

# The Distributional Impacts of Transportation Networks in China

Lin Ma                  Yang Tang \*

## Abstract

This paper documents the evolution of transportation networks in China and evaluates its impact on regional income differences. We provide the first panel dataset of the transportation network in China by standardizing a comprehensive collection of maps on the road, railroad, and waterway transportation between 1995 and 2016. Our dataset corrects the inconsistencies in construction standards and infrastructure qualities over time and space. Based on this dataset, we show that the expansion of transportation networks has significantly reduced the income gap between cities. This is mainly because inter-city trade is more responsive to network improvements than migration. As a result, although small cities shrink in population size due to emigration, the negative impacts are quantitatively dominated by improved market access, leading to gains in real wages.

**Keywords:** regional trade; migration; welfare; economic geography

**JEL Classification:** F1; F4; R1; O4

---

\*National University of Singapore (ecsml@nus.edu.sg) and Nanyang Technological University (tangyang@ntu.edu.sg), respectively. We are grateful for comments by Davin Chor, Carlos Garriga, Maximiliano A. Dvorkin, Wen-tai Hsu, and participants at the Federal Reserve Bank of St.Louis, and Singapore Management University. Song Wang provides superb research assistance. Yang Tang acknowledges financial support from Singapore MOE-Tier 1 research fund(M4011853). We are solely responsible for the remaining errors.

# 1 Introduction

Over the past decades, developing countries have invested extensively in transportation networks to pave the way for prosperity.<sup>1</sup> While it is well-understood that the improvements in transportation infrastructure result in economic gains on average,<sup>2</sup> the distributional impacts are much less clear. On the one hand, improvements in trade frictions allow the remote and small cities to enjoy better access to the markets in large cities, which often leads to lower spatial inequality. On the other hand, however, better roads also improve factor mobility, which tends to drain resources from remote locations. In a world with agglomeration, the concentration of production factors could further strengthen the existing productivity advantage in large cities. Therefore, the distributional impact of transportation infrastructure is a quantitative question. This paper aims to provide an answer through a general equilibrium model that accommodates both channels in the context of China.

The first step towards a quantitative answer is a panel dataset depicting the development of transportation networks in China. Based on such a dataset, a researcher can then measure the distance between regions using different transportation modes, and exploit the time-series and cross-sectional variation in the measured distance. Unfortunately, the existing datasets are not up to the task because the digitized maps are not comparable across time and space. The incomparability is due to several reasons. First and foremost, the standards by which the roads are designed and constructed have significantly improved over the years, so certain highways constructed in the earlier years are no longer classified as so in the recent years. Moreover, even for two roads of the same vintage and classification, the road design standards vary by region and terrain. In many cases, the speed limit and capacity of highways in the inland regions are only about half of those in the coastal regions, leading to difficulties in cross-sectional comparison. In addition, the map publishers often use different map-projections over the years, causing discrepancies in cross-time comparison. As a result of all these issues, the construction of a panel dataset by directly stacking geo-referenced

---

<sup>1</sup>For example, [Gurara et al. \[2018\]](#) estimated that the developing countries, on average, spend around 6 percent of GDP on infrastructure each year. The World Bank predicts that in the next decade, developing countries need to spend around 3.3 percent of GDP on transportation networks to meet the future demands for mobility, according to [Rozenberg and Fay \[2019\]](#).

<sup>2</sup>The literature of transportation costs on economic performance is extensive; see [Redding and Turner \[2015\]](#) for a summary of the literature.

maps over time is likely to lead to measurement errors. The lack of a time- and space-consistent documentation of the networks prevents many researchers from utilizing the time series variations in empirical studies and restricts the possibility of structural evaluations.

To this end, our first contribution is to provide a standardized panel to document the evolution of the transportation networks in China from 1995 to 2016. We digitize a series of high-resolution paper maps from several publishers and extract four modes of transportation: national road, highway, railroad, and water-borne. To standardize the quality of the roads across time and space, we have carefully documented the changes in road design over several rounds of revision using the publications from the Ministry of Transportation in China. In each revision, we have also described how the road design varies across regions stipulated by the published engineering standards. We finally approximate the quality of each road by its “design speed”. Comparing the standard practices in the literature of categorizing the roads into “highways” and “national roads”, the design speed of a road is a more precise indicator of the road’s quality to reflect the variation over time and space. We also allow the speed of railways to vary to capture the quality change over time. To address the time-inconsistency issue introduced by map projections, we propose an iterative procedure to ensure that every pixel of the same road always has the same coordinates since its first entry into the dataset. As compared to the existing datasets, our method creates the first dataset of the transportation network in China that is time- and space-consistent. This dataset can be useful for both empirical and structural work in this literature.

Based on the new dataset, we show that the investments in transportation infrastructure have reduced the average travel time between cities by 72 percent. Moreover, improvements in connectivity are not distributed evenly. While well-connected cities in the initial year receive a modest improvement at around 55 percent, the initially remote cities can enjoy a reduction as high as 95 percent. The uneven improvement in connectivity is due to the asymmetry in the initial network layout, in which the coastal cities were already well-connected to each other. In contrast, the infrastructure in the inland regions was sparse. Under this circumstance, linking a remote city with the extensive network along the coast will reduce the travel time from an inland city to all the coastal cities in the existing network, therefore substantially reduce its average transport friction. On the other direction, the additional ac-

cess to the remote city barely affects the average connectivity of the coastal cities. However, one cannot simply equate the remarkable gain in connectivity in the remote cities to the reduction in spatial inequality for the reasons that we have discussed above. The improved connection with the large cities not only facilitates the goods flows, but also makes it easier for production factors to migrate away. In the end, the overall effects need to be evaluated through the lens of a general equilibrium model.

To structurally evaluate the impacts of the transportation network, we extend the trade models with migration decisions in Tombe and Zhu [2019] by introducing urban elements such as productivity agglomeration, amenity, and congestion; we also embed the model in a multi-period set up to evaluate the impacts over time. Both trade and migration flows are subject to bilateral frictions that, in turn, depend on the underlying infrastructure, as well as policy restrictions. In this setup, improvements in transportation networks affect welfare mainly through the two channels outlined above. In the first channel, better infrastructure lowers trade frictions, which in turn facilitates inter-city trade. The smaller cities tend to benefit more as the improved access to the markets in large cities significantly reduces their price index. In the second channel, better infrastructure also reduces the frictions of migration. In a world with agglomeration, the expansion of road networks attracts people from small cities to migrate to large cities. The same agglomeration forces also imply that the population's concentration in large cities will widen the existing productivity differences, leading to higher spatial inequality.

Intuitively, the overall impacts of transportation networks on spatial inequality in the model thus depend on four parameters: the agglomeration and congestion elasticity, as well as the elasticity of migration and trade with respect to the underlying infrastructure. The first two parameters together determine the *direction of migration flow* given an improvement in the infrastructure, as higher agglomeration and lower congestion forces attract workers into large cities. The elasticity of migration and trade, in turn, drive the *magnitude of the migration and trade flow*. Consider an extreme case where the agglomeration force is absent, better transportation network disperse the population distribution, and thus always lower the spatial disparity regardless of the migration and trade elasticities. However, in a world with a relatively strong agglomeration force, whether the transportation network reduces the

spatial disparity is a quantitative issue. This depends on whether goods or people are more mobile in response to improved road networks.

We pin down these parameters through several methods. To identify the trade and the migration elasticity to the underlying transportation networks, we rely on the structure equations obtained from the trade balance and migration decisions in our model. Since the placements of the road networks might be endogenous, we use a hypothetical network following [Faber \[2014\]](#) as an instrumental variable for the actual network. We find that inter-city trade is much more responsive than migration: the trade elasticity is around 0.89, and the migration elasticity is only around 0.34. The finding is not surprising: as transportation costs take up a large part of the inter-city trade costs, they are only responsible for a small proportion of the migration costs [[Morten and Oliveira, 2016](#)]. The barriers to migration, especially in the context of China, come from the policy restrictions, rather than the difficulties of travel.

To identify the agglomeration, congestion, and the other parameters of the model, we follow the methods of moments estimation from [Ma and Tang \[2020\]](#). Our moment conditions cover the magnitude of the inter-city trade as well as the city-specific population growth rates. The result shows that the agglomeration elasticity significantly dominates the congestion elasticity, and thus the improvements in infrastructure on average lead to population concentration. Our estimates are consistent with the basic pattern of migration flows in China: migrants flow from the small and inland cities into the large coastal ones.

The parameterized model enables us to evaluate the impacts of transportation networks by comparing two sets of simulations. In both simulations, we start with the initial population distribution in the year 1995 and solve the model forward year-by-year until 2016. In the “baseline” simulation, we use both the trade and migration costs matrices based on the actual transportation networks each year. In the other counterfactual simulation, we fix both the frictions to the initial year. Comparing the two sets of results reveals the impacts of transportation network expansion.

The expansion of transportation networks has significantly reduced spatial inequality. In the counterfactual world, without improvements in the infrastructure, the disparity in real wages between cities has gradually become widened over time. For example, the logarithm of

the standard deviation has increased by 1.68 percentage points, and the Herfindahl-Hirschman Index (HHI) and the Gini coefficient see similar magnitudes of increment. However, in the baseline simulation where the transportation networks expand as in the data, the spatial inequality in real wage declines by all three measures. The spatial inequality in the city-level output shows a similar pattern.

The reason that transportation networks acted as an equalizer is that inter-city trade is much more responsive to transportation network expansion than migration. To highlight this point, we run two auxiliary counterfactual simulations. In the first decomposition exercise, we assume that the improvements in transportation networks only reduce trade costs but not the migration costs; following the same logic, in the other exercise, the migration frictions decline over the years while the trade costs remain at their 1995 level. The simulations show that if better infrastructure only facilitates inter-city trade, then spatial inequality between the cities decline even further as compared to the case in which both costs decline. On the contrary, if the expansion of the traffic networks only facilitates the movements of people but not goods, spatial inequality will soar up.

Our paper contributes to the large literature on the economic impacts of infrastructure improvements [Baum-Snow, 2007; Banerjee et al., 2012; Faber, 2014; Redding and Turner, 2015; Donaldson and Hornbeck, 2016; Qin, 2016; Lin, 2017; Baum-Snow et al., 2018]. In the context of China, most of the research is empirical, and we are among the first to evaluate the impacts of transportation network in a general equilibrium setup. Many reduced-form research document a “tunnel effect”, in which smaller and peripheral cities tend to lose from road expansions due to factor mobility [Faber, 2014; Qin, 2016; Baum-Snow et al., 2018]. We complement this line of work by showing how spatial disparity depends on the tradeoff between goods and factor mobility in general equilibrium. The “tunnel effect” could be dominated by the positive impacts arising in the goods market due to a high trade elasticity. Our key observation that transportation network acts as an equalizer echos the insight from Banerjee et al. [2012], in which they argue that the distributional impacts of roads depend on the relative mobility of goods and capital in a highly stylized model. In addition, this paper’s contribution is the time-consistent panel dataset on the transportation networks of China that can be used by future empirical and structural studies.

The rest of the paper is organized as follows. Section 2 introduces the panel dataset on the transportation networks of China; Section 3 presents the model and Section 4 discusses the quantification strategy. Section 5 reports the results and Section 6 concludes.

## 2 The Transportation Networks of China

In this section, we outline the algorithm to digitize and estimate the transportation networks of China from 1995 to 2016. We first discuss the issues related to the digitization of the transportation atlas; we then proceed to reveal the underlying issues of standardizing the road networks across time and space. The estimation strategy of transportation costs based on the digitized maps is then offered. Lastly, we present the basic empirical patterns on the evolution of transportation networks over this period in China.

### 2.1 From Physical to Digital Maps

Our starting point is to collect the published transportation atlas of each year. In this exercise, the sources of the physical maps include libraries, book dealers, and map collectors. We restrict our attention to a national atlas with a scale of at least 1:9 million. The resolution for maps smaller than this scale is too low and not suitable for color identification once digitized. Table A.5 summarizes the basic information of the physical maps in our collection for reference.

The maps in our collection are not consistently scaled or projected due to the limited availability of maps. No single publisher provides a complete series of maps, and our collection ends up from several publishers. Unfortunately, the publishers differ in the scales and projection methods in their maps. For example, while the national publisher such as *Sino Maps* and *Di Zhi Publication* uses Albers projection with the reference point at 25N and 47E, some of the provincial publishers such as *Guangzhou Publishers* use other projection methods with different reference points. Moreover, measurement errors emerge among maps with identical projection parameters due to the noise arisen at the designing, printing, scanning, and color-identification stages. As a result of these noises, the same road might be drawn slightly differently across the maps, leading to inconsistency in geographic coordinates from

geo-referenced maps. Since our end goal is to evaluate the impacts of the network expansion, we must ensure that the existing infrastructure is represented consistently over time.

