DISCUSSION PAPER SERIES

DP15372

CHINA'S DAZZLING TRANSPORT-INFRASTRUCTURE GROWTH: MEASUREMENT AND EFFECTS

Peter Egger, Gabriel Loumeau and Nicole Loumeau

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

CHINA'S DAZZLING TRANSPORT-INFRASTRUCTURE GROWTH: MEASUREMENT AND EFFECTS

Peter Egger, Gabriel Loumeau and Nicole Loumeau

Discussion Paper DP15372
Published 15 October 2020
Submitted 13 October 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

CHINA'S DAZZLING TRANSPORT-INFRASTRUCTURE GROWTH: MEASUREMENT AND EFFECTS

Abstract

We document an unprecedented change in the size and the quality of China's transport-infrastructure network between 2000 and 2013. This documentation is based on hand-collected and digitized data on roads and railways. The changes are summarized and portrayed as shortest-possible transport times of people and goods between 330 prefectures of mainland China. A quantitative model of China's prefectures and a Rest of the World, featuring both goods trade and migration, suggests that the long-run consequences of the transport-infrastructure changes induce regional convergence of lagging-behind prefectures in terms of population density and, to a lesser extent, in terms of real per-capita income. Not only changes in highway and high-speed-railway networks but also ones in lower-level road and railway networks are quantitatively important. Key drivers behind the effects are the facilitation of goods transport as well as technology diffusion, while the reduction of mobility costs and the diffusion of amenities appear less important.

JEL Classification: F14, R13, R41

Keywords: Transport Infrastructure, regional economics, General-equilibrium models, structural estimation, migration, Transportation problem

Peter Egger - egger@kof.ethz.ch
ETH Zurich and CEPR

Gabriel Loumeau - loumeau@ethz.ch
ETH Zurich

Nicole Loumeau - nicole.loumeau@kof.ethz.ch
ETH Zurich

China's Dazzling Transport-infrastructure Growth: Measurement and Effects

Peter H. Egger[*] ETH Zurich CEPR, CESifo

Nicole Loumeau[†] ETH Zurich

Gabriel Loumeau[‡] ETH Zurich

October 13, 2020

Abstract: We document an unprecedented change in the size and the quality of China's transport-infrastructure network between 2000 and 2013. This documentation is based on hand-collected and digitized data on roads and railways. The changes are summarized and portrayed as shortest-possible transport times of people and goods between 330 prefectures of mainland China. A quantitative model of China's prefectures and a Rest of the World, featuring both goods trade and migration, suggests that the long-run consequences of the transport-infrastructure changes induce regional convergence of lagging-behind prefectures in terms of population density and, to a lesser extent, in terms of real per-capita income. Not only changes in highway and high-speed-railway networks but also ones in lower-level road and railway networks are quantitatively important. Key drivers behind the effects are the facilitation of goods transport as well as technology diffusion, while the reduction of mobility costs and the diffusion of amenities appear less important.

Keywords: Transport infrastructure; Regional economics; General-equilibrium models; Structural estimation; Migration; Transportation problem.

JEL classification: C68; F14; R13; R41.

[*]ETH Zurich, Department of Management, Technology, and Economics, Leonhardstrasse 21, 8092 Zurich, Switzerland; E-mail: egger@kof.ethz.ch; + 41 44 632 41 08. Egger gratefully acknowledges funding from the Swiss National Science Foundation under grant no. 100018_169537.

[†]ETH Zurich, Department of Management, Technology, and Economics, Leonhardstrasse 21, 8092 Zurich, Switzerland; E-mail: loumeau@kof.ethz.ch; +41 44 632 67 58.

[‡]*Corresponding author:* ETH Zurich, Department of Management, Technology, and Economics, Leonhardstrasse 21, 8092 Zurich, Switzerland; E-mail: nicole.loumeau@kof.ethz.ch; +41 44 632 33 48.

