often much bigger effects that we find in Scenario 2.

construction of the NEN, see Figures 1a and 2a), but given that they start from an initially much lower base, the absolute change in their incomes is still substantially lower than that in rich (coastal) China. Besides this rise in inter-prefectural real income inequality, abandoning the Hukou restrictions also results in a much larger rise in the urban-rural wage gap of the average prefecture. This average does however hide substantial variation between prefectures, the urban-wage gap actually falls in the same number of prefectures as in Scenario 1. Interestingly (see Table 3), urban-rural inequality increases the least in the fastest urbanizing prefectures. This is very different from Scenarios 1 and 1b, where rising urbanization and rising urban-rural inequality went hand in hand. It can be explained by the fact that the fastest urbanizing places are also often places losing population (see below for more on this). This outflow of mostly rural inhabitants results in a relative increase in rural wages, mitigating the rise in urban-rural inequality in these places.

Figure 4. The spatial distribution of people and urbanization in the absence of the Hukou system

When abandoning the Hukou system, the spatial distribution of people across Chinese prefectures also becomes much more unequal. As Figure 2c shows, inland China sees a large outflow of people towards the coast (and some places in the North(-East)). As a result our simulations predict that a China without Hukou restrictions would be a China where most people live in the prefectures along its (southern) coastline (see Figure 4). The main remaining large inland population centres, are the currently already very large inland cities (notably Chengdu, Chongqing and Wuhan).⁴¹ Interestingly, this much more

⁴¹ Bosker *et al.* (2012) found an even stronger agglomeration pattern emerging with the abandonment of the Hukou system, with only 52 cities surviving in their baseline simulations. Moreover, their agglomeration pattern is quite different from the one we find. Instead of our “coastal + initially large inland cities” core, their core consists of many more inland cities as well as several place in China’s North-East. These differences can be attributed to four important differences between our and their counterfactual simulations. First, Bosker *et al.* (2012) abstract from the quality and quantity of infrastructure development in China, approximating travel times between cities by simple geodesic distances. Second, they consider only 264 prefectures, compared to the 332 that we consider (our sample also covers the prefectures in the west and north of China). Third, their simulations rely on ad-hoc migration dynamics that, importantly, do not consider individual specific differences in location preferences into account. The fact that we do this prevents prefectures from being completely abandoned in a no-Hukou scenario (in their baseline scenario more than 200 prefectures “empty out”). Finally, we explicitly consider each prefecture’s

pronounced agglomeration pattern of people across Chinese prefectures, does not, however, also mean a further widening of today's difference in urbanization between inland and coastal China. China's overall urbanization rate would rise by 35 percent (resulting in urbanization levels that are more similar to those found in many developed countries). This urbanization is however not concentrated in the same (few, mostly coastal) prefectures that attract many migrants. The degree of urbanization increases in almost all prefectures, and, interestingly, it is larger in the initially less urbanized places (see Figure 1c), which is very different to what we found in Scenarios 1 and 1b. Despite the fact that an abolishment of the Hukou restrictions would result in a stronger concentration of people in mainly coastal prefectures, it is at the same time accompanied by an "urban catch-up" of inland China (see Figure 2c; or the negative correlation between changes in population size and changes in urbanization rate in Table 3). This urban catch-up is however mainly driven by inland rural workers that, no longer held back by the Hukou restrictions, now move to the booming urban areas on the coast.

The relationship between the changes in the spatial distribution of income and people respectively, also often differs when comparing the impact of the construction of the NEN to the impact of abandoning the Hukou system. When abandoning the Hukou restrictions, the positive relationship between subsequent real income growth and a prefecture's initial urbanization rate is weaker compared to the two NEN-scenarios, and its relationship with initial population (or economic) size even turns negative (The Figure in Appendix D provides additional detail on these associations). This is partly a reflection of the NEN's bias towards improving the connectivity of the already large (urbanized) prefectures. Such a bias is not present when abandoning the Hukou restrictions. This difference is particularly evident for many prefectures in the South West of China (see Figure 2 and also Figure A2 in the Appendix). These places were among the least benefitting places from the construction of the NEN, and as a result saw a decrease in both people and real income per capita. When abandoning the Hukou system, these places still witness a net outflow of people, but are among the fastest growing prefectures in terms of real income per capita.

