Hukou and highways

The impact of China's spatial development policies on urbanization and regional inequality

Maarten Bosker[1], Uwe Deichmann[2] and Mark Roberts[3]

November, 2014

Abstract

China has used two main spatial policies to shape its geographic patterns of development during recent decades: restricted labor mobility through the Hukou residential registration system and massive infrastructure investment, most notably a national expressway network (NEN). We develop a structural new economic geography model to examine the impacts of these policies. We fit this model to the data, and simulate various counterfactual scenarios that allow us to compare each policy's respective impact on regional economic development and urbanization patterns across China. We find large overall economic benefits of both the construction of the NEN and of completely abolishing the Hukou system. However, they also result in much more pronounced agglomeration patterns and increased regional economic inequality. The construction of the NEN has only reinforced existing urbanization patterns, while a removal of the Hukou restrictions would also promote urbanization in currently lagging (inland) regions mostly by stimulating rural out-migration. The overall economic benefits are smaller in case of the construction of the NEN. The initially lagging regions not connected to the NEN do not benefit that much from its construction. A removal of the Hukou restrictions instead allows everyone to gain by moving to where he/she is most productive.

JEL Codes: R11, R23, R42

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

Sectors: Transport, Urban

[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 Laura Hering for helpful comments and suggestions. Financial support from the World Bank's Knowledge for Change Program is gratefully acknowledged. 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.

1. Introduction

In recent decades, China's economy has been characterized by rapid economic development accompanied by equally rapid urbanization. The resulting gains have not, however, been shared equally across the country. Rather, they have been mainly concentrated in urban areas and those in its coastal regions in particular. The urbanization of China is expected to continue, or even speed up, in the coming decades. By 2050, the urban share of China's population is forecast to reach almost 76 percent, compared to its current level of around 50 percent.[1] The consequences of this rapid development and urbanization are regarded as one of the defining trends shaping China's, and even the world, economy. 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. The "National New-type Urbanization Plan", released in early 2014, therefore marked a clear shift in the Chinese government's stance towards urbanization. In the plan the government acknowledges that increased urbanization will be one of the defining features of China's continued economic development in the coming decades.[2] It also sets out policies by which it aims to achieve its urbanization goals.

Two policies stand out in this regard.[3] The first is the Hukou, or household registration, system. Every Chinese citizen's Hukou status is determined by his or her origin. It represents an entitlement to welfare benefits and public services (such as education, health care, etc.) in the parents' place of origin. 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, preferring the higher urban wages over the poverty in rural areas. Conservative estimates suggest a stock of rural-urban migrants (or "floating population") without a local Hukou of 117 to 145 million at the end of the last decade; informal estimates exceed 200 million (Chan 2013). Nevertheless, the number of migrants would likely have been much larger in the absence of the Hukou system. This is because many would-be migrants were, presumably, deterred from moving to the city because they were afraid of losing their

---

[1] World Urbanization Prospects: The 2014 Revision, http://esa.un.org/unpd/wup/.

[2] According to the new plan, "Domestic demand is the fundamental impetus for China's development, and the greatest potential for expanding domestic demand lies in urbanization." (http://news.xinhuanet.com/english/china/2014-03/16/c_133190495.htm)

[3] 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.

entitlements to public services (and that of their children since children of migrants have the same Hukou status as their parents), and the discrimination that they still face in most places. 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).

The second policy is the huge investment in large-scale infrastructure projects. In recent decades China has constructed an extensive 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 is to spread development from the more developed coast to the cities in the interior. The highways aim to make non-coastal urban areas more accessible. They are expected to make it more attractive for firms to start producing in interior cities, as it becomes cheaper for firms to import materials and intermediate goods into these cities and to ship final output to the rest of China and the world. Indirectly, this policy is also aimed at reducing migration. By contributing to the economic development of China's interior, workers' incentive to migrate to the coastal cities is expected to diminish as they can now find jobs (at more similar real wage levels) in places that were previously unable to compete with the coastal regions. The recently unveiled "National New-type Urbanization Plan" calls for continued heavy investment in infrastructure.

In this paper, we consider the combined effects of these two main spatial development policies on China's economic geography. Building on earlier work by Roberts *et al.* (2012) and Bosker *et al*. (2012), we incorporate labor mobility into a structural new economic geography (NEG) model to assess how the rapid construction of the intercity national expressway network (NEN) and the Hukou system have jointly shaped China's spatial economy. In particular, we ask how these policies have affected real incomes across urban and rural areas, urbanization rates in different parts of the country, as well as the distribution of both people and economic activity across China. With spatially detailed data and a more comprehensive new economic geography (NEG) model, we are able to derive disaggregated estimates of the impacts of China's two main spatial policies and also to separate out their effects. The main advance of this work over previous studies is that we consider both the effects of the NEN and the Hukou system using a single unified framework, whereas these have only previously been studied separately. We do this by incorporating labor mobility into the structural NEG framework which Roberts et al (2012) used to analyze the impacts of the NEN. In doing so, we make use of prefecture level data on the stock of migrants by rural and urban place of residence to empirically identify the motives for internal migration.[4]

Our main findings are as follows. First, our analysis confirms 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. Geography related amenities are less important, especially for people's decision to move to urban areas. We find only weak evidence of rainfall and temperature differences affecting people's decision to migrate to rural areas.

Second, both the NEN and abolishing the Hukou system will result in more agglomeration of economic activity, although an end to migration restrictions would yield a somewhat more balanced urbanization pattern as lagging regions lose relatively more rural than urban population. This is consistent with the historical experiences of other countries where infrastructure improvements have been found to reinforce economically beneficial agglomeration processes when labor can move freely (e.g., World Bank 2009).

Third, removing the Hukou restrictions does indeed bring large overall welfare gains. So does transport infrastructure investment, but its effect is smaller given that it is mostly confined to the place whose connectivity is improved as a result of the NEN. With labor mobility restrictions in place this does not translate into productivity increases in lagging regions. Allowing people to move to places where their productivity is highest has the largest overall economic benefits, despite coming at the cost of increased regional inequality.

Finally, we find differences in which places gain from which spatial policy. The largest cities and most urbanized prefectures gain most from the NEN construction. Although they are not necessarily the initially richest places, these cities were the ones who saw the largest improvements in their connectivity due to the construction of the NEN. Hukou reform would instead benefits the initially richest prefectures located along China's Southern coastline. These cities are initially not necessarily the most urbanized nor those with the largest urban population, but their economic success will attract migrants from across the country when the Hukou restrictions are loosened.

The structure of the remainder of the paper is as follows. Section 2 describes our overall analytical approach and review prior related work. Section 3 briefly reviews China's two main spatial

[4] Empirical analysis of internal migration decisions within China has thus far been limited by data availability. Our analysis of these decisions, therefore, represents an important contribution to the literature in its own right.

development policies— transport investments and migration restrictions. We present our modeling strategy in Section 4. Section 5 presents the data, estimation and modelling strategy, and Section 6 discusses the estimated impacts of highway construction and migration restrictions on both overall and spatial economic outcomes. Section 7 concludes.