To overcome these challenges, we adopt an iterative process to construct the panel dataset, and our goal is to minimize the inconsistency of the infrastructure buildup across the maps. In the rest of the section, we outline the iterative method and provide more details in Appendix B. We start by digitizing the map in the initial year, which we denote as  $t_0$ . The dimension of the digitized map is 12669 pixels in width and 8829 pixels in height, which implies that each pixel roughly represents an area of 500 square meters in the real world. We then extract four modes of transportation by color identification: national road, highway, railway, and water. Table A.6 in the appendix provides the details on the correspondence between map legends and the transportation mode. Similar to many exercises in digitizing physical maps, the color-identified network contains gaps due to overlapping between roads and labels. To fill in these gaps, we use both standard image-processing algorithms, such as image dilation and contraction, and manual correction. The outcome of this exercise is four binary maps by the modes of transportation in the initial year  $t_0$ .

The map in all the following years is then constructed iteratively. For a given year  $t$  and mode  $m$ , we compare the scanned map in year  $t$  with the map in the previous  $t - 1$  of the same mode to construct the binary map in year  $t$ . We follow the general rules:

1. if a pixel is empty in year  $t - 1$  but not empty in year  $t$ , then it is classified as non-empty in year  $t$ . This case indicates new construction in year  $t$ .
2. if a pixel is empty in both year  $t - 1$  and year  $t$ , then it will be classified as empty in year  $t$ . This case corresponds to the continued existence of infrastructure.
3. if a pixel is non-empty in both year  $t - 1$  and year  $t$ , then it shall be classified as non-empty in year  $t$ . This case signals the continued absence of infrastructure.

The case omitted from the above rule is a non-empty pixel in year  $t - 1$  and becomes empty in year  $t$ . In theory, this case corresponds to destruction, demolition, or de-classification of an existing road, and thus, the pixel shall be classified as empty in year  $t$ . However, upon closer examination, we found these events very scarce. Most of these cases are actually

because the road is drawn with slightly different curvature between year  $t - 1$  and  $t$ . As a result, we manually correct these cases. If it is reasonable to believe that the differences are due to the variations in drawing, we re-classify all the relevant pixels in year  $t$  to their positions in year  $t - 1$  to ensure the consistent representation of existing infrastructure over the years.<sup>3</sup> Comparing to directly geo-referencing the maps of each year and stacking them to create the panel data, our approach, by design, ensures that every pixel of the same roads is always assigned the same coordinates over the years.

## 2.2 Standardizing Roads Across Time and Space

The outcome of the iterative methods outlined above is the binary maps for all the transportation modes in all the years. However, two roads constructed in different years are still hardly comparable in terms of capacity and quality due to the frequent changes in the standards of highway construction. Together with the rapid economic growth, the Chinese government had substantially upgraded the standards of highway engineering over the years. As a result, a “highway” constructed in 1988 might only be properly classified as a “second-rate road” by the 2014 standards. The inconsistency induced by such changes is widely considered as a major obstacle to utilize the time-series changes in transportation networks of China in empirical research.<sup>4</sup> Moreover, even if two roads are built at the same time, the quality might still differ significantly by region as stipulated by law. While the highways in the eastern flood plains are constructed to the highest standards allowing for 120 km/h travel, those in the mountainous regions are built with half of the design speed at 60 km/h. This inconsistency is mostly ignored in the literature, which often implicitly assume that the highways across the entire country are directly comparable in terms of traffic volume and speed. One of our contributions in this paper is to properly document the changes in highway engineering over time and region and then re-classify all the existing roads to a

---

<sup>3</sup>In the re-classification we also override the new construction to ensure that the same road is represented consistently across maps.

<sup>4</sup>For example, [Baum-Snow et al. \[2018\]](#) observes that “...however, the growth and improvement of China’s road network was so dramatic that roads that were important enough to merit inclusion on the 1990 map probably bear little resemblance to roads that meet this standard in 2010, even if both roads receive the same designation in the legend. Thus, we are reluctant to exploit the time-series variation in our measures of highways.”

single standard. In this section, we first provide a brief overview of the evolution of the highway engineering standards in China and then outline our procedure to standardize the roads across time and space. The details of the standardization procedure are provided in Appendix B.

The Ministry of Transportation of China publishes the standards of roads construction in the *Technical Standard of Highway Engineering*.<sup>5</sup> The standards had been revised nine times from 1951 to 2014, among which four revisions are relevant for our sampling period: 1988, 1997, 2003, and 2014. Each revision redefines a wide range of technical details of highway construction, such as the expected traffic volume, design speed, lane width, shoulder width, among many others. In the earlier years, all road designs depend on the terrain. For example, in the 1988 revision, the highways are constructed to allow for 120km/h design speed in the plains and low rolling hills (LRH), 100km/h in high hills, and 60 to 80km/h in the mountain regions. The quality of lower-tier roads also differ by terrain and allow lower speed limit.<sup>6</sup> Over time Ministry gradually relaxed the dependency on terrain for highway design mostly due to two factors: 1) the rugged terrain no longer restricts highway construction as China advances its engineering technology, and 2) the Ministry realized that a highway designed to a low capacity could hardly be upgraded to accommodate the rapidly increasing traffic volume. As a result, by the time the 2003 revision was drafted, all the newly constructed highways are required to have a design speed of 120 km/h, and the lower speed is only permitted in the areas with low expected traffic volume or high mountains.

For the first-rate roads, a term typically interchangeable with the more commonly used term “national roads”, the dependency on terrain is mostly maintained. In the plains and LRH, the design speed of the first-rate roads was always 100km/h throughout all four revisions, and in the hills and the mountains, the design speed was upgraded from 60 to 80km/h in the 2003 revision. To sum up, the evolving standards imply that the roads constructed in different years and regions are not comparable — a highway with 60 km/h design speed in

---

<sup>5</sup>Despite the word “highway” in the title, the standards regulate all inter-city roads, including the highways (Gao Su Gong Lu) and normal roads (Yi Ban Gong Lu) from the first-rate to the fourth-rate. “Highway” in the title is the official-translation of “Gong Lu”, which means inter-city roads. The “Urban roads” (Cheng Shi Dao Lu), the roads for intra-city transportation, are not regulated by these standards.

<sup>6</sup>For example, in 1988, the first-rate roads are designed for 100km/h in the plains, and 60km/h in the mountains, and for the second-rated roads, the design speeds are lowered to 80 and 40km/h respectively.

Table 1: Design Speed (km/h) of Roads by Standard Revision and Terrain

Revision	Highways			First-Rate Roads		
	Plains/LRH	Hills	Mountains	Plains/LRH	Hills	Mountains
1988	120	100	60/80	100	60	60
1997	120	120	60/80	100	60	60
2003	120	120	80	100	80	80
2014	120	120	80	100	80	80

Note: This table summarizes the design speed of the highways and the first-rate roads in various revisions of the *Technical Standard of Highway Engineering*. The “first-rate roads” is one tier below the highways in terms of design speed and quality. It is often considered interchangeable with the more common term, “national roads” (Guo Dao). The definition of the terrains is provided in the *Land Regulations in Highway Engineering*, also published by the Ministry of Transportation. “LRH” refers to “Low Rolling Hills”. See the details in appendix B.

1988 cannot accommodate the same traffic volume as a “highway” by the newer standards. Similarly, a road in the mountain regions is not comparable to a road in the plains either.

To standardize the roads across vintage and region, instead of categorizing the roads as “highways” and “national roads” commonly found in the literature, we delineate each road by its design speed. The design speed is an ideal choice in our context for two reasons. First and foremost, the design speed of a road varies by the time and the region of construction, which precisely allows us to compare the roads across these two dimensions. Moreover, most of the technical dimensions of highway engineering are highly correlated with the design speed. Therefore this variable is probably the best one-dimensional measure to indicate the quality of a road.<sup>7</sup> Finally, in this exercise, we assign each road pixel a value of design speed as described in Table 1, which is sourced from the four revisions of the *Technical Standard*. The definition of the terrains is based on the *Land Regulations in Highway Engineering*, also published by the Ministry of Transportation, together with the United Nations definition of terrains from Kapos et al. [2000]. In the appendix we explain in detail how the design speed change over time for both the highways and the first-rate roads, as well as the other technical details.

<sup>7</sup>For example, the highway standards require the roads to be built wider, less curvy, and with better material, in order to support a higher design speed and traffic volume.

## 2.3 From Digital Maps to Transportation Networks

In this section, we discuss the procedure to estimate trade and travel frictions based on the digitized infrastructure maps. We adapt the methods from the earlier work in [Allen and Arkolakis \[2014\]](#) and [Ma and Tang \[2020\]](#) to estimate the geographic costs of moving across space. We only provide the outline of the estimation, and refer the readers to [Ma and Tang \[2020\]](#) for more details.

The first step of the estimation converts the binary map into the mode-specific distance map between city centers. For each transportation mode  $m$ , we assign a cost value to all the pixels on the binary map. The cost parameter indicates the relative difficulty of traveling under the mode and is proportional to the speed at which the vehicle is allowed to travel. For example, the design speed is 120km/h for highways in the plain area, and 100km/h for the first-rate roads in 2014 as discussed in the previous part. Based on these, on national road and highway networks, crossing pixels with no road access have a normalized cost of 10. The cost is 2.5 and 3.75 for highway and national roads, respectively. The travel costs on waterway and railway networks are constructed similarly. In both modes, pixels with navigable waterways or railways have a normalized travel cost of 1, and all the other pixels have a cost of 10. Afterward, we use the Fast Marching Method (FMM) algorithm to create a distance matrix between all pairs of cities  $i$  and  $j$  for the year  $t$ , mode  $m$ , which we denote as  $d_{ijtm}$ .

Our second step aggregates the year-mode-specific distance matrix up to a year-specific distance matrix. At this stage, we differentiate between goods and passenger transportation, capturing the idea that even though both types of traffic flow rely on the same underlying transportation network, the cost of utilization might be different depending on the type of traffic. For this reason, we allow for separate cost parameters for goods,  $\{a_m^g, b_m^g\}$ , and passenger  $\{a_m^p, b_m^p\}$  transportation, which we will explain in detail later on. Subsequently, the aggregated transportation matrix will also be dependent on the traffic in question. Specifically, following [Ma and Tang \[2020\]](#), we assume that the reduced-forms relationship to

aggregate up  $d_{ijtm}$  across modes:

$$T_{ijt}^{\chi} = \frac{1}{\sum_m \exp(-a_m^{\chi} d_{ijtm} - b_m^{\chi})}, \chi = \{g, p\}.$$

In the above equation,  $T_{ijt}^{\chi}$  is the transportation cost between city  $i$  and  $j$  in year  $t$  for traffic type  $\chi$ , where  $\chi = g$  indicates goods, and  $\chi = p$  denotes passenger traffic.  $a_m^{\chi}$  is the variable cost, and  $b_m^{\chi}$  is the fixed cost of using mode  $m$ . The reduced-form relationship is the equilibrium condition from a discrete choice model of transportation as outlined in [Allen and Arkolakis \[2014\]](#), which we have abstracted away from the paper. Intuitively, the equation states that  $T_{ijt}^{\chi}$  is a weighted harmonic average of  $d_{ijtm}$ , where the weight is the inverse of the costs of using mode  $m$ . If  $a_m^{\chi}$  and  $b_m^{\chi}$  are high for mode  $m$ , then the total transportation cost summarized in  $T_{ijt}^{\chi}$  will be less reliant on  $d_{ijtm}$ . Also, note that the function form implies that both goods and people move along the same infrastructure as summarized by  $d_{ijtm}$ , and the transportation costs differ only because of the variable and fixed costs.

We next estimate the cost parameters  $\{a_m^{\chi}, b_m^{\chi}\}$  using the data on network traffic. The idea to identify the cost parameters is straightforward: if the cost parameters are high, then we shall expect less traffic on the mode in question. [Ma and Tang \[2020\]](#) showed that the traffic volume through city  $i$  by mode  $m$ , denoted as  $\nu_{imt}^{\chi}$ , equals

$$\nu_{imt}^{\chi} = \sum_{j=1}^J \exp(-a_m^{\chi} d_{ijtm} - b_m^{\chi}) + \sum_{j=1}^J \exp(-a_m^{\chi} d_{ijtm} - b_m^{\chi}).$$

Subsequently, the fraction of trade volume under mode  $m$  in city  $i$ ,  $s_{itm}^{\chi}$  is:

$$s_{itm}^{\chi} = \frac{\nu_{imt}^{\chi}}{\sum_{m'=1}^M \nu_{im't}^{\chi}}. \quad (1)$$

The data counter-part of  $s_{itm}^{\chi}$  comes from the *China City Statistics Yearbook 2005*. The yearbook reports the traffic volume through each city by modes of transportation in each year.<sup>8</sup> We take the average of  $s_{itm}^{\chi}$  across the years, and estimate  $\{a_m^{\chi}, b_m^{\chi}\}$  using a non-

---

<sup>8</sup>The yearbook reports the quantity shipped in metric tons in each city under transportation mode  $m$ . We use the estimated value-per-ton from [Ma and Tang \[2020\]](#) to convert the traffic in physical units to monetary terms.