1 Introduction

China's emergence on the global economic landscape has become a major topic in political as well as academic discussions. Understanding how one of the poorest secluded economies became a major economic force within only a few decades is of wide interest. Clearly, the trade-policy environment and the absorption of foreign technology are important factors here. However, the degree to which China gained market access – not only abroad but also at home – through a gigantic effort in building transport infrastructure is little understood. In general, transport-infrastructure networks are particularly important for economies which are large in terms of both geographical area and population. China is a prime example for that. It lagged significantly behind the western world in terms of the size and quality of its transport-infrastructure network up until the end of the 20th century. However, since the early 1990s, and particularly since its membership in the World Trade Organization (WTO) in 2001, there was an enormous surge in transport-infrastructure projects in China.[1] In 2000, the total length of the Chinese highway network amounted to 64% of the length of the U.S. Interstate Highway System, and no high-speed railway (HSR) lines were developed in the country. A mere thirteen years later, China's highway network represented 183% of the total length of the U.S. Interstate Highway System, and its HSR network was the longest in the world by a 40% margin.

To measure and study this unique and unprecedented case of transport-infrastructure growth, we compiled the largest and most comprehensive existing dataset based on hard copies of annual road and railway atlases for China in digital format. This dataset distinguishes between three hierarchical layers of roads (highways, province-level roads, and prefecture-level roads) and two hierarchical layers of railways (high-speed- and standard-railway lines). Moreover, it permits documenting the following stylized facts about the evolution of mainland China's overland-transport-infrastructure network between the years 2000 and 2013. First, the total length of the road network increased by about 28% until 2013, starting from 371,385 kilometers in 2000. The total length of highways alone almost tripled, from 50,127 kilometers in 2000 to 142,983 kilometers in 2013. In turn, the total length of Chinese regional (non-highway) roads increased by 77,855 kilometers or 21%. Over the same time span, 12,996 kilometers of high-speed-railway (HSR) lines were added to the network, making it the longest high-speed rail network in the world, and 20,141 kilometers of standard rail lines were built between 2000 and 2013. Second, the functions

[1] In 2017, the World Bank funded such projects all over the world worth $1,1 billion. Of those, $450 million were invested in China (Projects & Operations, The World Bank). In the same year, the estimated private investments in transport infrastructure in China amounted to $13 billion (Private Participation in Infrastructure Project Database, The World Bank). Similarly, India saw its share of infrastructure investment to GDP double between 2004 and 2016 (Infrastructure Investment, OECD.)

of these different network layers are different: highways and HSR lines mainly connect big centers (the former for people and particularly goods, and the latter only for people), standard railways provide access between large and medium-sized centers (for both goods and people), and regional roads play an important role in making highways and the HSR network accessible. We use typical maximum-speed limits for the different road layers, and recorded average speeds for the transport of goods and people for the different railway layers. Given typical transport mode usage patterns, we see a drastic decline in travel times – along the fastest route – for both goods and people over the 13-year time span. The average fastest travel time between all prefecture pairs declined by 13% for goods and by 50% for people between 2000 and 2013.

Clearly, these massive changes in transport opportunities in China must have meant a drastic decline of (explicit or implicit) transport costs for people and goods. In turn, this must have translated into changes in closely-related economic outcomes, in particular, the distribution of people and incomes across China. To shed light on the latter, we outline, calibrate and simulate a quantitative model of 330 prefectures plus a Rest of the World (RoW). Key features of the model are the following. First, goods and people are mobile between prefectures on the one hand, and between China and the RoW on the other hand (entailing trade and migration). Moving is costly for both goods and people. The moving costs are informed by the data on China's transport infrastructure network. Second, locations are attractive because of production and consumption potential. For the former, a composite of exogenous and endogenous technology components, and for the latter a composite of exogenous and endogenous amenity components are introduced. Endogenous technology depends on regions' own and others' population density, and so do endogenous amenities. We hypothesize – and the data suggest – that there are technology as well as amenity spillovers from better transport-accessible regions in the average prefecture. Endogenous technology and amenities adjust sluggishly in response to past population density: a higher density promotes subsequent technology and it reduces subsequent amenities.