Another striking difference lies in the disparity in initial prefecture characteristics that are associated with an increase in population and/or urbanization rate. In the case of the construction of the NEN, we mostly find that the initially largest urbanized places are the ones attracting most new migrants. When abandoning the Hukou restrictions, the initially more urbanized prefectures still tend to attract most new migrants, but the association with initial population size turns negative (see Table 3). Instead, the positive association with initial real income per capita strengthens, explaining the concentration of

urban and rural part so that not only inter-prefecture migration considerations determine the spatial equilibrium but also intra-prefecture migration considerations.

people in the currently wealthy coastal prefectures in South-East China that arises when abandoning the Hukou system.

We are aware that all effects we find in this Scenario 2 appear very large (especially so when compared to Scenarios 1 and 1b), but one has to remember that they reflect the (very) long run impact of a relaxation of the Hukou restrictions (see also footnote 36). It would probably take years, if not decades, for these effects to play out. It is also quite important to note that our model abstracts from congestion forces other than rising competitive pressure on wages in the most agglomerated prefectures⁴² (although one could argue that some of these congestion forces are to some extent captured by each location's unobserved amenities, see equation (7), or (8), and the discussion in section 3). We do not explicitly consider pressure on the housing market, pollution issues, nor the often stressed problems with the provision of public services that will most likely emerge with population movements of the scale suggested by our simulations.⁴³

6. Conclusions

China's urban population share is expected to rise to about 75 percent by the middle of this century. By then, more than a billion people will live in Chinese cities. Investments in connective infrastructure and migration policies have the potential to significantly shape the speed and pattern of this massive urbanization process. A better understanding of the impacts of these policies will help to achieve the greatest social and economic benefits in each and every part of China.

This paper compares the spatially differentiated impacts of the construction of China's national expressway network (NEN) with those from relaxing its long-standing Hukou restrictions on migration. To do this we develop a structural new economic geography model that, importantly, can be empirically implemented using readily available data on the urban and rural parts of 331 Chinese prefectures. We fit our model to the data, and simulate counterfactual scenarios that allow us to compare each policy's respective impact on regional economic development and urbanization patterns across China.

⁴² Our model also does not incorporate remittances. All wages earned in a particular location are spent there (either on local products, or products imported from other parts of China or from elsewhere).

⁴³ We could extend our model to incorporate one or more of these issues. But it is important to note that in order to be able to also implement such an extension empirically, requires detailed information on (each of the) congestion forces in each of the locations that we consider; and to have an idea on how they depend on a location's population size. Bosker *et al.* (2012) for example do this by exogenously fixing each location's housing stock. This poses a strong additional congestion force, as rental prices go up when more people migrate to a location. This can be viewed as more or less the polar opposite of our model that implicitly assumes that a location's housing stock perfectly adapts to changes in population (so that these changes leave no effect on the cost of living other than through their effect on the consumer price index). Reality lies in between these two polar cases: cities do respond to population changes by building additional housing units (in fact in some cities in China house prices have fallen despite a growing population as a result of overconstruction).

We find the largest welfare gains for most people when abolishing the Hukou restrictions, because this would enable each individual to move to where he/she is most productive. Continuing to restrict migration with the objective of more balanced regional development therefore has high aggregate costs. Relaxing the Hukou system, which Chinese policy makers are considering, will moreover result in a much more evenly urbanized China, with urbanization levels approaching those observed in most developed countries. On the other hand this rise in income and urbanization will most likely come with a strong further concentration of people and economic activity, very often in China's coastal regions that today already constitute the wealthiest parts of the country.

This is not to say that further investment in infrastructure is not important, but we do find that the construction of the NEN has resulted in a (much) more modest increase in overall Chinese income per worker of about 6.5 percent. So far, these large-scale infrastructure investments tended to favor the already large, urbanized places. And, with the Hukou system in place, many people in lagging regions are prevented from benefitting from the resulting increase in real incomes in these now better-connected places. As a result, the welfare effect of the construction of the NEN has been small, or even negative, in many of the currently already lagging Chinese prefectures. This is quite the opposite of the spread of economic activity that the large scale investment in the NEN meant to achieve.

Our analysis extends only to 2007. After this year it becomes more difficult to isolate the effect of the NEN. Since then, further highway construction and the development of an extensive high speed rail network has linked major cities even more closely together. Future analysis could extend our framework to assess the impact of these investments, also in light of the recent gradual relaxation of the Hukou regulations in several parts of China.