2. Motivation and general approach

We are not the first to look at the aggregate and/or spatial development 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)[5] among others, have considered the effect of China's large-scale infrastructure investments. They consider the effect of the construction of the NEN on economic activity, wages and/or population growth. Roberts et al. (2012) simulate the impact of the NEN using a structural NEG model, whose key parameters they assign through a combination of estimation and calibration. They explicitly distinguish between a rural and an urban sector, and consider the fact that the NEN not only reduced trade costs for urban varieties of goods, but also rural varieties.[6] It allows them to assess the overall welfare implications of the NEN in different Chinese prefectures, as well as the impact of the NEN on intra-prefectural urban-rural real wage differentials. They find that although the NEN increased overall Chinese welfare, it did not have the intended effect of decreasing real income inequality between prefectures, nor did it reduce urban-rural wage inequality. Faber (2014) provides a combination of reduced form estimates of the causal effect of the NEN on economic activity and population growth, and a calibration of a structural model that allows him to generalize his reduced form estimates.[7] His main finding is that the NEN has reinforced the concentration of economic activity in the largest cities. 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, his results show positive welfare impacts of the NEN in all Chinese regions. One shortcoming of both the Roberts *et al*. and Faber papers, however, is that they fail to consider the impact of labor mobility on their findings. The

[5] 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.

[6] Faber (2014) considers a model where agricultural products are costlessly traded, whereas Banerjee et al. (2012) assume that one of the two goods each region produces is not traded.

[7] Banerjee et al. (2012) use a very similar empirical strategy as employed by Faber (2014). Unlike Faber, however, they do not calibrate the model they propose, only showing reduced form estimates.

models they simulate/calibrate both assume that labor is completely immobile, arguing that the Hukou system effectively keeps everyone from migrating.

Papers that have looked at the spatial economic consequences of the Hukou system are, among others, Whalley and Zhang (2007) and Bosker *et al.* (2012)[8]. These studies rely on simulations of a "no-Hukou" scenario that is based on a structural regional economic or, in the case of Bosker *et al*., NEG model. Whalley and Zhang (2007) present simulation results that indicate that the migration restrictions imposed by the Hukou system have strong welfare effects. Their removal would result in both an increase in overall Chinese welfare and a substantial decrease in regional inequality. A big part of this reduction is the result of people moving from the currently underdeveloped interior to the richer provinces on the coast. Their model, however, considers at most China's 31 provinces, thereby abstracting from the substantial variation in regional economic outcomes between prefectures in the same province. Moreover, their model does not consider trade costs, so that regional economic outcomes do not depend in any way on differences between regions' accessibility. Bosker *et al.* (2012) do take trade costs into account. They simulate a structural NEG model under different labor mobility regimes considering a subset of 264 prefectures (out of a total of 333 prefectures). They find that a removal of the Hukou restrictions will result in a much stronger core-periphery pattern than that which already exists today. Their analysis focuses exclusively on the consequences of a removal of the Hukou restrictions on the distribution of people across China's prefecture cities. They do not consider the overall welfare nor income inequality consequences. Also, despite taking trade costs explicitly into account, they approximate them by using the great-circle distance between prefectures. As such, they abstract from the large differences in actual trade costs between different parts of China that are the result of the unequal investment in both the quality and quantity of infrastructure.

Set against this background, our paper's main contribution is to consider both of China's main spatial development policies simultaneously. Our model setup and accompanying empirical strategy allow us to compare the effects of each of these policies on China's spatial economy separately. Moreover, we can also assess whether the construction of the NEN has changed the expected effect of abolishing the Hukou restrictions. To do this, we extend the NEG model introduced by Roberts *et al.*

[8] 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 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).

(2012).[9] In this model, the construction of the NEN changes the geography of production, resulting in substantial changes in relative real wages between prefectures. As discussed above, however, the model's assumption that labor is immobile means that people do not migrate in response to these differences in real wages. This, in effect, represents an extreme Hukou scenario.[10] However, if the Hukou restrictions were removed, people would respond to these differences by gravitating towards places which offer higher real wages. Depending on the exact differences in real wages between places, they will either reinforce the existing agglomeration patterns, or, alternatively, mitigate the agglomerative pressures by moving to China's inland cities that, because of the NEN, now offer them higher expected real incomes.

In our model, we allow for migration both within (from rural to urban areas) and between prefectures. We follow Behrens *et al.* (2013) and Tabuchi and Thisse (2002), and explicitly model people's migration decisions as depending on real wage differences and differences between locational (dis)amenities. Moreover, we allow people to have individual-specific unobserved preferences for living in the rural or urban part of each of China's prefectures. By explicitly modelling people's migration dynamics in this way, we improve 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[11] in combination with various *ad hoc* assumptions on migration costs.

Using data on the urban and rural part of each of the 331 Chinese prefectures in our sample, we estimate the main parameters of our model. Hereby we rely partly on Roberts *et al.* (2012), adopting their estimates for the model parameters that do not relate to people's migration dynamics. A main contribution of our paper is that by using data on the number of migrants in the rural and urban part of each prefecture we estimate how real wages, geographical amenities and public amenities shape the migration decisions of people in China who, despite the current Hukou restrictions, decided to move between or within prefectures anyway. Only relatively few studies exist that estimate the relative importance of 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).[12] We circumvent this problem by

[9] This model itself builds on the original NEG model of Krugman (1991a, 1991b).

[10] Another way of viewing the paper of Roberts *et al*. (2012) is that it only considers the short-run equilibrium impacts of the NEN, whilst, in this paper, we consider the long-run impacts which follow when labor is allowed to respond to the short-run changes in the spatial configuration of real wages induced by the NEN.

[11] Drawing largely on Poncet (2006) for evidence on migration flows.

[12] Poncet (2006) also relies on migrant flows at the provincial levels, 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

relying on the equilibrium conditions of our model. This allows us to identify the important factors driving migration using only information on the total stock of migrants in each prefecture's urban and rural parts.

Based on our estimated model parameters we then simulate the impact of the construction of the NEN and/or a relaxation of the Hukou system. Specifically, we consider three scenarios. The first two consider *the impact of the construction of the NEN*. **Scenario 1** considers the impact of the NEN under the current Hukou restrictons which allow only for partial labor mobility. In **Scenario 2** we instead consider the impact of the NEN in the complete absence of the Hukou restrictions, thereby implying free labor mobility. Comparing these two "NEN-scenarios" allows us to verify claims that increased labor mobility may mitigate the inequality increasing effect of the NEN found in notably Roberts et al. (2012) and Faber (2014). [13] Finally, **Scenario 3** instead considers *the impact of a complete abolishment of the Hukou system*. In this scenario, we take the construction of the NEN as given and consider the spatial impact of allowing all Chinese citizens to freely choose their preferred place of residence. Comparing this scenario with Scenarios 1 and 2 reveals whether, and if so how, China's two main spatial development policies have very different impacts on its regional development. Moreover, Scenario 3 can be compared to earlier results in Bosker et al. (2012) to see whether the impact of abolishing the Hukou system is very different when taking China's large-scale infrastructure investments seriously.