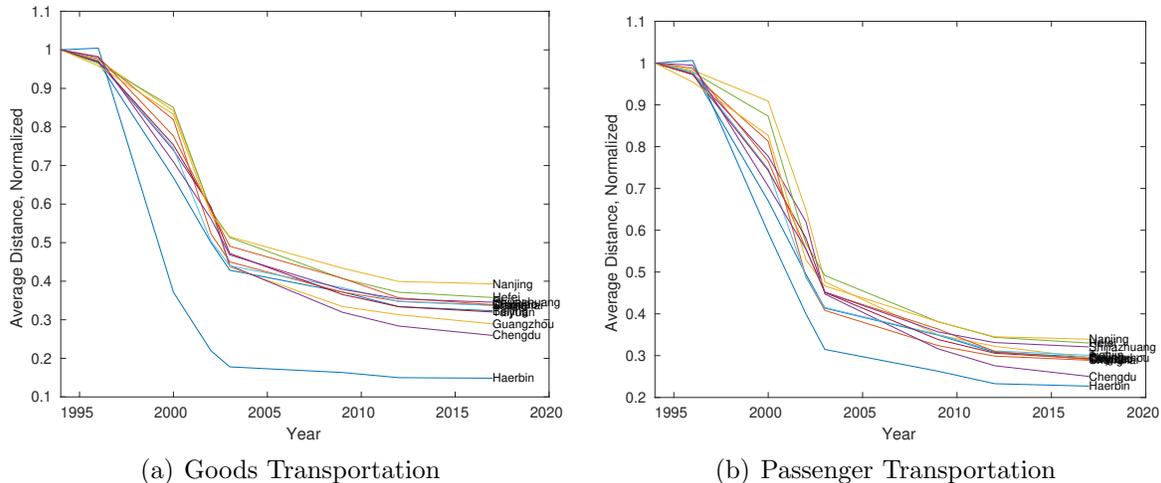


Figure 1: The Evolution of Transportation Networks

Note: The figures present the evolution of average transportation costs among a selection of origin cities. Each line represents the simple average of  $T_{ijt}^g$  from a given origin to all the other destination cities. The average transportation costs are normalized to 1 in the initial year.

linear least-squares routine to minimize the distance between  $s_{imt}^x$  and the data counterpart. Similar to Ma and Tang [2020], the estimated  $\{a_m^x, b_m^x\}$  can capture the main feature of the data: the vast majority of intercity traffic is carried out via road transportation, while railroads and water-borne transportation is relatively minor.

With the estimated  $\{a_m^x, b_m^x\}$ , we have finally reached the ultimate goal of the entire exercise—the transportation network matrix  $T_{ijt}^x$ . Before we turn to the quantitative model to evaluate the distributional impacts of the network expansion, we first take a look at the estimated  $T_{ijt}^x$  and understand several empirical patterns on the evolution of transportation networks in China.

## 2.4 The Evolution of Transportation Networks in China

The estimated  $T_{ijt}^x$  matrix offers a first peek into the evolution of transportation networks in China. Figure 1 to 3 presents the basic findings. Two basic empirical messages emerge.

The first message is that the expansion of transportation networks significantly reduced friction in both goods and passenger transportation. On average, the costs of transportation in both modes dropped by around 72 percent from 1995 to 2016. The majority of the reduction occurred during the earlier years before 2005, and during recent years, the changes

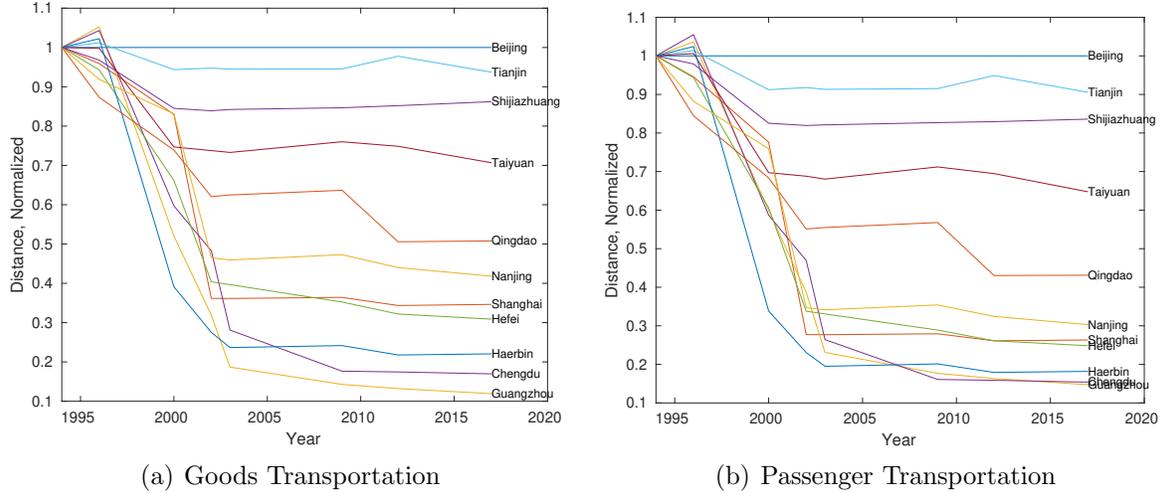


Figure 2: The Evolution of Transportation Networks, Distance from Beijing

Note: The figures present the transportation costs between the origin city and Beijing over time. The transportation costs are normalized to 1 in the initial year.

had been mostly gradual. The decreasing marginal returns in connectivity is intuitive. While the first road connecting two cities tend to result in a substantial reduction in travel frictions, the same might not be valid for the fifth road.

The second message that emerges from these graphs is that initially remote cities received a larger reduction in frictions. While the well-connected cities such as Beijing only enjoyed a reduction at around 67 percent, a remote city such as Wulumuqi could benefit from a reduction at as much as 96 percent. This pattern is true for both the average friction as shown in Figure 1, and even more pronounced in the pair-wise frictions as shown in Figure 2. Figure 3 plots the reduction in average transportation costs against the initial cost, and a clear negative relationship emerges in both goods and passenger transportation network. The uneven improvement in connectivity is due to the asymmetry in the initial network layout, in which the coastal cities were already well-connected to each other, whereas the infrastructure in the inland regions was sparse in 1995. Under this circumstance, linking a remote city into the extensive network along the coast will reduce the travel time of the inland city to all the coastal cities in the existing network, therefore substantially reduce its average transport friction. On the other direction, the additional access to the remote city barely affects the average connectivity of those coastal cities.

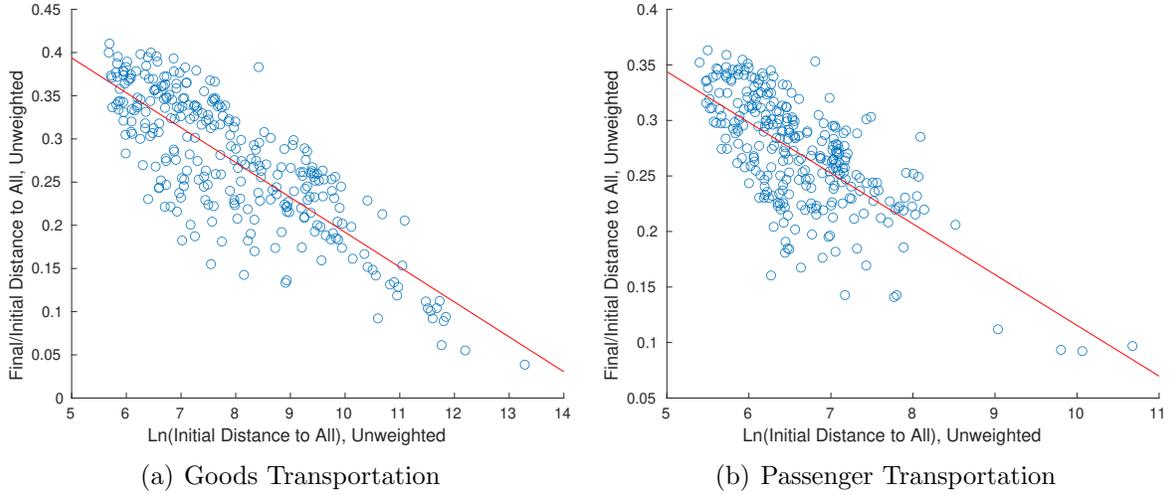


Figure 3: The Reduction in Transportation Costs v.s. Initial Costs

Note: The figures present the reduction in the average transportation costs against the initial costs. Again the average transportation cost is the simple average of  $T_{ijt}^g$  from a given origin to all the other destination cities.

Before moving onto the structural model, we emphasize that one cannot simply equate the relative gain in connectivity in the remote cities to a relative gain in the real wage, thus reducing spatial inequality. The improved connectivity might work as a double-edged sword. While the improved connection with the large cities facilitates the goods flows, it also makes it easier for production factors to migrate away. The overall effects of the transportation improvements need to be evaluated through a general equilibrium model, which we present in the next section.

### 3 The Model

Our model is based on [Eaton and Kortum \[2002\]](#). We add in the endogenous migration decisions similar to [Tombe and Zhu \[2019\]](#) and embed the model in a multi-period setting with urban elements.

#### 3.1 General Environment

The economy contains a mass  $\bar{L} > 0$  of individual workers, and  $J > 1$  geographically segmented cities, indexed by  $j = 1, 2, \dots, J$ . There is a  $(J + 1)$ th location that we use to

capture the rest of the world denoted as “ROW”. Individuals can migrate between the  $J$  cities subject to frictions, but they can not migrate between ROW and the  $J$  cities. Firms can trade between all the  $J + 1$  locations subject to variable trade costs. Time, indexed by  $t$  is discrete and infinite.

Individuals living in city  $j$  at time  $t$  obtain utilities according to:

$$U_{jt} = \phi_{jt} \cdot c_{jt}.$$

where  $c_{jt}$  is the consumption of a CES aggregation of intermediate goods indexed by  $\omega$  on the real interval  $[0, 1]$  with an elasticity of substitution denoted as  $\eta$ :

$$c_{jt} = \left( \int_0^1 (q_{jt}(\omega))^{\frac{\eta-1}{\eta}} d\omega \right)^{\frac{\eta}{\eta-1}}. \quad (2)$$

Individual utility also depends on the location-specific amenities denoted as  $\phi_{jt}$ , which in turn depends on an exogenously fixed component  $\bar{\phi}_j$ , as well as the population size of the city at period  $t$ ,  $L_{jt}$ :

$$\phi_{jt} = \bar{\phi}_j \cdot (L_{jt})^\alpha. \quad (3)$$

## 3.2 Production and Trade

The production side of the economy follows [Eaton and Kortum \[2002\]](#): market structure is perfectly competitive and every city is able to produce every variety  $\omega \in [0, 1]$ . The production function for variety  $\omega$  in city  $j$  at time  $t$  takes the form:

$$q_{jt}(\omega) = A_{jt} \cdot z_j(\omega) \cdot \ell_{jt},$$

where  $\ell_{jt}$  is the labor input.  $A_{jt}$  is the city-specific productivity that depends on an exogenous component,  $\bar{A}_j$ , as well as a time-varying part that is a function of the population size in

the city to capture the idea of agglomeration:

$$A_{jt} = \bar{A}_j \cdot (L_{jt})^\beta. \quad (4)$$

The variable  $z_j(\omega)$  is the city-variety specific productivity which is drawn from an *i.i.d* Frechet distribution with the parameter  $\theta$ :

$$F(z) = \exp(-z^{-\theta}).$$

Trade is subject to the standard iceberg costs: in order for 1 unit of product to arrive in city  $i$  from city  $j$ ,  $\tau_{ijt} > 1$  units of goods need to be produced and shipped. We assume  $\tau_{iit} = 1$  for all  $i$  and the standard triangular inequality  $\tau_{kit}\tau_{ijt} \geq \tau_{kjt}$  for all  $i, j$  and  $k$ . The trade costs in turn depend on the time-varying transportation network that we will specify later.