References

- Banerjee, A., Duflo, E., & Qian, N. (2012). *On the road: Access to transportation infrastructure and economic growth in China* (No. w17897). National Bureau of Economic Research.
- Baum-Snow, N., & Turner, M. (2012). Transportation and the decentralization of Chinese cities.
- Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2012). Roads, railways and decentralization of Chinese cities. Mimeo.
- Behrens, K., Mion, G., Murata, Y., & Südekum, J. (2013). Spatial frictions (No. 7175). Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit.
- Bosker, M., Brakman, S., Garretsen, H., & Schramm, M. (2012). Relaxing Hukou: Increased labor mobility and China's economic geography. *Journal of Urban Economics*, 72(2), 252-266.
- Chan, K. W. (2013). China: internal migration, in *The Encyclopedia of Global Human Migration*. Immanuel Ness and Peter Bellwood (eds.), Blackwell Publishing, London.

- Chan, K.W., & Buckingham, W.(2008). Is China abolishing the hukou system? *The China Quarterly*, 195, 582-606
- Cheng, T. & Selden, M. (1994) “The Origins and Social Consequences of China’s Hukou System.” *The China Quarterly* 139:644-668
- Chreod Ltd. (2013), Local area unit statistical data set for 2000 and 2010, Beijing.
- Desmet, K. and E. Rossi-Hansberg (2013) Urban accounting and welfare. *American Economic Review*, 103-6: 2296-2327.
- Ethnologue (2004), World Language Mapping System, SIL International, www.ethnologue.com.
- Faber, B. (2014). Trade Integration, Market Size, And Industrialization: Evidence From China’s National Trunk Highway System. *Review of Economic Studies*, forthcoming.
- Fingleton, B., (2005). Towards applied geographical economics: modelling relative wage rates, incomes and prices for the regions of Great Britain. *Applied Economics*, 37 (21), 2417–2428.
- Fingleton, B., (2007). New economic geography: some preliminaries. In: Fingleton, Bernard (Ed.), *New Directions in Economic Geography*. Edward Elgar, Cheltenham.
- Fujita, M., Venables, A.J., Krugman, P.R. (1999). *The Spatial Economy: Cities, Regions and International Trade*. MIT Press, Cambridge. Mass.
- Giulletti, C., Wahba, J., and Zenou Y. (2014). Strong versus weak ties in migration. IZA discussion paper, no. 8089.
- Krugman, P., (1991a). *Geography and Trade*. MIT Press, Cambridge, Mass.
- Krugman, P. (1991b). Increasing returns and economic geography. *Journal of Political Economy*, 99 (3), 483–499.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (ed.), *Frontiers in Econometrics*. Academic Press, New York, 105–142.
- NASA (2009), Surface meteorology and Solar Energy (SSE) at 1° resolution, U.S. National Aeronautics and Space Administration (NASA), <http://eosweb.larc.nasa.gov/sse/>.
- NBS (2014), Statistical Communique of the People’s Republic of China on the 2013 National Economic and Social Development, National Bureau of Statistics, February 24, 2014.
- Nunn, N., & Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, 94(1), 20-36.
- Poncet, S. (2006). Provincial migration dynamics in China: Borders, costs and economic motivations. *Regional Science and Urban Economics*, 36(3), 385-398.
- Roberts, M., Deichmann, U., Fingleton, B., & Shi, T. (2012). Evaluating China's road to prosperity: A new economic geography approach. *Regional Science and Urban Economics*, 42(4), 580-594.
- Rozelle, S., Taylor, J. E., & DeBrau, A. (1999). Migration, remittances, and agricultural productivity in China. *American Economic Review*, 89(2), 287-91.
- State Council (2014), Opinions of the State Council on further advancing the reform of the household registration (Hukou) system, No. Guofa [2014]25, The State Council of the People's Republic of China, Beijing.
- Südekum, J. (2005). Increasing returns and spatial unemployment disparities. *Papers in Regional Science*, 84 (2), 159–181.
- Tabuchi, T., & Thisse, J. F. (2002). Taste heterogeneity, labor mobility and economic geography. *Journal of Development Economics*, 69(1), 155-177.