Besides being able to jointly consider the impact of China's two main spatial development policies, our model and accompanying dataset also allow us to consider a richer set of outcomes compared to earlier contributions looking at the impact of either the NEN or the Hukou system. The former focus exclusively on the impact of the construction of the NEN on real income (inequality) between prefectures (Roberts *et al*., 2012; Faber, 2014), or between the urban and rural areas within prefectures (Roberts *et al.*, 2012). The latter (Bosker *et al.*, 2012), focus exclusively on the spatial distribution of people and firms across China's prefectural cities, without detailing the effects on real income inequality or urbanization rates[14]. The fact that we explicitly consider both the urban and rural part of each prefecture allows us to provide a much more complete picture of the impact of the

within provinces. Other studies (notably Rozelle *et al.*, 1999 and Zhao, 1999) provide evidence on the determinants of people's migration by relying on small-scale surveys in particular areas of China. Although very interesting, it is not obvious that their results generalize to other parts of China.

[13] We also compare our findings to the no labor ("extreme hukou") mobility regime, as used in Roberts *et al.* (2012) or Faber (2014), but refer to Roberts *et al.* for an in depth discussion of this scenario.

[14] 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.

NEN and/or Hukou system on China's spatial development. In particular, we consider their impact on both real income inequality between and within (urban vs. rural) prefectures, evaluate how they affect the spatial distribution of people across China's prefectures, and at the same time assess their impact on the urbanization rates in different parts of China.[15]

3. A brief history of the NEN and the Hukou system[16]

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, and 18 East-West connections, giving it the unofficial name "7918 network". The first phase of the network—connecting all cities with population above 500,000 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 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 impacts from highways from those of railway lines would be very difficult.

While large scale transport infrastructure investments have been a fairly recent feature of China's spatial policies, China's unique 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

[15] 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.

[16] 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 and, in particular, in the references we cite).

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).

Despite their restrictive nature, however, 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). The total stock of non-Hukou population was thought to be above 145 million in 2010. 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 been considerably lower than it would otherwise have been.

4. The model

We build on the NEG model by Roberts *et al.* (2012). In this section we briefly set out the main features of their model. In doing so, we focus on the way transport costs determine real incomes in the urban and rural sectors of each prefecture by influencing market access given that this provides the main conduit through which the NEN affects outcomes. The model in Roberts et al. (2012) assumes no labor mobility, neither between prefectures nor between the urban and rural sector within each prefecture. We set out in detail in section 4.1 how we extend the model to incorporate both types of labor mobility.

The model in Roberts *et al.* (2012) is an elaboration of the original NEG model of Krugman (1991a, 1991b). Thus, as with Krugman's original model, the model consists of two sectors – an urban sector which is characterized by internal economies of scale and monopolistic competition, and a constant returns rural sector in which perfect competition prevails. However, whereas Krugman's original model is restricted to two regions, the model in Roberts *et al* follows both Fujita *et al*. (1999) and Fingleton (2005, 2007) in generalizing it to many regions. Likewise, the Roberts *et al*. model extends both "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. 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 the urban and rural sectors, not to mention the overall level of income. Most notably, a prefecture's wage in each sector depends on two factors – (i) the region's (exogenous) level of labor efficiency in that 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 NEN – affects the overall spatial equilibrium and, in particular, levels of real income and wages both across across and within prefectures.

Intuitively, a decline in transport costs associated with 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 transporting their output to another prefecture *j*, increases demand for prefecture *i*'s output and, therefore, also its RMA. Given labor immobility this pushes-up the urban wage. Operating against this, however, the reduction in transport costs also exposes urban firms in region *i* to greater competition from urban firms located in both prefecture *j* and other prefectures with which it is indirectly connected through prefecture *j*. This resulting increase in competition results in countervailing downward pressure on prefecture *i*'s RMA. The overall impact on the urban wage depends on which of these opposing forces dominates.[17]

4.1 Introducing labor mobility

Our main addition to the Roberts *et al.* (2012) model is the introduction of labor mobility, both between prefectures as well as between the urban and rural parts of each prefecture. In order to be able to also model the migration between the urban and rural part of each prefecture we assume that the urban part of each prefecture focuses exclusively on the production of the urban commodity and its rural part specializes in the production of the rural commodity. 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.[18]

Roberts *et al.* (2012) consider the 2007 distribution of people across China's prefectures (and their urban and rural part respectively) as reflecting the equilibrium of their NEG model without labor

[17] 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.

[18] 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).

mobility. Basically, their model rationalizes the observed real wage differences between urban and rural areas and between different prefectures in China by the fact that people cannot move between places in response to these differences (so that they persist in equilibrium). However, although restricted by the Hukou system, people in China did move (in fact estimates are that about 260 million people have migrated over the last three decades, mostly from rural to urban areas; see World Bank and DRC, 2014). To do justice to these migration flows, we adapt the Roberts *et al.* (2012) model to allow for labor mobility both between and within provinces. The simplest way to do this would be to allow people to move in response to the real wage differences generated by the original model (see Roberts *et al.*, 2012). However, labor mobility in response to these real wage differences would imply the long-run equilibrium prediction that real wages are equalized, both across the urban and rural sector, as well as across all prefectures in China. Any real wage differences would immediately result in a migration response, driving these differences to zero again.

In order to rationalize the 2007 scenario as a spatial equilibrium of the Roberts *et al.* (2012) model with (some) labor mobility, we therefore need to make a further modification to the model. We do this in the simplest possible way. In particular, we follow Tabuchi and Thisse (2002) and Behrens *et al.* (2013) by adapting people's utility function as follows.[19] We assume that a person's likelihood to choose to live in location i, where this location can be either the urban or rural part of a prefecture,is based on the utility he or she derives from living in that place. This utility depends linearly on real income W_i earned in location i (see Appendix A), location i's amenities, A_i, and an idiosyncratic preference for living in location i that differs for each individual j, ε_{ij}:[20]

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

People are free to choose their location, and do so by moving to that location which provides them with the highest utility. The probability that individual j chooses to live in location i is:

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

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

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

[19] 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 heterogenous 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 few cities remain in existence).

[20] Note that all individuals equally value amenities and real income.

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. In this case people's idiosyncratic preferences for each location are very heterogenous 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, which is given by $\boldsymbol{P_i}$:

$$\boldsymbol{P_i} = P(U_{ij} > max_{k \neq i}\, U_{kj}) = exp((W_i + A_i)/\mu) \,/\, \textstyle\sum_k[\exp((W_k + A_k)/\mu)] = L_i \,/(\sum_k L_k) \qquad (4)$$

In the Roberts *et al.* (2012) model, the persisting real wage differences were rationalized by assuming no labor mobility (with this equilibrium defined by equations [A1] – [A5] in Appendix A[21]). In our extended model, (4) shows that these differences can still persist in a spatial equilibrium where people do move between locations, but are offset by differences in the amenities offered at different locations, as well as by allowing different people to have different idiosyncratic preferences for living in each location.