Denote the price of variety  $\omega$  from the sellers in city  $j$  at the market in city  $i$  as  $p_{ijt}(\omega)$ , the consumers in city  $i$  buy from the supplier with the lowest price:

$$p_{it}(\omega) = \min_{j=1, \dots, J+1} \{p_{ijt}(\omega)\}.$$

**International Trade** The cities can also trade with the ROW. We assume there are only  $H$  port cities (indexed as  $h = 1, 2, \dots, H$ ) in the economy which can directly trade with the ROW. The non-port cities have to first connect to the nearest port in order trade with the ROW. This trading structure allows us to classify all the cities into the mutually-exclusive sets  $\{H, J_{1t}, J_{2t}, \dots, J_{Ht}\}$ , where  $H$  is the set of the port cities, and  $J_h$  is the set of the non-port cities that use port  $h$  for to trade in year  $t$ :

$$J_{ht} = \{j \in J \setminus H : h = \operatorname{argmin}_{x \in H} (\tau_{xjt})\}.$$

### 3.3 Migration

Individuals decide where to migrate at the beginning of period  $t$  to maximize the current period utility. Denote  $V_{jt}$  to be the indirect utility from living in city  $j$  at time  $t$ :

$$V_{jt} = \phi_{jt} \cdot \frac{w_{jt}}{P_{jt}}, \quad (5)$$

where  $w_{jt}$  is the nominal wage rate, and  $P_{jt}$  is the ideal price index in city  $j$  at period  $t$ :

$$P_{jt} = \left( \int_0^1 (p_{jt}(\omega))^{1-\eta} d\omega \right)^{\frac{1}{1-\eta}}.$$

In addition to the indirect utility, each worker also draws an idiosyncratic preference shock toward each city  $\{e_{jt}\}_{j=1}^J$ , where  $e_{jt}$  is *i.i.d* across location, time, and individual. We assume that  $e_{jt}$  follows a Frechet distribution with the CDF:

$$F(e_{jt}) = \exp \left[ - (e_{jt})^{-\kappa} \right],$$

where  $\kappa$  is the shape parameter. Lastly, moving from city  $j$  to  $i$  also incurs an origin-destination specific cost, denoted as  $\lambda_{ijt} \geq 1$ , that acts as a discount on utility. If a worker moves from city  $j$  to city  $i$ , the utility she receives in the end is:

$$\frac{V_{it} \cdot e_{it}}{\lambda_{ijt}}.$$

Similar to the iceberg trade cost, we assume  $\lambda_{iit} = 1$  and  $\lambda_{kit}\lambda_{ijt} \geq \lambda_{kjt}$  for all  $i, j$  and  $k$ . The costs of migration include the financial costs of moving over a long distance and thus depends on the underlying transportation networks. In addition, it also captures various policy barriers that deter migration such as Hukou, working permits, and other bureaucratic red tapes.

Taking all the three elements discussed above into consideration, a worker living in city  $j$  will migrate to city  $i$  at period  $t$  if and only if living in city  $i$  provides him with the highest

utility among all the  $J$  cities:

$$\frac{V_{it} \cdot e_{it}}{\lambda_{ijt}} \geq \frac{V_{kt} \cdot e_{kt}}{\lambda_{kjt}}, \forall k = 1, 2, \dots, J.$$

### 3.4 The Equilibrium

At the beginning of period  $t$ , individuals take the initial population distribution from the previous period,  $\{L_{j,t-1}\}_{j=1}^J$ , as given.  $\tau_{ijt}$  and  $\lambda_{ijt}$  matrices are also public information based on the current state of transportation infrastructure. The equilibrium at period  $t$  is then defined as a vector of prices  $\{w_{jt}, p_{jt}(\omega)\}$ , a vector of quantities  $\{q_{jt}(\omega), c_{jt}, L_{jt}\}$  such that:

- Every individual maximizes his utility by choosing the location and the consumption bundle in period  $t$ .
- Every firm maximizes its profit.
- Labor market clears in each location.
- Trade balance.

### 3.5 Model Solution

**Conditional on a population distribution**  $\{L_{jt}\}$ , the solution of the model in period  $t$  is similar to a standard Eaton-Kortum model. Perfect competition implies that the price charged by the sellers from city  $j$  in city  $i$  for variety  $\omega$  is:

$$p_{ijt}(\omega) = \frac{\tau_{ijt} w_{jt}}{z_j(\omega) A_{jt}}.$$

The price paid for a particular variety  $\omega$  in city  $i$  is given by the minimum:

$$p_i(\omega) = \min_j \left\{ \frac{\tau_{ijt} w_j}{z_j(\omega) A_{jt}} \right\}.$$

Given the Frechet distribution, the price index in city  $i$  is:

$$\begin{aligned} P_{it} &= \left( \int_0^\infty p^{1-\eta} dG_i(p) \right)^{\frac{1}{1-\eta}}, \\ &= \Psi \left( \sum_{j=1}^J (w_{jt}\tau_{ijt})^{-\theta} (A_{jt})^\theta \right)^{-1/\theta}, \end{aligned} \quad (6)$$

where  $G_i(p)$  is the CDF of prices in city  $i$ , and  $\Psi$  is the Gamma function evaluated at  $1 + (1 - \eta/\theta)$ .

The share of total expenditure in city  $i$  on the goods from city  $j$  is thus:

$$\pi_{ijt} = \frac{(w_{jt}\tau_{ijt})^{-\theta} (A_{jt})^\theta}{\sum_{k=1}^J (w_{kt}\tau_{ikt})^{-\theta} (A_{kt})^\theta}.$$

The expression for the bilateral trade flow from  $j$  to  $i$  is thus:

$$X_{ijt} = X_{it}\pi_{ijt} = w_{it}L_{it} \frac{(w_{jt}\tau_{ijt})^{-\theta} (A_{jt})^\theta}{\sum_{k=1}^J (w_{kt}\tau_{ikt})^{-\theta} (A_{kt})^\theta}. \quad (7)$$

From above it is straightforward to solve for the equilibrium real wage in city  $j$  at time  $t$ .

Given the model solution conditional on  $L_{jt}$ , we can now solve for the migration decision. Since  $V_{jt} = \phi_{jt}w_{jt}/P_{jt}$ , the distribution of the idiosyncratic preference shock implies that the fraction of the population that moves from city  $j$  to city  $i$  in period  $t$  is:

$$m_{ijt} = \frac{(V_{it})^\kappa (\lambda_{ijt})^{-\kappa}}{\sum_{m=1}^J (V_{mt})^\kappa (\lambda_{mjt})^{-\kappa}}.$$

Note that in the above expression,  $\kappa$ , the parameter in the idiosyncratic shock, is also the migration elasticity with respect to migration barrier. Denote the denominator in the above equation (raised to a power of  $1/\kappa$ ) as  $\Pi_{jt}$ :

$$\Pi_{jt} = \left[ \sum_{m=1}^J (V_{mt})^\kappa (\lambda_{mjt})^{-\kappa} \right]^{\frac{1}{\kappa}}. \quad (8)$$

The  $\Pi_{jt}$  term, intuitively, is the expected utility for a worker originally in location  $j$  at time

*t*. The migration flow from city *j* into city *i* is:

$$L_{ijt} = (V_{it})^\kappa (\lambda_{ijt})^{-\kappa} (\Pi_{jt})^{-\kappa} L_{j,t-1}. \quad (9)$$

Total population size in city *i* is thus solved as:

$$L_{it} = \sum_{j=1}^J L_{ijt} = (V_{it})^\kappa \sum_{j=1}^J \frac{(\lambda_{ijt})^{-\kappa}}{(\Pi_{jt})^\kappa} \times (L_{j,t-1}). \quad (10)$$

## 4 Quantification

We quantify the model to 291 prefecture-level cities in China, plus one additional location representing the rest of the world (ROW). Goods and people can move between the Chinese cities with frictions, but migration between China and the ROW is not allowed.

Our sample contains the largest number of cities from which we can observe both the population and economic output from 1995-2016. One particular challenge of working with the prefecture-level administrative region in China is that cities' boundary and definition have undergone substantial changes during the sample period. For example, many new prefecture cities are created by re-classifying and re-grouping emerging towns from rural prefectures (Di Qu); meanwhile, sub-prefecture administrative regions (Xian) might be re-assigned to different cities. To address these issues, we re-align the city boundaries and construct a geographically consistent panel of population and output based on the city boundaries in the year 2016 following the methods detailed in [Ma and Tang \[2020\]](#). In short, we document the change of boundaries in each year based on the official decrees and re-construct the population and output for each city by aggregating the county-level data of each year.

In the rest of this section, we briefly outline the estimation procedure for all the parameters in the model.

### 4.1 Trade Elasticity

We have estimated both goods and passenger transportation networks,  $T_t^g$  and  $T_t^p$ , of each year in Section 2. Based on the structural equations obtained from the model section, we

proceed first to estimate the trade elasticity.

Our starting point is to assume that the iceberg trade cost is a function of the transportation networks between the origin and destination city. Let

$$\tau_{ijt} = \begin{cases} \bar{\tau} \cdot (T_{ijt}^g)^\psi, & i \neq j \\ 1, & i = j \end{cases}$$

, where  $T_{ijt}^g$  is the goods transportation network between the cities. We then note that the export volume from a non-port city  $j \in J_h$  to the ROW can be written as:

$$\begin{aligned} \log(X_{ROW,j,t}) &= \log(X_{ROW,t}) - \log \left[ \sum_{k=1}^J (w_k^t \tau_{ROW,k,t})^{-\theta} (A_{kt})^\theta \right] \\ &\quad - \theta \log(w_{jt}) + \theta \log(A_{jt}) - \theta\psi \log(T_{ROW,h,t}^g) - \theta\psi \log(T_{hjt}^g). \end{aligned}$$

Adding in the origin-city fixed effects absorbs the costs and productivity of city  $j$ , and port-city fixed effects absorb the distance between port  $h$  and the ROW. The only variation that remains after these fixed effects is the distance between city  $j$  and the port city  $h$ , leading to our estimation equation:

$$\log(X_{ROW,j,t}) = \delta_0 + \delta_{jt} + \delta_{ht} - \theta\psi \log(T_{hjt}^g) \quad (11)$$

Estimating the above equation using OLS might lead to bias as the placement of the transportation infrastructure is potentially endogenous to population, productivity, and other factors that also drive the city-level exports. To alleviate the issue of endogeneity, we introduce an instrument variable for  $T_{hjt}^g$  following the hypothetical network methods in [Faber \[2014\]](#). In short, we start by approximating the cost of construction on every pixel of our map using the elevation, slope, and terrain cover data. Based on the construction costs, we construct the minimum spanning tree (MST) connecting a certain number of cities. We build the MST separately for each mode of transportation and then aggregate to a hypothetical  $T_{ijt}^g$  matrix from the MST networks following the same step as outlined in Section 2. Finally, we use the MST-based  $T_{ijt}^g$  matrix as the instrumental variable for the actual transportation network. The MST networks are functions of geographical characteristics of China, which

eliminates the possibility of reverse-causality. The exclusion restriction assumes that the difficulty of trespassing a pixel based on terrain characteristics will only affect city-level trade flows through the availability of transportation infrastructure on that pixel.

The estimation above leads to an estimate of the composite parameters  $\theta\psi$  of 0.89, with an standard error of 0.03. Note that based on the estimate of  $\theta\psi$ , we can already construct the iceberg trade costs matrix up to an exponent of  $\theta$  and a scale  $\bar{\tau}^\theta$ :  $(\tau_{ijt})^\theta = \bar{\tau}^\theta (T_{ijt}^g)^{\theta\psi}$ . We will show in the later part of the section that  $(\tau_{ijt})^\theta$  allows us to back out the city-level amenity up to a scale.

## 4.2 Migration Elasticity

We estimate the migration elasticity with respect to transportation networks following a similar manner. We start by modeling the migration friction as a function of the underlying infrastructure and the entry barriers:

$$\lambda_{ijt} = \bar{\lambda} \cdot (T_{ijt}^p)^\xi,$$

where  $\bar{\lambda}$  governs the overall magnitude of the matrix, and  $T_{ijt}^p$  is the passenger transportation network.

The migration flow from equation (9) leads to the following estimation equation:

$$\begin{aligned} \log(L_{jit}) &= \underbrace{\kappa \log(V_{jt})}_{\delta_{jt}} + \underbrace{\log(L_{it}) - \kappa \log(\Pi_{it})}_{\delta_{it}} - \kappa\xi \log(T_{jit}^p) \\ &= \delta_{jt} + \delta_{it} - \kappa\xi \log(T_{jit}^p) + \mu_{jit}, \end{aligned}$$

where  $\delta_{jt}$  and  $\delta_{it}$  absorbs the destination- and origin-fixed effects. Note that the impacts of destination-specific entry barriers, such as the Hukou system, are eliminated by the fixed effects and thus do not bias our estimation of  $\kappa\xi$ . We use the same MST-based instruments as in the previous section to alleviate the concerns for endogenous placements of the transportation networks. In the end, the IV-based regression estimates that  $\kappa\xi$  to be 0.38, with a standard error of 0.005. Similar to the trade elasticity, the estimate of the compound

parameter  $\kappa\xi$ , together with the fixed effects  $\delta_{it}$  from the same regression, allow for the construction of the migration cost matrix raised to a power  $(\lambda_{ijt})^\kappa = (\bar{\lambda}\delta_{it})^\kappa (T_{ijt}^p)^{\kappa\xi}$ .