The quantification rests on two key pairs of parameter estimates: the ones linking travel times for goods and people to the respective moving costs; and the ones determining the transport-related diffusion processes of technology and amenities. For either pair of parameters endogeneity is a concern, as China's infrastructure planning may be partly informed by latent drivers of goods trade, migration, endogenous technology, and endogenous amenities. To address this problem, we devise two alternative and independent novel instrumental-variable (IV) approaches. A first approach solves an optimal-transport problem[2] using only Chinese topological information, on top of the locations of prefectures. The second IV approach, instead, uses only historical information

[2]See Monge (1781) and Kantorovitch (1958) for the initial formulation of the relevant class of problems, and Villani (2003) for a more recent and comprehensive approach.

in the construction of the instruments. The historical approach combines the population count in 1953 and the path of courier routes employed during the Chinese Ming dynasty.[3] Both instruments deliver similar answers regarding the parameters of interest to this study.

We learn the following from the quantitative model when shocking China's transport network in 2000 by using its configuration of 2013 instead. The improvements (through greater length and quality) lead to a relative convergence of population density and, to a lesser extent, of real per-capita incomes across China relative to a no-change long-run equilibrium. The reason is that the network changes – in particular, in the regional road system but also in the standard-railway system – connects originally relatively remote and less central prefectures to the improved faster networks. This provides them better access to China's central prefectures and to the rest of the world. Two components contribute the lion's share to the change: reduced goods-transport costs and greater diffusion of technology. Reduced mobility costs for people and amenity diffusion (i.e., better access to amenities in close-by prefectures) partially off-set each other so that their net effect appears of relatively lesser importance.

The paper proceeds as follows. We outline our relationship to earlier economic work on transport infrastructure networks in broad brushes in the subsequent section. We describe the evolution of the Chinese transport network between 2000 and 2013 in Section 3. In Section 4, we present the theoretical framework which features four channels through which transport systems may cause economic effects. Before presenting the counterfactual analysis in Section 6, we outline the model calibration in Section 5. Section 7 concludes.

2 Literature

This paper's agenda relates to three strands of earlier work. First, there is growing interest in the measurement of road- and railway-transport infrastructure in economics. However, at least at the international level, data are extremely scarce. This is why much of the measurement focuses on individual, large countries such as the United States,[4] India,[5] or China,[6] and it focuses on railways[7] or the highway system.[8] Some of that literature works with track lengths (i.e., the mileage between locations; see Duranton and Turner, 2012; Duranton et al., 2014; Donaldson, 2018) and a smaller

[3] The Ming dynasty ruled over China between 1368 and 1644; see Berman and Zhang (2017).

[4] Chandra and Thompson (2000); Baum-Snow (2007); Michaels (2008); Duranton and Turner (2012); Duranton et al. (2014); Donaldson and Hornbeck (2016)

[5] Donaldson (2018)

[6] Banerjee et al. (2012); Faber (2014); Baum-Snow et al. (2017); Baum-Snow and Turner (2017); Qin (2017); Baum-Snow et al. (2018); Barwick et al. (2018)

[7] Qin (2017); Donaldson and Hornbeck (2016); Donaldson (2018)

[8] Chandra and Thompson (2000); Baum-Snow (2007); Michaels (2008); Duranton and Turner (2012); Faber (2014); Duranton et al. (2014); Baum-Snow et al. (2018)

body of work maps track lengths into travel speeds (see Baum-Snow et al., 2018). But the latter are typically associated with overall transport, not distinguishing between goods and people. We add to this literature by compiling a novel, hand-collected and digitized dataset of China's road system which does not only cover highways but also provincial and prefecture roads. We augment this dataset with one on railways, covering not only the HSR network but also standard lines. Furthermore, using information on the physical network in conjunction with information on travel speeds and shortest paths, we establish a dataset of fastest travel times on the road and railway networks which differ between people and goods.