5. Estimating the main model parameters

Having set out our model, detailing how we allow for 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.

5.1 Data

Our analysis uses a mix of spatial data extracted using geographic information system (GIS) techniques and statistical data. We build on Roberts *et al.* (2012, Appendix A) for data on the NEN and prefecture level information on urban and rural wages, income, investment per worker, human

[21] See also equations [3] – [7] in Roberts *et al.* (2012).

capital (years of education), and land area for 331 prefectural cities (three prefectural cities were excluded due to missing data) and similar administrative units.[22] 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 ($t_{ii} = 0$).[23]

Largely because of its unique registration system, Chinese migration statistics are notoriously difficult to interpret (Chan, 2013). 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" for both years. Information on amenities (the share of households with access to tap water, toilets and natural gas supply) come from the same data set. [24]

Finally, data on geographic characteristics/amenities assumed to influence migration decisions were derived as follows. 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°. The dominant language for each prefectural region comes from a GIS data set of linguistic regions called the World Language Mapping System (Ethnologue, 2004).

[22] 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).

[23] In effect, 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 a detailed study on inter-prefectural infrastructure.

[24] 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.

5.2 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).

5.2.1 The determinants of migration in China

To identify the relative importance of real wages and amenities in people's migration decisions, we make use of the spatial equilibrium condition (4). From this condition it follows that we can do this using only information on each location's stock of migrants (see also Behrens *et al.*, 2013).

The strong assumption underlying our identification is that China's spatial economy is in equilibrium in 2007.[25] 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 into equation (4): they denote the W_i for each prefecture's rural and urban part respectively. This allows us to back out the indirect utility levels that correspond to the 2007 spatial equilibrium, *now allowing for labor mobility*. 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 Shijazhuang prefecture). Next divide each location's $\mathbf{P}_i$ by that in this baseline location (and use the normalization $W_1 + A_1 \equiv 0$):

$$\mathbf{P}_i / \mathbf{P}_1 = exp((W_i + A_i)/\mu) / 1 = L_i / L_1 \qquad (5)$$

From this it immediately follows that each location's log population relative to that in the baseline location is directly related to the indirect utility derived from the real income and amenity levels in that location (relative to the baseline location):

$$\ln[L_i / L_1] = (W_i + A_i)/\mu \qquad (6)$$

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

[25] 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).

regression of each location's log population relative to the baseline location on its calibrated W_i and its observed amenities:

$$\ln[L_i / L_1] = \beta_0 + \beta_1 W_i + \beta_2 A_i^{obs} + \varepsilon_i \quad (7)$$

where β_0 captures the reference location's share in Chinese population. Moreover, assuming that the fitted residuals can be interpreted as the unobserved part of each location's amenities, we also obtain an estimate of $A_i^{unobs} = \widehat{\varepsilon_i}$, and their relative importance in individuals' location choice.[26]

5.2.2. Estimation results

We use the above outlined strategy to estimate the relative importance of real wages, various geographical amenities, and different public amenities in determining people's location choice. 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 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.[27]

[26] 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.

[27] Full results are available upon request.

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, instead of regressing our real wage and amenity data on log population relative to a baseline location, we focus on each location's log *migrant stock* relative to that in the baseline location as our main independent variable. For most people in China the costs of migration are simply too high, in a large part due to the Hukou restrictions; the inability to sell one's land is another important reason (World Bank and DRC, 2014). This implies that a large share of the population is not making an active location choice as implied by our model (see (2)). If these restrictions are big enough (which they arguably are), it is hard to identify the willingness to move in response to amenity and/or real wage differences that people in principle have. In fact, one can even argue that, when estimating (7) using total population shares, 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] Migrants (about 260 million of them over the last three decades), however, did make this choice. In particular, they decided to move despite the heavy costs of migration they faced. We effectively identify people's willingness to migrate in response to real wage and/or amenity differences from the variation in location choice of people who did decide to migrate. This implies that in our counterfactual "no Hukou" scenario we make the implicit assumption that current migrants' preferences are representative of those of the entire Chinese population.[29]

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.

[28] See also Desmet and Rossi-Hansberg (2013) who attempt to capture Hukou impacts in this way.

[29] 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.

Table 1a. Determinants of the stock of migrants

	(1)	(3)		(4)		(5)
Dep. Variable:	ln migrants	ln migrants		ln migrants		ln migrants
ln real wage	0.986	0.923		0.749		0.608
	[0.00]***	[0.00]***		[0.00]***		[0.008]***
		urban	rural	urban	rural	no split Geo
ln rugg	-	-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-coordinate), ln(y-coordinate), ln(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).*

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.[30] In our preferred specification in column 4, a 1% higher real wage corresponds to attracting a 0.75% 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

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

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.

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.[31] The most striking difference is the fact that we find a much weaker association between real wages and these shares.

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.[32] 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, 1999), all our evidence supports the idea that real wages are indeed a very important driver of migration decisions in present day China.

[31] 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.

[32] 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].

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		0.371		0.051	
	[0.100]*		[0.046]**		[0.022]**	
	Urban	rural	urban	rural	urban	rural
ln rugg	-0.122	-0.174	-0.128	-0.166	-0.012	-0.009
	[0.005]***	[0.001]***	[0.006]***	[0.001]***	[0.278]	[0.172]
ln cooling days	-0.072	-0.031	0.012	-0.028	0.032	-0.093
	[0.814]	[0.911]	[0.968]	[0.919]	[0.628]	[0.266]
ln heating days	0.263	0.439	0.271	0.419	-0.028	-0.070
	[0.043]**	[0.001]***	[0.073]*	[0.002]***	[0.626]	[0.003]***
ln rainfall	-0.027	0.677	-0.011	0.670	-0.079	-0.017
	[0.847]	[0.012]**	[0.943]	[0.012]**	[0.178]	[0.555]
D yangtze	0.045	-0.194	0.041	-0.247	0.022	0.007
	[0.763]	[0.331]	[0.788]	[0.241]	[0.52]	[0.842]
nat.res.index	-0.001	-0.039	-0.002	-0.036	-0.002	0.009
	[0.932]	[0.062]*	[0.87]	[0.112]	[0.465]	[0.181]
% hh water	-0.756		-0.560		0.104	
	[0.012]**		[0.066]*		[0.008]***	
% hh toilet	1.121		1.259		0.109	
	[0.023]**		[0.012]**		[0.098]*	
% pop gas	0.397		0.337		-0.182	
	[0.029]**		[0.081]*		[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-coordinate), ln(y-coordinate), ln(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).*

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 part (they are the sum of each location's residual and its corresponding provincial fixed effect[33]).

[33] Which is different depending on whether the location is rural or urban.

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

Now that we have obtained estimates of the key migration parameters to add to the values of the structural model parameters derived by Roberts *et* al. (2012) , we are 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 set out each of these three scenarios below, detailing how we generate the counterfactual spatial equilibria that each of these scenarios rely on. Our main focus in all three scenarios is on how the NEN and/or Hukou restrictions shape the spatial distribution of real income and people across China. We consider both inequality between prefectures as well as inequality between the urban and rural part within each prefecture. More specifically, our analysis focuses on the distribution of four main variables across China's prefectures:[34] real income per worker, the urban-rural real wage gap, urbanization rates, and overall population. 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).