### 4.3 The Productivity and Amenity

In this section, we outline the estimation of  $\bar{A}_j$ ,  $\bar{\phi}_j$ , the time-invariant components in city-specific productivity and amenity as specified in equation (4) and (3). To estimate  $\bar{A}_j$ , we extend the methods as detailed in [Ma and Tang \[2020\]](#), which backs out the fixed component of productivity as a residual wage net of agglomeration and market access captured in  $L_{jt}$  and the  $\tau_{ijt}$  matrix. We apply this method to our sample of 291 cities, and use the year 1995, one year prior to our initial year in simulation to estimate  $\bar{A}_j$ .

To estimate  $\bar{\theta}_j$ , we turn to the definition of the indirect utility of living in city  $j$ ,  $V_{jt}$ , as in equation (5). Raising both sides of the equation to the power  $\kappa$ , and then taking logs, we have:

$$\log [(V_{jt})^\kappa] = \kappa \log (w_{jt}) - \kappa \frac{1}{\theta} \log \left[ (P_{jt})^\theta \right] + \alpha \kappa \log (L_{jt}) + \kappa \log (\bar{\phi}_{jt}). \quad (12)$$

In the above equation, while  $w_{jt}$  and  $L_{jt}$  are observable in the data,  $(V_{jt})^\kappa$  and  $(P_{jt})^\theta$  are model-constructs without a counter-part in the data. To solve for these two parts, we apply the methods outlined in [Allen and Donaldson \[2018\]](#) to our context, using  $(\tau_{ijt})^\theta$  and  $(\lambda_{ijt})^\kappa$  that we have backed out from the previous two steps.

**The indirect utility,  $(V_{jt})^\kappa$**  To back-out  $(V_{jt})^\kappa$ , we turn to the **population movement conditions**. The starting point is the equilibrium condition for population movement in period  $t$  as defined in equation (9), which dictates that the population in city  $i$  shall equal to the population flows into the city from all the  $J$  locations:

$$L_{it} = \sum_{j=1}^J L_{ijt} = \sum_{j=1}^J \left( \frac{V_{it}}{\lambda_{ijt}} \right)^\kappa (\Pi_{jt})^{-\kappa} L_{j,t-1},$$

where  $\Pi_{jt}$  is the option value of living in city  $j$  as defined in equation (8). Rearranging the equation, we have:

$$\frac{1}{(V_{it})^\kappa} = \sum_{j=1}^J \frac{(\lambda_{ijt})^{-\kappa}}{\sum_{m=1}^J (V_{mt})^\kappa (\lambda_{mjt})^{-\kappa}} \frac{L_{j,t-1}}{L_{it}}. \quad (13)$$

Conditional on the estimated  $(\lambda_{ijt})^\kappa$  from the previous part and the population distribution in period  $t-1$  and  $t$ , one can back out  $(V_{it})^\kappa$  by solving equation (13) as a system of non-linear equations. Note that as  $(V_{it})^\kappa$  is the unknown variable, one can solve equation (13) without knowing the value of  $\kappa$ .

**The price index,  $(P_{jt})^\theta$**  Similarly, we can also solve for  $(P_{jt})^\theta$  from the zero profit condition and the trade balance equations. Our starting point is the **zero profit condition**, which implies that the total cost of production in a city,  $w_{jt}L_{jt}$  should equal to sum of sales to all the  $J+1$  locations :

$$\begin{aligned} w_{jt}L_{jt} &= \sum_{i=1}^{J+1} X_{ijt} = \sum_{i=1}^{J+1} w_{it}L_{it} \frac{(w_{jt}\tau_{ijt})^{-\theta} (A_{jt})^\theta}{\sum_{k=1}^{J+1} (w_{kt}\tau_{ikt})^{-\theta} (A_{kt})^\theta} \\ &= \sum_{i=1}^{J+1} w_{it}L_{it} (w_{jt}\tau_{ijt})^{-\theta} (A_{jt})^\theta (P_{it})^\theta \Psi^{-\theta}. \end{aligned} \quad (14)$$

Recall that  $V_{it} = \phi_{it}w_{it}/P_{it}$ , and therefore  $(P_{it})^\theta = (\phi_{it})^\theta (w_{it})^\theta (V_{it})^{-\theta}$ . With this insight, we can further rewrite the zero profit condition as:

$$w_{jt}L_{jt} = \sum_{i=1}^{J+1} (w_{it})^{1+\theta} \left[ (\bar{A}_{jt})^\theta (\bar{\phi}_{it})^\theta \Psi^{-\theta} \right] L_{it} (w_{jt}\tau_{ijt})^{-\theta} (L_{jt})^{\beta\theta} (V_{it})^{-\theta} (L_{it})^{\alpha\theta}.$$

If we define

$$K_{jit} = (\bar{A}_{jt})^\theta (\bar{\phi}_i)^\theta \Psi^{-\theta},$$

the above can be further collected into:

$$(w_{jt})^{1+\theta} L_{jt} = \sum_{i=1}^{J+1} (w_{it})^{1+\theta} K_{jit} (\tau_{ijt})^{-\theta} (V_{it})^{-\theta} (L_{it})^{1+\alpha\theta} (L_{jt})^{\beta\theta}. \quad (15)$$

We also define  $Y_{jt} = w_{jt}L_{jt}$ , and the zero profit condition from equation (14) can be expressed as:

$$(w_{jt})^\theta = \sum_{i=1}^{J+1} \frac{Y_{it}}{Y_{jt}} (\tau_{ijt})^{-\theta} (A_{jt})^\theta (P_{it})^\theta \Psi^{-\theta}.$$

Let

$$\Omega_{jt} = \Psi \frac{w_{jt}}{A_{jt}} \quad (16)$$

denote the marginal cost of production up to a constant  $\Psi$ , from above we have the equilibrium equation from the zero-profit condition:

$$(\Omega_{jt})^\theta = \sum_{i=1}^{J+1} \frac{Y_{it}}{Y_{jt}} (\tau_{ijt})^{-\theta} (P_{it})^\theta. \quad (17)$$

Meanwhile, the **trade balance condition** also implies that the total imports of city  $j$  must also equal to its total exports, which in turn is the same as the total income,  $w_{jt}L_{jt}$ . Similar to before, substituting in the solution of the price index as in equation (6) and the definition of  $V_{jt}$ , we have:

$$w_{jt}L_{jt} = \sum_{i=1}^{J+1} (w_{jt})^{1+\theta} L_{jt} (w_{it}\tau_{jit})^{-\theta} \left[ \bar{A}_i (L_{it})^\beta \right]^\theta (V_{jt})^{-\theta} \left[ \bar{\phi}_j (L_{jt})^\alpha \right]^\theta \Psi^{-\theta}. \quad (18)$$

That is,

$$(w_{jt})^{-\theta} (V_{jt})^\theta = \sum_{i=1}^{J+1} K_{ijt} (w_{it}\tau_{jit})^{-\theta} (L_{it})^{\beta\theta} (L_{jt})^{\alpha\theta}. \quad (19)$$

We can also directly express the trade balance condition as:

$$\begin{aligned} w_{jt}L_{jt} &= \sum_{i=1}^{J+1} w_{jt}L_{jt} (w_{it}\tau_{jit})^{-\theta} (A_{it})^{\theta} (P_{jt})^{\theta} \Psi^{-\theta}, \\ (P_{jt})^{-\theta} &= \sum_{i=1}^{J+1} (w_{it}\tau_{jit})^{-\theta} (A_{it})^{\theta} \Psi^{-\theta}. \end{aligned}$$

Similarly, using the notation of  $\Omega_{it} = \Psi w_{it}/A_{it}$ , from above we have:

$$(P_{jt})^{-\theta} = \sum_{i=1}^{J+1} (\tau_{jit})^{-\theta} (\Omega_{it})^{-\theta}. \quad (20)$$

The above equation, together with equation (17), defines a system of equations with  $(P_{jt})^{\theta}$  and  $(\Omega_{jt})^{\theta}$  as unknowns, conditional on the  $(\tau_{jit})^{\theta}$  estimated from the previous parts and  $Y_{jt}$  taken from the data. Similar to the trick in solving equation (13), one can back-out  $(P_{jt})^{\theta}$  without specifying the value of  $\theta$ .

With  $(V_{jt})^{\kappa}$  and  $(P_{jt})^{\theta}$  from the previous two steps, it is straightforward to back-out  $\bar{\phi}_j$  up to the power of  $\kappa$  from equation (12). We use  $t = 1995$ , one year before our baseline simulation to compute the “raw” amenity as

$$\log(\bar{\phi}_j)^{\kappa} = \log[(V_{jt})^{\kappa}] - \kappa \log(w_{jt}) - \kappa \frac{1}{\theta} \log[(P_{jt})^{\theta}] + \alpha \kappa \log(L_{jt}).$$

However, directly using the amenity from the above equation might be problematic due to the potential endogeneity embodied in the initial wage and population data. To alleviate this bias, we further project  $\log(\bar{\phi}_j)^{\kappa}$  onto an array of time-invariant geographic characteristics of each city,  $z_j$ , using OLS. The  $z_j$  vector includes the elevation, terrain slope, precipitation, and the distance to coast to capture the quality of living in a location. Denote the parameter vector of the projection as  $\gamma_{\phi}$ , we then use the predicted value,  $\log(\bar{\phi}_j)^{\kappa} = \gamma_{\phi} z_j$ , as our final estimate of the fixed component of amenity, up to an exponent of  $\kappa$ .

## 4.4 All the other parameters

At this stage, we have identified all the parameters of the model up to a scale except for  $\Theta = \{\bar{\tau}, \alpha, \beta, \eta, \theta, \kappa, \bar{\lambda}\}$ . We discipline these parameters by matching a vector of moments between the model and the data. We denote the moment conditions in the data as the vector  $\bar{\mathcal{S}}$ , and the counter-parts in the model as  $\mathcal{S}(\Theta)$ .

The first element in the  $\mathcal{S}$  vector is the aggregate magnitude of inter-city trade in China. Following [Ma and Tang \[2020\]](#), the data moment comes from the *Investment Climate Survey* in China from the [World Bank \[2005\]](#). Based on the firm-level survey, around 62.5 percent of the total shipments were destined towards the consumers outside of the firm’s home city. Accordingly, we compute the internal-trade-to-GDP ratio in our model in the year 2005 as the counter-part. The magnitude of internal trade identifies  $\bar{\tau}$ , the scale parameter in the iceberg trade costs matrix.

The rest of the vector  $\mathcal{S}$  consists of the population growth rates of all the 291 Chinese cities between the years 1995 and 2016. The long-term trend in the population movement in the data is driven by the changes in migration and trade friction, as well as agglomeration and congestion forces. Similarly, in the model, the population growth rates are jointly determined by the agglomeration forces ( $\alpha$  and  $\beta$ ), as well as the migration elasticity ( $\kappa$ ) and the cost of moving  $\bar{\lambda}$ . To this end, the long-term population growth rates can pin down the rest of the  $\Theta$  vector that is broadly related to population movements.

In the end, our estimation strategy is to back-out  $\Theta$  to minimize the Euclidean distance between the data and the model moments:

$$\min_{\Theta} [\bar{\mathcal{S}} - \mathcal{S}(\Theta)] \mathcal{W} [\bar{\mathcal{S}} - \mathcal{S}(\Theta)]',$$

where  $\mathcal{W}$  is the weighting vector. We weight the population growth rates using the city-level GDP in the initial year 1995, and then set the weight to the trade volume moment to be the average of the weight assigned to the other 291 population-based moments. The weighting vector prioritizes population growth rates in the largest cities; we prefer the weighting because impacts of agglomeration, congestion, and migration frictions are best identified in the largest cities. We use the iterative particle swarm optimization algorithm developed in

Ma and Tang [2020] to solve the minimization problem and use Gauss-Newton Regression (GNR) to estimate the standard errors.

Table 2: Parameters, Estimated

name	estimate	s.e.
$\alpha$	-0.0151	0.0560
$\beta$	0.3345	0.0363
$\theta$	4.0044	0.5571
$\kappa$	1.6291	0.1733
$\bar{\tau}$	3.3312	1.7307
$\bar{\lambda}/1000$	27.8503	6.0394

Note: This table reports the estimation results. The standard errors are computed using Levenberg-Marquardt.  $\alpha$  and  $\beta$  is the congestion and agglomeration elasticity, respectively.  $\theta$  and  $\kappa$  is the trade and migration elasticity, respectively.  $\bar{\tau}$  is the scale parameter of trade costs, and  $\bar{\lambda}$  is the scale parameter of migration costs.

Table 2 presents the estimation results. Most of the parameters are pinned down with relatively low standard errors, except for  $\alpha$ , the congestion disutility parameter. The reason for the weak identification is that for population movement, what matters is  $\alpha - \beta$ , the difference between congestion and agglomeration, but not the two parameters separately. Nevertheless, our estimation procedure can pin down the said difference at around 0.33, and shows that between 1995 and 2016, the agglomeration forces far dominate the congestion, leading to “net agglomeration”. We estimate that the trade elasticity to be around 4.0, which is well within the range of the estimates commonly seen in the trade literature.<sup>9</sup> Our migration elasticity is lower than the trade elasticity at around 1.62, again suggesting the population flow is much less sensitive to changes in migration costs than good flow. The estimates from the structural estimation are consistent from the instrument-variable based estimates reported earlier.