- Whalley, J., & Zhang, S. (2007). A numerical simulation analysis of (Hukou) labour mobility restrictions in China. *Journal of Development Economics*, 83(2), 392-410.
- World Bank (2007), Domestic trade impacts of the expansion of the national expressway network in China. EASTR Working Paper No. 8, Transport Sector Unit, Infrastructure Development East Asia and Pacific Region, The World Bank, Washington, DC.
- World Bank (2009), *World Development Report 2009: Reshaping Economic Geography*, Washington, D.C.
- World Bank and DRC (2014), *Urban China – Toward efficient, inclusive, and sustainable urbanization*, World Bank, Washington, D.C., and Development Research Center of the State Council, Beijing.
- Zhang, J., and Zhao, Z. (2013). Measuring the income-distance trade-off for rural-urban migrants in China, IZA discussion paper, no.7160.
- Zhao, Y. (1999a), Labor migration and earnings differences: the case of rural China. *Economic Development and Cultural Change*, Vol.47(4), 767-782.
- Zhao, Y. (1999b), Leaving the countryside: rural-urban migration decisions in China. *American Economic Review*, 89(2): 281-286.
- Zheng, S., & Kahn, M. E. (2013). China's bullet trains facilitate market integration and mitigate the cost of megacity growth. *Proceedings of the National Academy of Sciences*, 110(14), E1248-E1253.

Appendix A

Equilibrium in the Roberts *et al.* (2013) model is described, for each region, by the following set of five simultaneous non-linear equations:

$$w_i^U = \left[\sum_{j=1}^N Y_j (G_j^U)^{\sigma-1} (T_{ij}^U)^{1-\sigma} \right]^{\frac{1}{\sigma}} \quad [A1]$$

$$w_i^R = \left[\sum_{j=1}^N Y_j (G_j^R)^{\eta-1} (T_{ij}^R)^{1-\eta} \right]^{\frac{1}{\eta}} \quad [A2]$$

$$G_i^U = \left[\sum_{j=1}^N \kappa_j^U \lambda_j (w_j^U T_{ij}^U)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad [A3]$$

$$G_i^R = \left[\sum_{j=1}^N \kappa_j^R \phi_j (w_j^R T_{ij}^R)^{1-\eta} \right]^{\frac{1}{1-\eta}} \quad [A4]$$

$$Y_i = \theta \kappa_i^U \lambda_i w_i^U + (1 - \theta) \kappa_i^R \phi_i w_i^R \quad [A5]$$

where $i = 1, \dots, N$ and N is the total number of regions. Meanwhile, w_i and G_i are the nominal wage per efficiency worker and price index respectively with the superscript U (R) denoting the urban (rural sector). T_{ij} is the iceberg transport cost incurred in shipping a unit of output from region i to region j , where these costs are assumed to take the form $T_{ij}^x = 1 + t_{ij}^{\tau_x}$ where $x \in \{U, R\}$, $0 < \tau_U, \tau_R < 1$ are scalar parameters and t_{ij} is the optimal travel time by road between regions i and j . σ and η are the elasticities of substitution in the urban and rural sectors respectively, whilst $\kappa_i^U \lambda_i$ is the number of urban efficiency workers in region i , which is equal to the product of urban labor efficiency κ_i^U and (raw) labor units λ_i . Likewise, $\kappa_i^R \phi_i$ is i 's number of rural efficiency workers, which is equal to the product of rural labor efficiency κ_i^R and rural labor units ϕ_i . Y_i is the income level, which is equal to the weighted sum of the number of urban efficiency workers multiplied by w_i^U , and the number of rural efficiency workers multiplied by w_i^R . The weights θ and $1 - \theta$ are approximated by the respective urban and rural shares of total employment in the Chinese economy, and efficiency levels in both sectors are measured relative to the minimum observed level of urban labor efficiency across all regions. In other words, $\kappa_i^U = E_i^U / E_{\min}^U$ and $\kappa_i^R = E_i^R / E_{\min}^R$, where E_i denotes region i 's absolute level of labor efficiency and $E_{\min}^U = \min(E_1^U, \dots, E_N^U)$. In equations [A1] and [A2], nominal wages per efficiency

worker in region i are determined by the region's level of real market access (RMA) in the urban and rural sectors respectively. However, we are more interested in wages per worker, which are given by $w_i^{*x} = \kappa_i^x w_i^x$. Given this, equations [A1] and [A2] may be re-written as $w_i^{*U} = \kappa_i^U w_i^U = \kappa_i^U (RMA_i^U)^{1/\sigma}$ and $w_i^{*R} = \kappa_i^R (RMA_i^R)^{1/\eta}$ respectively. RMA_i^U is basically equal to a weighted sum of aggregate income levels across all regions, including region i , where the weights are determined by both G_i^U and the cost of transporting urban goods from i to each region. Regions with high levels of w_i^{*U} will be well-connected to other regions with high levels of income and a high urban price index. The interpretation of RMA_i^R is similar except that it relates to the cost of transporting rural goods.