6.1 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 and rural part 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. We generate our counterfactual scenarios starting from this "NEN + Hukou" situation by either removing the NEN so that travel times between cities are changed back to what they were before the NEN, and/or completely removing the restrictions on labor mobility so that each person is free to move to the location of his/her choice.

<u>Scenario 1: The impact of the NEN under the current Hukou system</u>

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 (see equation 7 above and Appendix A) back to the situation before the NEN was built.

[34] 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.

To find the spatial equilibrium corresponding to this "new" situation, we first numerically solve for the (short run) equilibrium values for w^U_i, w^R_i, G^U_i, and G^R_i, 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^U_i, w^R_i, G^U_i, and G^R_i to calculate real wages in the urban and rural part of each prefecture: $\omega^U_i = w^U_i/P_i$ and $\omega^R_i = w^R_i/P_i$ respectively, where $P_i = (G^U_i)^{\theta}(G^R_i)^{1-\theta}$ and G^U and G^R are the urban and rural price indices respectively

Subsequently we use these counterfactual real wages to change each location's population relative to that in our baseline location (defined earlier) according to equation (7)[35]:

$$\ln(L_i^{U*}/L_1^*) = \widehat{\beta_0} + \widehat{\beta_1}\omega_i^{U*} + \widehat{\beta_2}A_{i,U}^{obs} + \widehat{\varepsilon_{i,U}^{obs}} \quad (8)$$

, and similar for the rural population in prefecture *i*, $\ln(L_i^{R*}/L_1^*)$. In a Hukou scenario with restricted labor mobility, we can then calculate each location's counterfactual population as follows:

$$L_i^{*,Urban} = L_i^{non-migrant} + L_{CHINA}^{migrants} exp\left(L_i^{U*}/L_1^*\right)/\sum_k[exp\left(L_i^{U*}/L_1^*\right) + \exp(L_i^{R*}/L_1^*)] \quad (9)$$

, and again similar for $L_i^{*,Rural}$. That is, we assume that under the Hukou scenario the only people who are potentially willing to migrate (despite the severe consequences that this entails because of the loss of Hukou rights), are those people who actually migrated to another location with the Hukou system in place (i.e. the "revealed" migrants in our dataset).

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 get at how, we numerically resolve for the equilibrium urban and rural wages and price levels based on the equations in Appendix A that correspond to 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

[35] 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.

iterations). Upon convergence[36], 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 the NEN under free labor mobility (i.e. no Hukou restrictions)

In this scenario, we compare the counterfactual "NEN + no Hukou"- to the counterfactual "no NEN + no Hukou"- spatial equilibrium. We simulate these counterfactuals using the same iterative procedure as outlined for Scenario 1 above. We employ the NEN travel time matrix to determine transport costs in the "NEN + no Hukou" scenario, and the no NEN travel time matrix in the "no NEN + no Hukou" scenario, respectively.

We simulate a counterfactual scenario under a complete abolishment of the Hukou system as follows. Instead of updating each location's population according to (XXX) that makes the "Hukou assumption" that only migrants potentially respond to real wage differences between locations, we now allow all Chinese people to do so. In particular, we change (9) to:

$$L_i^{*,Urban} = L_{CHINA} exp\left(L_i^{U*}/L_1^*\right)/\sum_k [exp\left(L_i^{U*}/L_1^*\right) + \exp(L_i^{R*}/L_1^*)] \qquad (10)$$

again similarly for $L_i^{*,Rural}$. Both the "NEN + no Hukou" and "no NEN + no Hukou" counterfactual scenarios employ (10) when iteratively solving for their respective spatial equilibrium. Comparing the two reveals what the impact of the NEN would have been had there been no restrictions on labor mobility over the period of its construction. This scenario can be directly compared to Scenario 1. It provides evidence on whether, and if so how, the large investments put into the construction of the NEN would have had different effects on China's spatial economy had people's possibilities to migrate not been restricted by the Hukou system. We can use this information to further assess the postulated mitigating effect of free labor mobility on the increase in regional inequality associated with the construction of the NEN found in Roberts *et al.* (2012).

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

Scenario 3: The impact of abolishing the Hukou system

Our final scenario focusses on how the abolishment of the Hukou system would affect China's spatial development. It does so by comparing the observed 2007 "NEN + Hukou" spatial equilibrium to the counterfactual "NEN + no Hukou" spatial equilibrium, that we already simulated in Scenario 2 (see above).

Scenario 3 can be directly compared to Scenarios 1 and 2 to provide evidence as to 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 an abolishment of the Hukou restrictions at the prefecture level, but without explicitly taking the large-scale investments in infrastructure into account (see our earlier discussion in section 2). 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 considers the changes in, *inter alia*, real wages and urbanization levels.

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

In this section, we discuss our main findings. We first consider the results from Scenario 1 – i.e. the impact of the NEN given the Hukou restrictions by accounting for restricted labor mobility instead of in the no labor mobility assumption of earlier contributions (e.g. Roberts *et al.*, 2012 and Faber, 2014). Next, we turn to the Hukou restrictions and show that their removal is predicted to have much stronger impacts on regional welfare and agglomeration patterns in China than the current large-scale investments in infrastructure. Also, we find that both policies often have very different spatial effects. 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 – 5 further illustrate some of the most interesting correlations and provide detailed maps of the spatial distribution of these changes across China's prefectures.

6.2.1 The impact of the NEN with and without the Hukou restrictions

First we consider the impact of the NEN under different labor mobility regimes (i.e. Scenarios 1 and 2). 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). Both scenarios show an increase in real income per worker of about 6%, accompanied by an increase in regional inequality (as measured by the standard deviation of real income per worker across prefectures) of about 9%.

If anything, we find that allowing for labor mobility, albeit restricted by the Hukou system, results in more, not less, inequality, both within and between prefectures. The standard deviation of real income per worker across prefectures, as well as the average within prefecture urban-rural wage gap increase more in Scenario 1 than in the Roberts *et al.* (2012) extreme (zero labor mobility) Hukou case. Moreover, about 25 (8 percent) prefectures more now witness an increase in the urban-rural wage gap compared to the "extreme Hukou" case. The same holds when looking at population movements. These were assumed to be absent in the extreme Hukou scenario of Roberts *et al*. When allowing for restricted labor mobility, we observe a slight increase in agglomeration, with the standard deviation of both population and urbanization increasing. However, the population changes that we find are surprisingly small. Thus, China's overall urbanization rate e.g. increases by a mere 0.4 percent point. This very small increase does hide some significant spatial differences however. As shown in Figure 3b and the maps in Figure 4, the initially most urbanized places are the ones found to have urbanized fastest as a consequence of the NEN, and they also found to be the ones which have attracted the most new migrants (see also the correlations in Table 3). The construction of the NEN has only strengthened the existing agglomeration pattern in China, reinforcing its urban core at the expense of its rural hinterland. Nevertheless, do note that even the fastest urbanizing places never witness an increase in their urbanization rate of more than 4 percentage points.