## 5 Quantitative Results

In this section, we present the quantitative results based on the estimated parameters from the previous part. We start by assessing the welfare and the distributional impacts of the

<sup>9</sup>For example, Simonovska and Waugh [2014] estimated the trade elasticity to be between 2.79 and 4.46; similarly Bernard et al. [2003] and Eaton et al. [2011] suggested a range between 3.6 and 4.8 using firm-level data.

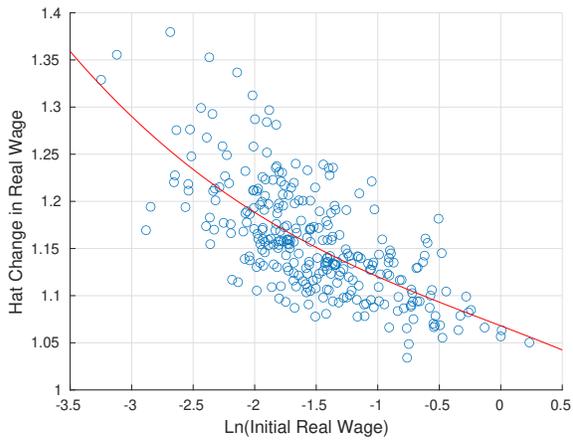
transportation network improvement over the entire period; we then decompose the impact into trade v.s. migration channels to highlight the mechanisms through which transportation networks affect welfare and spatial inequality.

## 5.1 Baseline Results

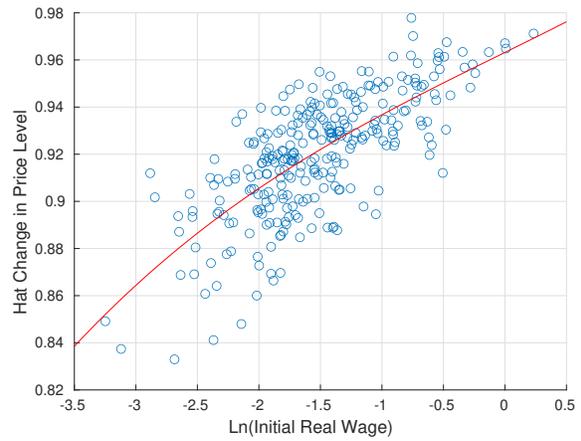
We evaluate the impacts of transportation networks by comparing two sets of simulations. In both simulations, we start with the initial population distribution in the year 1995 and solve the model forward year-by-year until 2016. In the “baseline” simulation, we use the  $\tau_{ijt}$  and  $\lambda_{ijt}$  matrices that are based on the actual transportation networks in each year that we have estimated in Section 2. On the contrary, in the “no change” counterfactual, we fix both the  $\tau_{ijt}$  and the  $\lambda_{ijt}$  matrices to their levels in the initial year. Comparing the two sets of results reveals the impacts of the expansion of transportation networks. The results are summarized in Table 3 and Figure 4.

The expansion of transportation networks over the sample period has significantly reduced spatial inequality. In the counterfactual world, without improvements in infrastructure, the disparity in real wages between cities has gradually become widened over time. As shown in the first column of Table 3, the logarithm of the standard deviation has increased by 1.68 percentage points, and the HerfindahlHirschman Index (HHI) and the Gini coefficient see similar magnitudes of increment. However, in the baseline simulation where the transportation networks expand as in the real world, the spatial inequality in real wage declines as measured by all three measures reported in the same table. The spatial inequality in the city-level output shows a similar pattern, as shown in the second and the fourth column of the table. Although all the inequality measures become widened in both simulations, the spatial disparity grows at a significantly lower speed in the “baseline” than in the “no change” counterfactual.

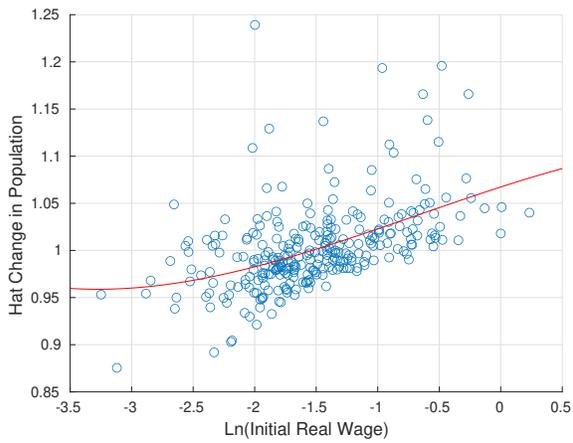
In theory, transportation networks bring about two counter-acting forces that could shape the spatial inequality in either direction. On the one hand, lower trade friction facilitates inter-city trade and reduces spatial inequality. The smaller cities tend to benefit more from trade liberalization as the improved market access to the large cities significantly reduces the price index. On the other hand, better infrastructure also reduces migration frictions,



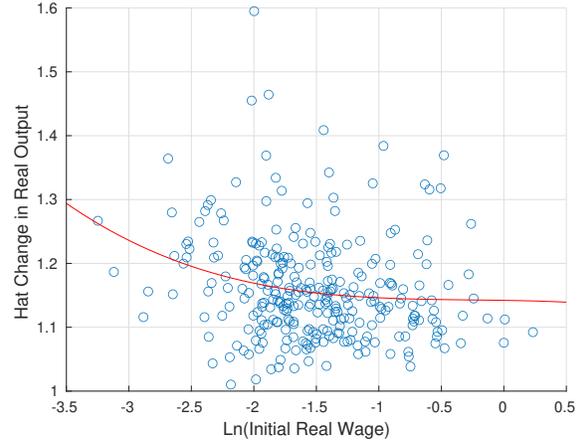
(a) Real Wage



(b) Price Index



(c) Population



(d) Output

Figure 4: The Impacts of Transportation Networks

Note: The figures present the impacts of transportation network improvements between 1995 and 2016 on real wage, price index, and population against the initial level of real wage in 1995. Each dot represents a prefecture-level city.

Table 3: The Impacts of Transportation Networks on Spatial Inequality

	No Change		Baseline		Trade Only		Migration Only	
	$w/P$	$Y$	$w/P$	$Y$	$w/P$	$Y$	$w/P$	$Y$
$\ln(\text{STD})_T - \ln(\text{STD})_1$	0.0168	0.0443	-0.0149	0.0197	-0.0193	0.0031	0.0225	0.0653
$100 \times (\text{HHI}_T - \text{HHI}_1)$	0.0127	0.1643	-0.0052	0.1128	-0.0085	0.0660	0.0175	0.2256
$\text{Gini}_T - \text{Gini}_1$	0.0105	0.0287	-0.0058	0.0208	-0.0087	0.0116	0.0141	0.0392

Note: This table reports the impacts of transportation networks between 1995 and 2016. “No Change” refers to the counter-factual in which the infrastructure was fixed at their 1995 level. “Baseline” is the case in which the improvements in transportation networks reduce both trade and migration costs. “Trade Only” and “Migration Only” refers to the case in which the improvements only reduce the trade and the migration costs, respectively. The first three rows report measures of spatial inequality across the cities over time, where 1 refers to the initial year and  $T$  refers to the final year.

thus increases the spatial inequality. In a world where agglomeration dominates the congestion force ( $\alpha + \beta > 0$ ), the expansion of the road networks allows the people from small and inland cities to migrate to large and coastal cities. In addition, the dominant agglomeration force implies that the population’s concentration in large cities will widen the existing productivity differences, leading to higher spatial inequality. The overall impacts of transportation networks on spatial inequality thus depend on the agglomeration( $\alpha$ ) and congestion( $\beta$ ) forces, as well as the elasticity of trade and migration flows relative to underlying infrastructure captured by  $\theta\psi$  and  $\kappa\xi$  in our model. Our quantitative results suggest that transportation networks acted as an equalizer in China during our sample years. This is fundamentally because 1) the Chinese cities are estimated to be in the net agglomeration regime ( $\alpha + \beta = 0.3194$ ), and 2) inter-city trade is much more responsive to transportation than migration ( $\theta\psi = 0.89$  and  $\kappa\xi = 0.34$ ).

To highlight the roles of trade and migration, we compare several key economic variables at the city-level in the year 2016. With a slight abuse of the notation in the literature, we use the “hat” of a variable to measure the impact of transportation networks. Specifically, we define  $\hat{x}_{it} \equiv x'_{it}/x_{it}$ , where  $x_{it}$  is the value of  $x$  in city  $i$  at year  $t$  in the “no change” counter-factual, and  $x'_{it}$  is the value in the “baseline” simulation. By this definition,  $\hat{x}_{it}$  measures the percentage impact of the transportation networks in city  $i$  at year  $t$ . Figure 4 presents  $\hat{x}_{it}$  in the final year, 2016, against the initial level of the real wage for four variables: real wage, price index, population, and output. The first panel in Figure 4 confirms the finding that transportation networks reduce spatial inequality. While all the cities enjoy a higher real wage due to the expansion of the road networks, the gain is negatively correlated with the

initial real wage. The real wage in large cities such as Beijing and Shanghai only improves by around 5 percent, and in smaller cities, the impact could be as high as 35 percent. The following three panels shed light on the different roles of trade and migration.

The second panel highlights that the equalizing effect of transportation networks comes from trade liberalization. The graph plots the hat-changes in the price index against the initial level of the real wage. While the price index in the large cities is barely affected by the expanding networks, the smaller cities see their price index drop by around 10 to 15 percent. In the model with productivity agglomeration and regional productivity differences, the goods in the larger cities are, on average cheaper than small cities as the producers tend to be more productive. In this case, better access to the goods market in the large cities tend to leave a larger impact on the price levels of the smaller cities, while the impacts on the price levels of large cities are much lower and negligible.

The impacts on population flow plotted in the third panel suggest that reduced migration frictions serve as a counter-force to increase spatial inequality. Since the transportation network better connects the remote cities to the rest of the country, people find it easier to migrate out to the coastal cities that offer a higher wage. Therefore, as shown in panel (c) of Figure 4, cities with initially higher real wages tend to receive a population inflow. In contrast, the initially poor and remote cities struggle to retain their workforce. The population concentration due to productivity agglomeration thus implies that spatial inequality tends to increase given the improvements in transportation networks. Finally, since the impacts on the per-capita output and population run in the opposite directions, the resulting impact on city-level output does not line up clearly against the initial real wage, as shown in the last panel of the same figure.

## 5.2 Decomposing the Impacts of Transportation: Trade v.s. Migration

In the previous part, we have compared a “no change” counterfactual simulation with a “baseline” simulation in which the frictions in both trade and migration decline. In this section, we separate the effects of trade and migration on spatial inequality by running two