Appendix B

In their paper, Roberts *et al.* (2013) estimate the main parameters of the urban and rural wage equations respectively (i.e. equations [A1] and [A2]). To deal with problems of endogeneity and spatial autocorrelation, they adopt a feasible generalized two-stage least squares (FGS2SLS) approach to the estimation of both equations. This strategy results in point estimates of σ and η – the elasticities of substitution for the urban and rural sectors respectively – of 6.424 and 4.887. The strategy also yields estimates of urban (κ^U) and rural (κ^R) labor efficiency for each region.

For the remaining model parameters in equations [A1] – [A5], Roberts *et al.* specify λ_i , ϕ_i and θ as taking their 2007 observed values. Meanwhile, they arrive at values of $\tau_U = 0.45$ and $\tau_R = 0.75$ for the two key transport function parameters via calibrating their model so as to achieve a good fit between their baseline solution and actual 2007 data subject to the condition of satisfactory regression diagnostics for the two wage equations.

Appendix C

Table A1 below further extends our findings reported in Table 1a and 1b by considering migrants distinguished by the location where they originated from. In particular, column 2 considers only migrants that came from the same province, whereas column 1 focuses on migrants originating from another province than the prefecture where they currently reside.

Interestingly real wages are important drivers of both types of migration, however the same increase in real wages affects inter-provincial migration to a much smaller extent than intra-provincial migration (the effect is more than two times smaller). Also the provision of public amenities has a larger effect on intra-provincial migration flows compared to inter-provincial migration. Both findings provide further

confidence in our baseline results⁴⁴, as it is expected that people are more sensitive (aware) to real wage difference with nearby locations⁴⁵.

Table A1. Ln Migrants, by different origin

Dep. Variable:	(1) other province		(2) same province	
ln real wage	0.449 [0.00]***		0.969 [0.00]***	
	<u>urban</u>	<u>rural</u>	<u>urban</u>	<u>rural</u>
ln ruggedness	-0.108 [0.017]**	-0.152 [0.005]***	-0.070 [0.259]	-0.123 [0.012]**
ln cooling days	0.019 [0.948]	0.055 [0.801]	0.204 [0.601]	0.104 [0.741]
ln heating days	0.221 [0.274]	0.444 [0.019]**	0.120 [0.736]	0.106 [0.403]
ln rainfall	-0.086 [0.577]	0.763 [0.002]***	-0.096 [0.708]	0.677 [0.009]***
D yangtze	0.043 [0.771]	-0.262 [0.077]*	0.080 [0.691]	-0.226 [0.452]
nat.res.index	-0.006 [0.619]	-0.013 [0.573]	-0.010 [0.571]	0.017 [0.604]
% hh water	-0.281 [0.329]		1.227 [0.003]***	
% hh toilet	0.916 [0.027]**		1.777 [0.001]***	
% pop gas	0.614 [0.001]***		0.428 [0.148]	
nr.obs	662		662	
R2	0.673		0.685	

Notes: all regressions include province-urban and province-rural fixed effects, as well as controls for a location's $\ln(x\text{-coordinate})$, $\ln(y\text{-coordinate})$, $\ln(\text{area})$ and four dummy variables denoting whether the dominant language spoken in each prefecture corresponds to one of four of China's main languages spoken (Mandarin, Yue, Wu and Jinyu). p -values, based on standard errors clustered at the province level, in brackets. ***, **, * denotes significance at the 1%, 5%, 10% respectively. In all columns we allow the coefficients for all geography-related variables to differ between the urban and rural parts of prefectures respectively (effectively this means that we include each geography variable interacted with our rural dummy as well as interacted with our dummy indicating the urban part of each prefecture).

⁴⁴ It would have been much harder to explain, had we found that inter-provincial migration responds much more strongly to real wage difference this intra-provincial migration.

⁴⁵ As evidenced by the fact that in China intra-provincial migration flows are much larger than inter-provincial flows, despite the often much larger real wage difference with prefectures in other provinces. Moving to a nearby location e.g. means that it is easier and cheaper to stay in touch or visit relatives and friends.

Appendix D. Relationship initial population size / urbanization and change in real income per worker in the 3 different Scenarios