Table 2 Counterfactual results

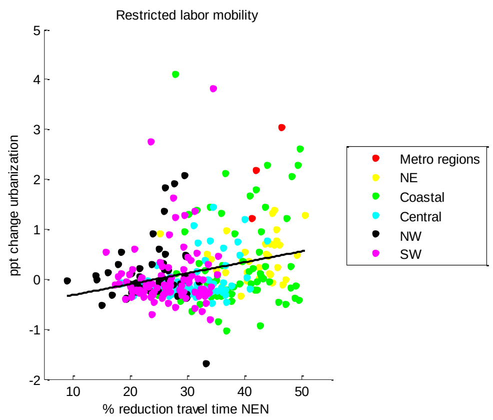
Counterfactual scenario:	**Roberts et al. (2012)** NEN vs. no NEN (extreme Hukou)	**Scenario 1** NEN vs. no NEN Hukou	**Scenario 2** NEN vs. no NEN no Hukou	**Scenario 3** Hukou vs. no Hukou NEN
change in aggregate Chinese				
real income (pw) (%)	6.0	6.6	6.0	169.7
sd real income pw (%)	8.6	9.0	4.7	50.2
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

Of course the fact that we only allow “revealed Hukou migrants”, about a fifth of the total Chinese population, to 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, our Scenario 2 results show that even in the absence of any Hukou restrictions, with everybody free to move to his/her place of preference, 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, this aggregate number now hides more substantial differences in the spatial distribution of the welfare and population effects that the construction of the NEN entails.

Two things stand out. First, with labor completely mobile we see a much smaller increase in regional inter-prefectural income inequality as a result of the construction of the NEN. The increase in the standard deviation of real income, as well as the correlation between initial real income per worker and the change in real income are about half that in Scenario 1 (see Table 3 and Figure 2a below). Figure 2a shows that now many initially wealthy places even observe a fall in real income per worker. Moreover, 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 the Hukou scenario). A higher degree of labor mobility indeed appears to mitigate the rise in regional income inequality association with the construction of the NEN, while delivering a similar overall welfare increase for the average Chinese citizen.

Second, the spatial impact of the NEN is very different in the two scenarios. In both scenarios the places whose connectivity is most improved because of the construction of the NEN gain most in terms of real income per worker and population (see the correlations in Table 3). However, the correlation of improved connectivity with real income per worker growth is strongest in the Hukou scenario (Scenario 1), whereas it is strongest with population growth in the no Hukou case (Scenario 2). Moreover, with the Hukou restrictions in place the places that benefit most from the construction of the NEN in terms of their connectivity also witness an increase in intra-prefecture urban-rural disparities, both in terms of population (urbanization) and welfare (urban-rural wage gap). In the absence of any restrictions on labor mobility this relationship with intra-prefecture inequality is absent. Figure 1 shows this in more detail for the NEN's impact on urbanization rates across China's prefectures.

Figure 1. % reduction in travel time because of the NEN and ppt change in urbanization



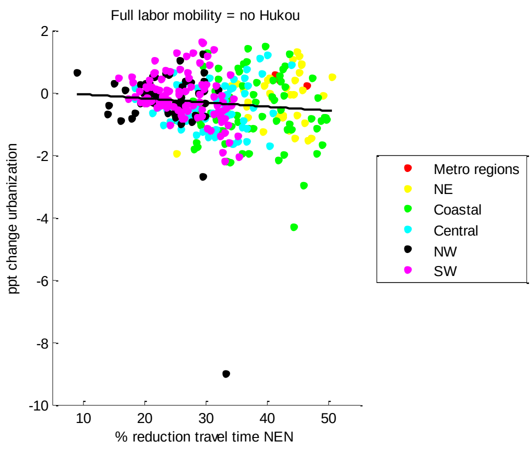


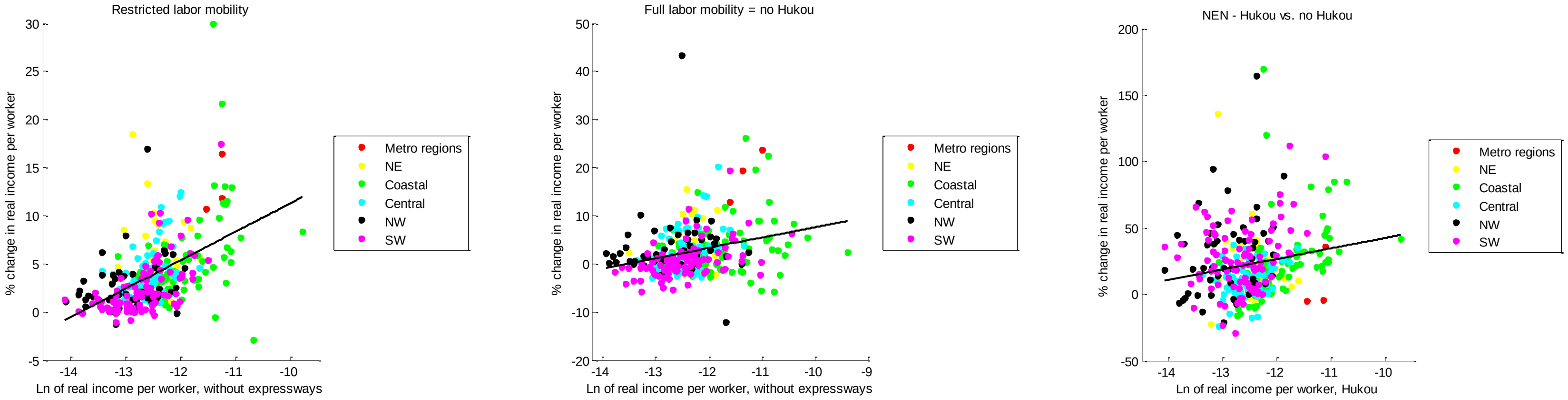
Table 3: Correlation matrix outcomes different scenarios

Correlations	Scenario 1 NEN vs. no NEN (Hukou)				Scenario 2 NEN vs. no NEN (no Hukou)				Scenario 3 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