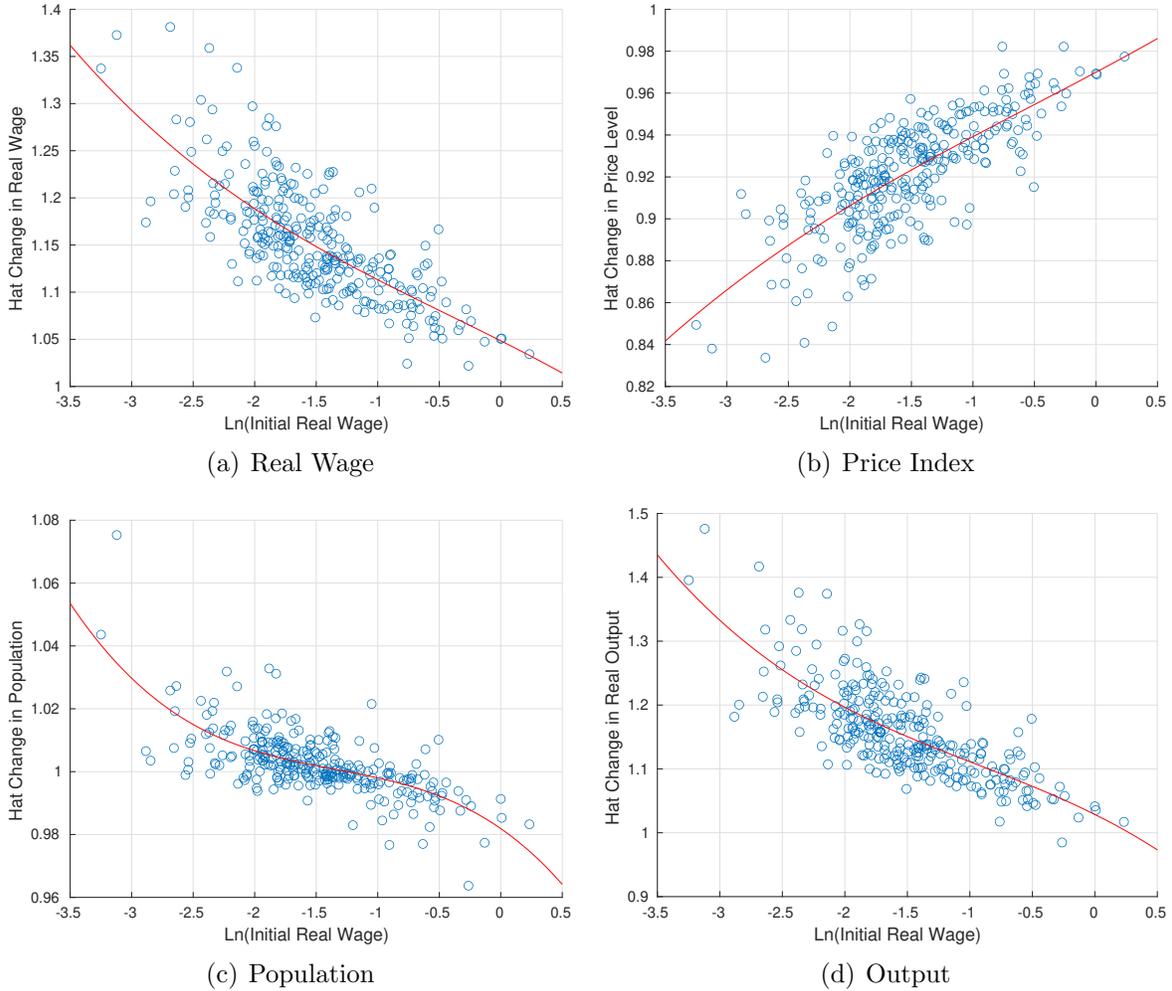
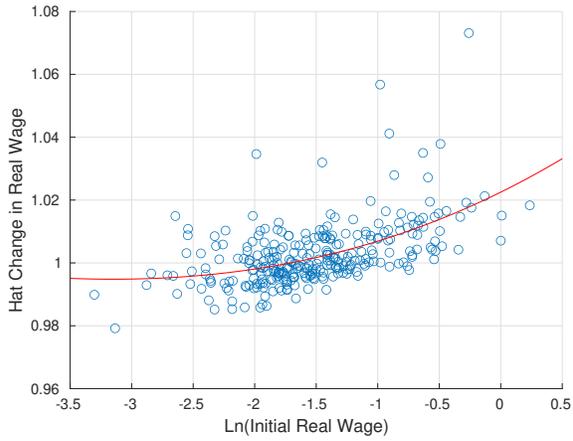


Figure 5: The Impacts of Transportation Networks via Trade

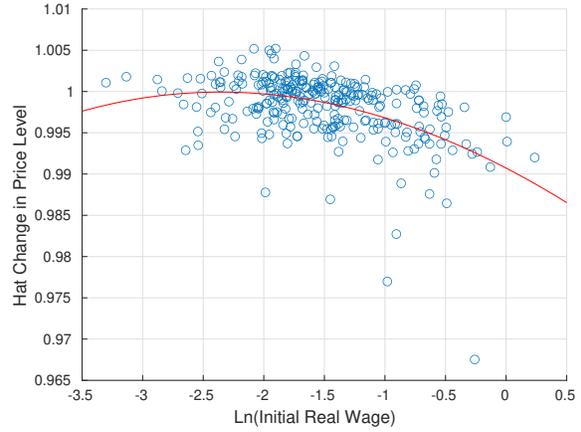
Note: The figures present the impacts of transportation network improvements between 1995 and 2016 on real wage, price index, and population against the initial level of real wage in 1995. Each dot represents a prefecture-level city.

auxiliary counterfactual simulations. In the first decomposition exercise called “trade only”, we assume that the improvements in transportation networks only reduce trade costs but not the migration costs. Following the same logic, in the “migration only” counterfactual simulation, the migration frictions decline over the years, but the trade costs remain at the 1995 level. Comparing each of the two cases with the “no change” counterfactual disentangle the contribution of each channel. The results are reported side-by-side in Table 3, as well as in Figure 5 and 6.

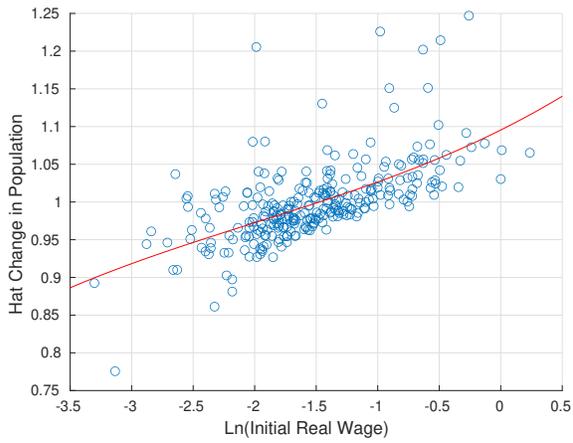
When better infrastructure only facilitates inter-city trade, spatial inequality declines



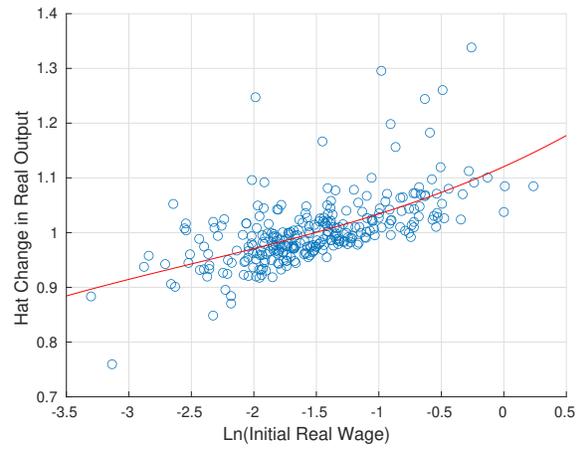
(a) Real Wage



(b) Price Index



(c) Population



(d) Output

Figure 6: The Impacts of Transportation Networks via Migration

Note: The figures present the impacts of transportation network improvements between 1995 and 2016 on real wage, price index, and population against the initial level of real wage in 1995. Each dot represents a prefecture-level city.

even more compared to the baseline case in which both costs decline. As shown in the fifth column of table 3, the logarithm of the standard deviation of real wage declines by 1.93 percentage points. In comparison, the same measure only declines by 1.49 percentage points in the baseline simulation. The other measures of spatial inequality convey the same message. The additional equalizing effect comes from the population movement as highlighted in Figure 5. In this figure, we compare the “no change” with the “trade only” counterfactual in the same manner as in Figure 4. Population movement in panel (c) shows that, contrary to the baseline case in which population flows into the large cities, trade liberalization on its own direct the population away from the large municipalities. With better access to the goods market in large cities, smaller cities are more attractive as the residents do not suffer as much congestion disutilities as those living in large cities. As a result, the smallest cities’ population roughly increases by 8 percent, while the largest cities modestly lose 2 to 4 percent of their initial population. Spatial inequality falls subsequently as population distribution is less concentrated. Lastly, the first two panels in Figure 5 are almost identical to the first two panels in Figure 4 in which both the trade and the migration forces are at play. The similarity suggests that the impact of trade liberalization on real wage and price index dominates the impact of factor movement.

On the contrary, if the expansion of the traffic networks only facilitates the movements of people, spatial inequality will soar up. The last column of Table 3 reports that in the “migration only” case, all measures of spatial inequality are significantly higher than the “no change” case. In other words, if transportation networks only reduce migration friction, then infrastructure improvements will lead to higher spatial inequality. Figure 6 highlights the reason behind the result. Holding the trade costs constant, a reduction in the migration frictions draws people away from the small to the large cities, as shown in panel (c). The resulting population concentration improves the productivity of the large cities relative to the small ones through agglomeration. As a result, the gaps in real wage and total output widen up, as shown in panel (a) and (d) of the figure.

Table 4: The Impacts of Transportation Networks on the First Moments

	No Change	Baseline	Trade Only	Migration Only
Aggregate Growth, 1995 to 2016	0.0857	0.2221	0.1943	0.1149
Trade Share, 2016	0.7298	0.7865	0.7868	0.7295
Migration Share, 1995 to 2016	0.0967	0.1222	0.0931	0.128

Note: This table reports the aggregate impacts of transportation networks between 1995 and 2016. “No Change” refers to the counterfactual in which the infrastructure was fixed at their 1995 level. “Baseline” is the case in which the improvements in transportation networks reduce both trade and migration costs. “Trade Only” and “Migration Only” refer to the case in which the improvements only reduce the trade and the migration costs, respectively. The first three rows report measures of spatial inequality across the cities over time, where 1 refers to the initial year, and  $T$  refers to the ending year.

### 5.3 The Impacts on the First Moments

In the previous parts, we have focused on the distributional impacts of transportation networks. In the current section, we briefly discuss the impacts of the transportation networks on the first moments of the key variables: the total output as well as the trade and the migration shares. Table 4 summarizes the results.

In the baseline case, the aggregate output has grown by 22.2 percent between 1995 and 2016. As compared to the 8.6 percent growth in the “no change” case, the infrastructure buildup has more than doubled the speed of economic growth<sup>10</sup>. The impact on aggregate economic growth is modest relative to the ten-fold economic growth in the data over the same period. The model is not designed to explain the first moments as we have abstracted away from many factors that lead to higher total factor productivity (TFP), such as technology improvements, physical capital formation, and human capital investments. The absence of a time trend in productivity does not affect the distributional results, as the aggregate productivity growth will only shift up the real wage of all the cities uniformly. Since the primary focus of this paper is the distributional impacts of transportation networks, we leave the elements of economic growth to future works. Nevertheless, our framework allows us to decompose the contribution of trade and migration to aggregate economic growth.

Trade liberalization explains around 80 percent of the aggregate impact from transportation networks expansion, while migration explains the remaining 20 percent. We again turn

<sup>10</sup>The aggregate economy in the “no change” simulation grows because people slowly move into the larger cities over the years. The concentration into the larger cities improves aggregate output because 1) larger cities are more productive (higher  $\bar{A}$ ), and 2) productivity agglomeration.

to the counterfactual simulation in which we only allow the improvements in infrastructure to affect one friction at a time. When migration frictions are held constant at the initial level, the reductions in iceberg trade costs leads to a  $19.4 - 8.6 = 10.8$  percent growth in total output. The impact of trade is around  $10.8/13.6 \approx 79.4$  percent of the overall impact. On the other hand, the reduction in migration frictions alone can only deliver  $11.5 - 8.6 = 2.9$  percent of economic growth, which is 21.3 percent of the overall growth observed in the baseline case. The net impacts of trade and migration add up to 13.7 percent, which is only slightly higher than the gain in the baseline case (13.6 percent). This suggests that the interaction between the two channels is negligible.

## 6 Conclusion

In this paper, we examine the distributional impact of transportation networks in China. To do this, we first provide a comprehensive and time-consistent panel dataset on the expansion of the transportation networks in China between 1995 and 2016. Based on this dataset, we further evaluate the distributional impacts and argue that infrastructure investments have significantly reduced spatial disparity. We show that the reduction in spatial disparity is rooted in the fact that goods are more mobile than people in response to the improved road networks. Therefore, the impacts of trade liberalization quantitatively dominate factor concentration, leading to a reduction in spatial inequality.

## References

- Allen, Treb and Costas Arkolakis**, “Trade and the Topography of the Spatial Economy,” *The Quarterly Journal of Economics*, 2014, 1085, 1139.
- **and Dave Donaldson**, “The geography of path dependence,” Technical Report, working paper 2018.

- Banerjee, Abhijit, Esther Duflo, and Nancy Qian**, “On the Road: Access to Transportation Infrastructure and Economic Growth in China,” Working Paper 17897, National Bureau of Economic Research March 2012.
- Baum-Snow, Nathaniel**, “Did Highways Cause Suburbanization?\*,” *The Quarterly Journal of Economics*, 05 2007, 122 (2), 775–805.
- , **J. Vernon Henderson, Matthew A. Turner, Qinghua Zhang, and Loren Brandt**, “Does investment in national highways help or hurt hinterland city growth?,” *Journal of Urban Economics*, 2018.
- Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum**, “Plants and Productivity in International Trade,” *American Economic Review*, September 2003, 93 (4), 1268–1290.
- Donaldson, Dave and Richard Hornbeck**, “Railroads and American Economic Growth: A Market Access Approach \*,” *The Quarterly Journal of Economics*, 02 2016, 131 (2), 799–858.
- Eaton, Jonathan and Samuel Kortum**, “Technology, Geography, and Trade,” *Econometrica*, Sep. 2002, 70 (5), 1741–1779.
- , – , and **Francis Kramarz**, “An Anatomy of International Trade: Evidence From French Firms,” *Econometrica*, 09 2011, 79 (5), 1453–1498.
- Faber, Benjamin**, “Trade Integration, Market Size, and Industrialization: Evidence from China’s National Trunk Highway System,” *The Review of Economic Studies*, 03 2014, 81 (3), 1046–1070.
- Gurara, Daniel, Vladimir Klyuev, Nkunde Mwase, and Andrea F Presbitero**, “Trends and Challenges in Infrastructure Investment in Developing Countries,” *International Development Policy—Revue internationale de politique de développement*, 2018, 10 (10.1).