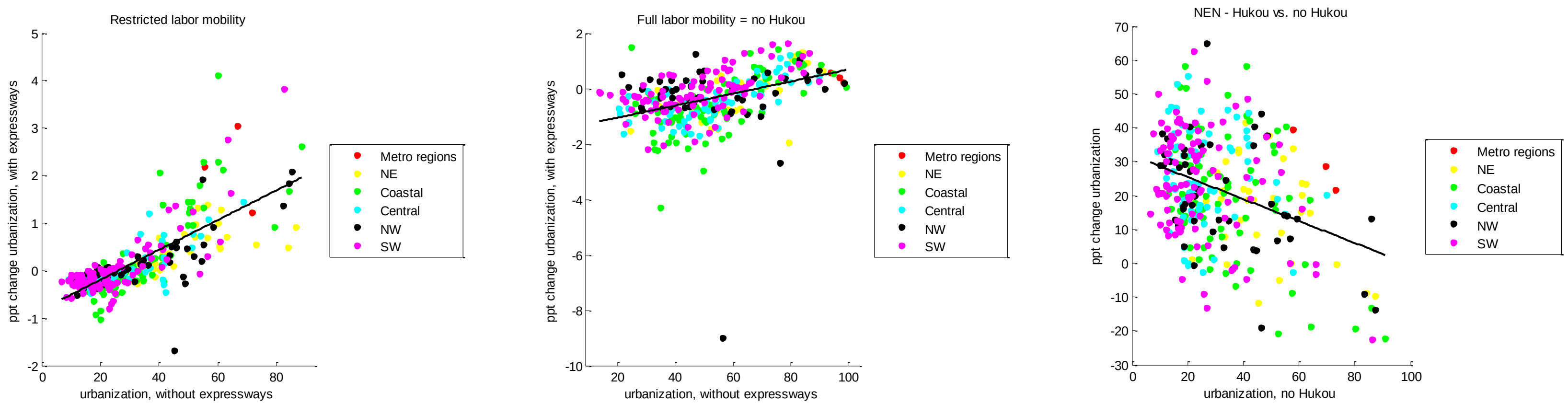
Notes: bold correlations are NOT significant at the 5% level.

Figure 2. Convergence or divergence? Real income per worker and urbanization

a. Real income per worker



b. Urbanization

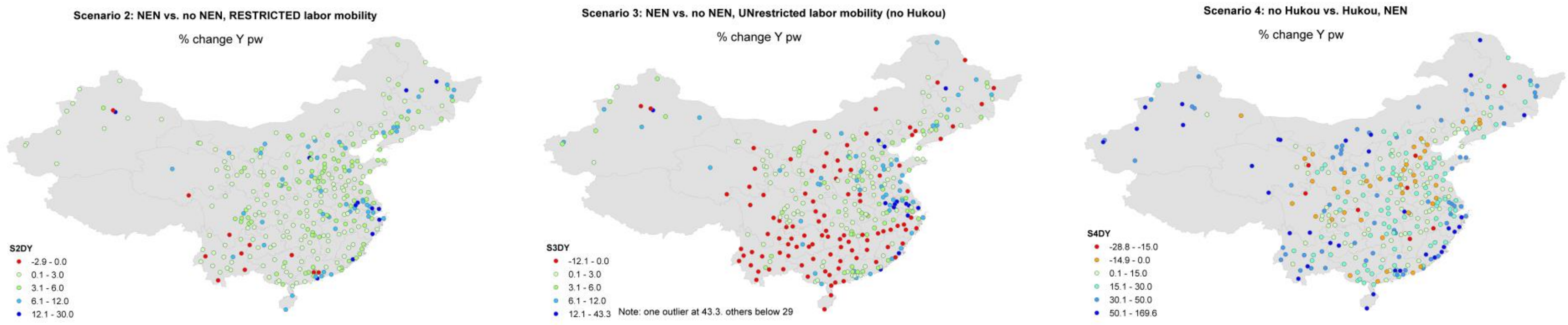


These differences between Scenarios 1 and 2 can be explained by the fact that without any Hukou restrictions in place, many more people respond with their feet to the changes in real incomes induced by the construction of the NEN (see the higher correlation between the change in real income per worker and the change in population in Scenario 2 shown in Table 3). 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 3c). It results in more spatial inequality in terms of population (the standard deviation in prefecture population increases almost twice as fast in Scenario 2), reinforcing the existing agglomeration pattern in China. However, it is also exactly this higher degree of labor mobility in the absence of the Hukou restrictions that mitigates 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 2 (again 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. As most of these “movers” come from the (poorer) rural part of their prefecture, this reduces the positive correlation between initial urbanization levels and subsequent urbanization growth. Urbanization rates now also go up in many prefectures in inland China despite the fact that they are often losing overall population (see Figure 3b and 3c). At the same time, the 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.

Figure 3. Spatial impacts in each of the three Scenarios (XXX Note: in these maps the scenario numbers should each be reduced by 1 XXX)

a. Percent change in real income per worker



b. Percent change in urbanization

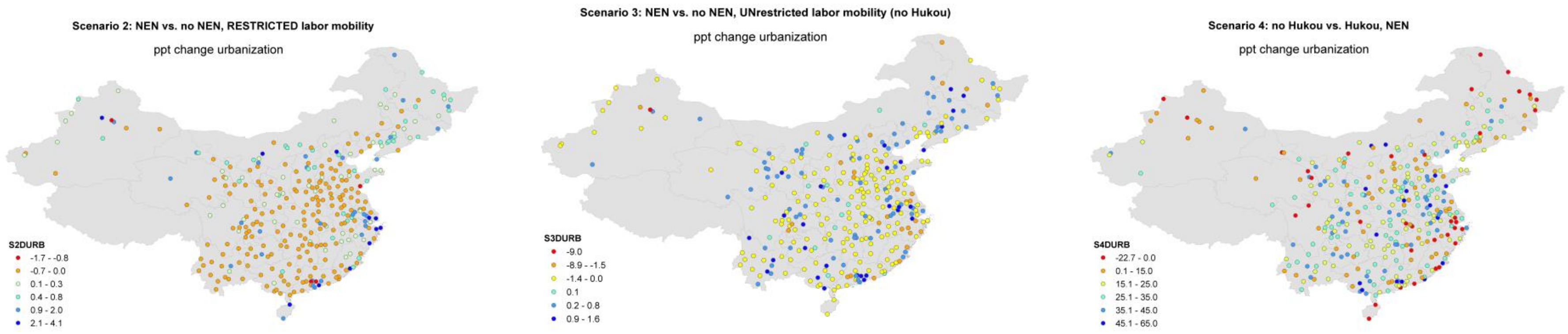
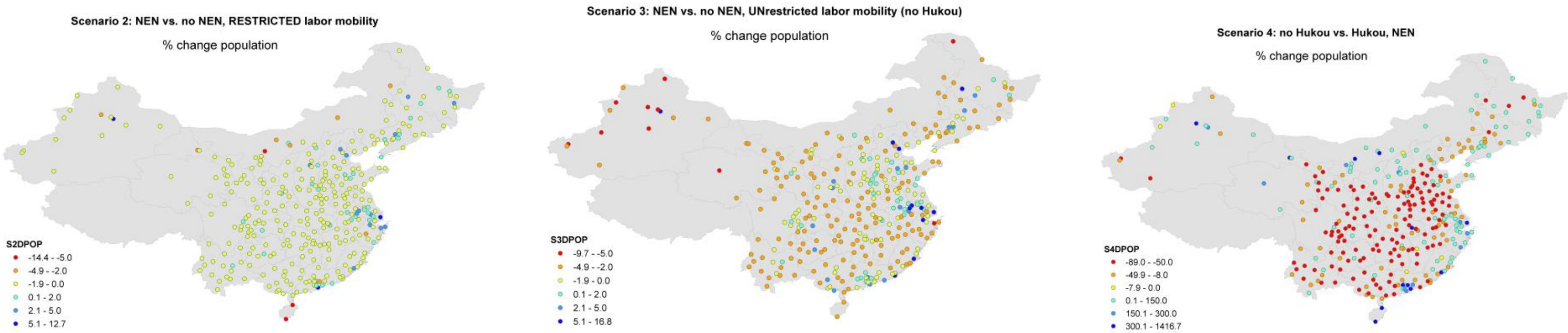