- Kapos, Valerie, J Rhind, Mary Edwards, MF Price, C Ravilious et al.**, “Developing a map of the world’s mountain forests.” *Forests in sustainable mountain development: a state of knowledge report for 2000. Task Force on Forests in Sustainable Mountain Development.*, 2000, pp. 4–19.
- Lin, Yatang**, “Travel costs and urban specialization patterns: Evidence from China’s high speed railway system,” *Journal of Urban Economics*, 2017, *98*, 98 – 123. Urbanization in Developing Countries: Past and Present.
- Ma, Lin and Yang Tang**, “Geography, trade, and internal migration in China,” *Journal of Urban Economics*, 2020, *115*, 103181.
- Morten, Melanie and Jaqueline Oliveira**, “Paving the Way to Development: Costly Migration and Labor Market Integration,” Working Paper 22158, National Bureau of Economic Research April 2016.
- Qin, Yu**, “No county left behind? The distributional impact of high-speed rail upgrades in China,” *Journal of Economic Geography*, 06 2016, *17* (3), 489–520.
- Redding, Stephen J. and Matthew A. Turner**, “Chapter 20 - Transportation Costs and the Spatial Organization of Economic Activity,” in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5 of *Handbook of Regional and Urban Economics*, Elsevier, 2015, pp. 1339 – 1398.
- Rozenberg, Julie and Marianne Fay**, *Beyond the gap: How countries can afford the infrastructure they need while protecting the planet*, The World Bank, 2019.
- Simonovska, Ina and Michael E. Waugh**, “The elasticity of trade: Estimates and evidence,” *Journal of International Economics*, 2014, *92* (1), 34 – 50.
- Tombe, Trevor and Xiaodong Zhu**, “Trade, Migration, and Productivity: A Quantitative Analysis of China,” *American Economic Review*, May 2019, *109* (5), 1843–72.
- World Bank**, “China - Investment Climate Survey 2005,” Technical Report, Enterprise Analysis Unit, World Bank Group 2005.

## A Tables and Figures

Table A.5: The Physical Maps

Year	Publisher	Scale	Projection
1994	Sino Maps	1:6 million	Albers, 25N, 47E
1995	Dizhi Maps	1:9 million	-
1996	Sino Maps	1:4.5 million	Albers, 25N, 47E
1997	Chengdu Maps	1:9 million	Albers
2000	Sino Maps	1:6 million	Albers
2002	Sino Maps	1:4.5 million	Albers, 25N, 47E
2003	Sino Maps	1:6 million	Albers, 25N, 47E
2007	Guangdong Maps	1:6 million	Lambert, 24N, 46N, 110E
2008	State Bureau of Surveying and Mapping	1:4.5 million	Albers
2009	Sino Maps	1:4.5 million	Albers, 25N, 47E
2012	Sino Maps	1:4.5 million	Albers, 25N, 47E
2013	Sino Maps	1:4.6 million	Albers, 25N, 47E
2017	Sino Maps	1:6 million	Albers

Note: This table presents the basic information on the physical maps in our collection. The Chengdu and Guangdong Maps are regional publishers, while the others are national. Both Albers and the Lambert projections are conic projections; Albers projection is an equal area projection while the Lambert projection is conformal. The coordinates following the projections are reference longitude and latitudes. The coordinate (25N, 47E) is commonly used for Chinese maps as it centers around Henan — the geographical center of China.

Table A.6: The Mapping Between Road Classification and Map Legends

Year	National Road	Highway	Railroad	High Speed Rail
1994	Zhong Yao Gong Lu	Gao Su Gong Lu	Dian Qi Hua Tie Lu	-
1994			Shuang Gui Tie Lu	-
1994			Dan Gui Tie Lu	-
1995	Zhu Yao Gan Xian Gong Lu	Gao Su Gong Lu	Dian Qi Hua Tie Lu	-
1995			Fu Xian Tie Lu	-
1995			Dan Xian Tie Lu	-
1995			Zhai Gui Tie Lu	-
1996	Zhong Yao Gong Lu	Gao Su Gong Lu	Dian Qi Hua Tie Lu	-
1996			Shuang Gui Tie Lu	-
1996			Dan Gui Tie Lu	-
1997	Zhu Yao Gong Lu	Gao Deng Ji Gong Lu	Tie Lu	-
2000	Guo Dao	Gao Su Gong Lu	Tie Lu	-
2002	Guo Dao	Gao Deng Ji Gong Lu	Tie Lu	-
2003	Guo Dao	Gao Deng Ji Gong Lu	Tie Lu	-
2007	Guo Dao	Gao Su Gong Lu	Tie Lu	-
2008	Guo Dao	Gao Su Gong Lu	Tie Lu	-
2009	Guo Dao	Gao Deng Ji Gong Lu	Tie Lu	-
2012	Guo Dao	Gao Deng Ji Gong Lu	Tie Lu	-
2013	Guo Dao	Gao Su Gong Lu	Tie Lu	Gao Su Tie Lu
2017	Guo Dao	Gao Deng Ji Gong Lu	Tie Lu	-

Note: This table shows the correspondence between map legends and the classification of transportation modes used in this paper. We directly report the Pinyin of the legends.

## B Map Digitization and Standardization

### B.1 Digitization

The collection of physical maps used in this paper are summarized in Table A.5 and A.6. After scanning these maps, we adopt standard digitization procedures: extract the modes of transportation by color identification and then geo-reference each map by inverting the projection methods. The geo-referenced maps are comparable across the years at the pixel level, despite the differences in projection methods and scale.

The transportation networks in the color-identified maps usually have paths that are 10 to 20 pixels in width. Given that each pixel corresponds to around 500 square meters in the real world, we need to shrink the color-identified maps down to one-pixel paths using the skeletonization algorithm. We use eight-connectivity at the pixel level, which means that each pixel in the network is considered to be connected to all eight of its neighbors, including the diagonal neighbors. We then break the entire skeleton network into “segments”, which are defined as the set of pixels between the branch and endpoints of the graph. The above procedure is applied to all modes of transportation and in all years.

To ensure that each segment is consistently represented over time, we use an iterative method. Conditional on the map in year  $t - 1$ , only new segments are added to the map in the year  $t$ . A segment is defined as “new” if 90% of its pixels are more than 10km away from the existing graph. If a segment is classified as “old”, then we use the pixels in year  $t - 1$  as the location in the year  $t$ . Adding new segments in this way might leave gap areas between the new and the old segments. We use the “dilate and connect” method to connect the new segments to the existing network if the gap indeed exists. It is reasonable to assume that any new construction is connected to the existing network.

### B.2 Standardization

To standardize the roads across time and space, we rely on the publications from the Ministry of Transportation from China: the *Technical Standard of Highway Engineering*. We use the following four revisions: 1988 (JTJ01-88), 1997 (JTJ01-97), 2003 (JTG B01-2003), and 2014

(JTG B01-2014). We use the “design speed” of highways and the first-rate roads in each publication to determine the design speed of newly-constructed highways and national roads, respectively, in the map legends.

The information on design speed comes from the *Technical Standard of Highway Engineering*. In the 1988 revision, the design speed, written as “Ji Suan Xing Che Su Du” as in Pinyin, depends on the terrain as stipulated in Chapter 2.0.2. The dependence on terrain is particularly emphasized for the construction of highways, due to the engineering difficulties and costs in construction. In the mountainous regions, the default design speed is 80km/h, and 60km/h speed is reserved for “particularly difficult segments”. Since no further detail is provided in defining the conditions for using the 60km/h design speed, we choose to use 80km/h for all the highways in the mountain areas constructed under the 1988 revision. The design speed for the first-rate roads is straightforward: 100km/h in plains and low rolling hills, and 60km/h in the high hills and mountains.

The major change in the 1997 revision is to remove the dependence on terrain for highway construction. This change is explicitly stated in the notes as “the design speed of highway is no longer linked to the underlying terrain, ..., under normal circumstances, the design speed should be 120km/h”. Under limited conditions, the design speed can be lowered to 100 or 80 km/h. In “particularly difficult segments”, 60km/h is still admissible. For this reason, we use a design speed of 120km for all terrains except for the mountains, which we still use 80km/h as design speed. The construction of the first-rate roads is still reliant on the underlying terrain, and there is no change in the design speeds.

The next revision in 2003 focused on the first-rate roads. In high hills and mountains, the design speed increased from 60km/h to 80km/h, the same as the highways. In the highway construction, the 60km/h design speed is explicitly discouraged, as the Ministry learned that it is difficult to upgrade a highway with a 60km/h design speed to accommodate the ever-increasing traffic volume. The 2003 revision still allows a 60km/h design speed for highways, but it stipulated that such segments cannot be longer than 15km. The 2003 revision also changed the Chinese translation of “design speed” from “Ji Suan Xing Che Su Du” to a more direct translation of “She Ji Su Du”, to be comparable to the international standards.

For our purposes, the design speed of the highways and the first-rate roads did not change

in the latest 2014 revision. The latest revision stated that the 2003 revision was already well-crafted, and was widely tested in practice. For consistency, the Ministry no longer saw a need to revise the design speeds further.

**Definition of Terrains** As shown in Table 1, the design speed of the roads differ by terrains, and there are four types of terrains used in the *Technical Standards of Highway Engineering*: the plains (Ping Yuan), the low rolling hills (LRH, Wei Qiu), the hills (Zhong Qiu), and the mountains (Shan Ling).

These types of terrains are defined in the *Land Regulations in Highway Engineering*, also published by the Ministry of Transportation. In the *Land Regulations*, the plains are the areas where the slopes are smaller than 3 degrees. The low rolling hills are the regions where the slopes are between 3 and 20 degrees, and the local elevation range (LER) is less than 200 meters. However, the *Land Regulations* groups the “hills” and the “mountains” into one category (Zhong Qiu Shan Ling), and defines it as the terrains with a slope greater than 20 degrees, or with LER higher than 200 meters. This aggregation is because by the time that the *Land Regulations* was published in 1999, the Ministry of Transportation no longer distinguished between these two terrains. The Ministry also published the official definition of the technical terms in the *Standard of Technical Terms for Highway Engineering*. Unfortunately, the distinction between “hills” and “mountains” are subjective and only descriptive in this official publication. The “hills” are defined as the areas with relatively large LER but without those characteristics of a typical mountain, such as a ridge, peak, or base. The “mountains”, on the other hand, are the terrains with large LER and observable characteristics. To quantify the design speed by terrain, we need a systematic way to distinguish between the hills and the mountains. This is because in the 1988 revision of the standards, the design speed of highways differs between the hills and the mountains, as seen in Table 1.

To this end, we turn to the official definition used by the United Nation Environment Programme World Conservation Monitoring Centre (UNEP-WCMC), which defines mountains as those with LER greater than 300 meters, according to Kapos et al. [2000]. Taking the *Land Regulations* and the UNEP definitions together, we define the “hills” as the terrains with a slope of more than 20 degrees, or with LER between 200 and 300 meters, and the

“mountains” as those with a slope of more than 20 degrees and with LER more than 300 meters. We use the data from GTOPO30 from the USGS for the slope and the LER data. The slope and LER are evaluated within a five-pixel radius following [Kapos et al. \[2000\]](#).

### B.3 Other Geographic Data Sources

In addition to the maps and publications mentioned above, we have also utilized the following datasets:

1. **GTOPO30** from USGS. This dataset provides the slope and elevation data with a grid spacing of 30 arc seconds, which is approximately 1 kilometer. We use this data source for two purposes: 1) to define the terrains as specified above, and 2) to approximate the construction costs of roads, which is used to calculate the minimum spanning tree following [Faber \[2014\]](#).
2. **Global Land Cover Characteristics Data Base** from USGS. This dataset provides land cover characteristics. We use this dataset together with GTOPO30 to approximate the construction costs of roads.
3. **NOAA Dataset** on temperature and precipitation. We use this dataset to estimate the fixed components of amenity in each prefecture.

### B.4 The Economic Data Sources

The economic datasets used in this project are:

1. The *One Percent Population Survey* in 2005. This dataset is used to back out the bilateral population flow between the prefectures.
2. The *Custom Transaction Dataset* between 2000 and 2005. This dataset is used to estimate the trade elasticity.
3. The *City Statistical Yearbooks*. This dataset provides the data on goods and passenger transportation volume through each city by mode of transportation. This dataset is used to estimate the geographic transportation costs between the prefectures.