Figure 3, continued

c. **Percent change in total population**



Summing up, in both Scenarios 1 and 2, the construction of the NEN results in a similar 6 percent increase in Chinese real income per worker, which is associated with a 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]

6.2.2 The impact of removing the Hukou restrictions

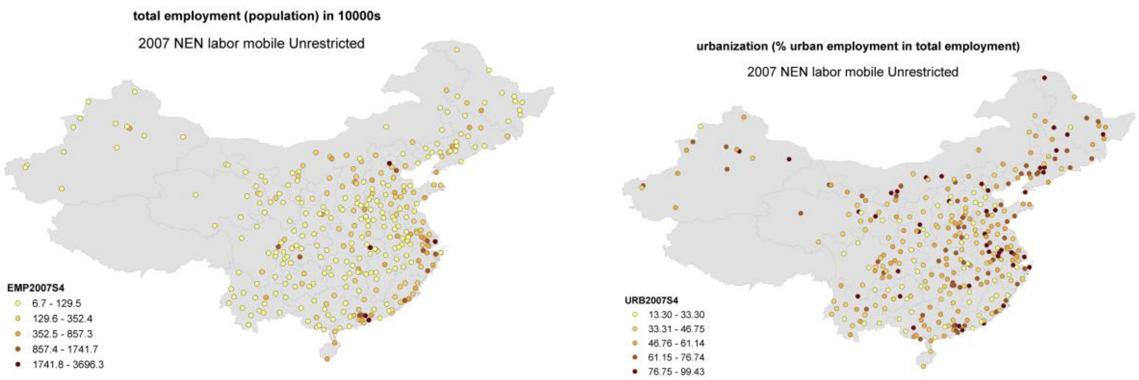
In Scenarios 1 and 2 we focussed on the spatial impact of the construction of the NEN under two different labor mobility regimes. In this section we turn our attention to the aggregate and spatial development impacts of China's other main spatial development policy: the Hukou system. In Scenario 3, we 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 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%. 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 3a shows that these are mainly located in China's coastal regions). Some initially poor

[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 500,000 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 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 restriction 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 3.

regions, in e.g. China's South West also benefit a lot (especially when compared to the impact of the construction of the NEN, see Figures 2a and 3a), but given that they start from an initially much lower base, their incomes still diverge from those in rich (coastal) China. Besides this rise in inter-prefectural real income inequality, we also see that abandoning the Hukou restrictions 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 2, 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 compared to the two NEN scenarios (i.e. compared to Scenarios 1 and 2). This is so despite the fact that more prefectures witness a population increase in Scenario 3 (see Table 2). As Figure 3c shows, inland China sees a large outflow of people towards the coast (and some places in the North(-East)). Figure 4 shows that our simulation indeed predicts that a China without Hukou restrictions would be a China where most people live in the prefectures along its (southern) coastline. 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 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. Urbanization rises in almost all prefectures, and, interestingly, this rise is larger in the initially less urbanized places (see Figure 2c), which is very different to what we found in Scenarios 1 and 2. Despite the fact that the lifting 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 3c; the negative correlation between changes in population size and changes in urbanization rate in Table 3 also aptly summarizes this pattern). This urban catch-up is mainly driven by a declining overall population 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 3 and also A2). These places were among the least benefitting places from the construction of the NEN, and

[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) completely abstract from the construction of the NEN, approximating travel times between cities by simple geodesic distances and starting their simulations in 2000 (before the NEN was built). Second, they consider only 264 prefectures, compared to the 332 that we consider (Figure 2 in their paper shows that they are primarily missing 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, fourth, we explicitly consider each prefecture's urban and rural part so that not only inter-prefecture migration considerations determine the spatial equilibrium but also intra-prefecture migration considerations.

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 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 3 appear very large (especially so when compared to Scenarios 1 and 2), but one has to remember that they reflect the (very) long run impact of a relaxation of the Hukou restrictions. 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 prefectures (although one could argue that some of these congestion forces are to some extent captured by each location's unobserved amenities, see equation (7) and the discussion in section 4). 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.

Nevertheless, we do think that our numbers convey the strong message that the impact of an alleviation of the strong restrictions on labor mobility that are present in China today will indeed result in big welfare gains for most people in China as each individual is now able to move to where he/she is most productive. Moreover, it will result in a much more 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, our simulations do find that they have resulted in a modest increase in overall Chinese income per worker of about 6 percent. But, so far, infrastructure investment has tended to favor the already large, urbanized places. And, with the Hukou system keeping many people in lagging regions from benefitting from the resulting increase in real incomes in these now better-connected places, the effects of infrastructure investment have been small, or even negative, in the already poorer Chinese prefectures. This is quite the opposite of the expected spread of economic activity that the large scale investment in the NEN meant to achieve.

7. Conclusions

China's urban population share is expected to rise by about 25 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. Better understanding the impacts of these policies will help adapt them to achieve the greatest social and economic benefits.

This paper compared the spatially differentiated impacts of the construction of China's national expressway network (NEN) with those from relaxing its long-standing migration restrictions. Despite the stated objective of promoting development in lagging regions, our analysis suggests that the NEN has led to further concentration of people and economic activity in high potential regions, mostly along the coast. Increased labor mobility following the end of Hukou would reinforce this trend. But, where the construction of the NEN has only reinforced existing urbanization patterns, a removal of the Hukou restrictions would also promote urbanization in currently lagging (inland) regions, be it by predominantly stimulating increased rural out-migration.

Our analysis also shows that pursuing these two policies has high overall economic payoffs. This is especially so when allowing people to more freely migrate to the places where they are most productive. Continuing to restrict migration with the objective of more balanced regional development therefore has high aggregate costs.

Our analysis extends only to 2007 because after this year it becomes more difficult to isolate the effect of the NEN. Since then, 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 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.

Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies*, rdu010.

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

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/ .

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., & DeBrauw, A. (1999). Migration, remittances, and agricultural productivity in China. *American Economic Review*, *89*(2), 287-91.

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.

Zhao, Y. (1999), Labor migration and earnings differences: the case of rural China. *Economic Development and Cultural Change*, Vol.47(4), 767-782.

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.

Appendixes

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:

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

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

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

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

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

where *i* = 1, …, *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 $\kappa_i^R = E_i^R / E_{\min}^U$, where E_i denotes region *i*'s absolute level of labor efficiency and $E_{\min}^U = \min(E_1^U, ..., E_N^U)$. 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.

Table A1. Ln Migrants, by different origin

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

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 results[42], as it is expected that people are more sensitive (aware) to real wage difference with nearby locations[43].

[42] 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.
[43] 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