A second strand of work we relate to is the one on quantitative regional economic models which account for transport infrastructure. Such models are important to obtain an estimate of otherwise unobservable transport costs or more generally mobility frictions. In such models, both goods (through interregional trade) and people (through migration; sometimes also through commuting) are mobile between subnational, regional aggregates. Most of the respective work makes use of information on the transport infrastructure in some way. Part of the work uses it (or some coarse imputation of it) to establish a measure of transport costs between regions by computing shortest travel paths (see Allen and Arkolakis, 2014; Donaldson and Hornbeck, 2016; Desmet et al., 2018). This is done to obtain a measure not only of interregional transport costs but even of trade itself, as the latter is typically not observable between subnational units. Another part of work uses interregional trade or worker and household flows and an estimate of bilateral costs to measure the link function of travel times into trade or mobility costs (see Combes and Lafourcade, 2005; Ahlfeldt et al., 2015; Alder and Kondo, 2018). This is done to be able to do counterfactual analyses regarding shocks of the transport infrastructure. We relate to this work by adopting an exponential functional form that relates travel time to transport costs.

In the aforementioned quantitative work there are typically two channels of influence of transport networks on outcomes: goods-trade costs and migration (and, eventually, commuting) costs (see Allen and Arkolakis, 2014; Ahlfeldt et al., 2015; Alder and Kondo, 2018; Desmet et al., 2018). We add to this by opening up two additional channels: technology diffusion and amenity diffusion. Some earlier empirical work established that technology spillovers are geographically bound and, in a way, travel in space at some cost (see Coe and Helpman, 1995; Rodriguez-Clare, 1996; Feenstra and Kee, 2008; Fracasso and Vittucci Marzetti, 2015).[9] While such evidence is not established for amenities,[10] it is clear that better (faster) transport systems reduce the need for

[9]In dynamic multi-region quantitative models, innovation is a local activity that depends on the factors available there, but technology dissipates completely to all regions within a relatively short period (see Desmet and Rossi-Hansberg, 2009, 2014; Desmet et al., 2018; Allen and Donaldson, 2018)

[10]Even though in a different context and using a static approach, let us note that recent works in urban economics analyzing amenity spillovers as part of local agglomeration forces are similar in spirit to our approach

a perfect co-location of people and amenities, as amenities can relatively quickly be reached on roads or railways. We document that, of all four channels, at least among prefectures in China, the goods-transport and the technology-diffusion channels are most important.

Finally, we relate to a strand of work which treats the transport infrastructure as endogenous. This may be key, when estimating either the link-function parameters between travel speeds and transport costs (for goods or people) or the parameters governing the degree of diffusion of technology or amenities. The root of the alleged endogeneity is that the (national or subnational) transport-infrastructure planner may be influenced by the distribution of endogenous factors in space (of incomes, prices, people, ...). This could lead to a bias of the parameters of interest and, in turn, of the simulated quantitative effects of infrastructure changes. Broadly speaking, related earlier work invoked one of the following instrumental-variable strategies to address this problem: the use of some kind of route-optimizing network based on geographical features of the world only (see Allen and Arkolakis, 2014; Donaldson and Hornbeck, 2016; Desmet et al., 2018); or the use of historical/planned transport routes (see Baum-Snow, 2007; Michaels, 2008; Duranton and Turner, 2012; Garcia-Lopez, 2012; Duranton et al., 2014; Baum-Snow et al., 2018); or the construction of hypothetical transport networks (see Faber, 2014; Alder and Kondo, 2018). Our contribution to this literature is two-pronged. On the one hand, we construct a hypothetical transport network employing a Monge-Kantorovich solution to the optimal-transport problem based on features of China's topology alone (i.e., without using any information on the distribution of people). On the other hand, we use historical information on the courier routes during the Ming dynasty (1368-1644, see Berman and Zhang, 2017) and historical population data at the prefecture level. Both instruments deliver similar answers.

3 Chinese Transport-infrastructure Networks

In the last three decades, the development of transport infrastructure networks has been a key focus for the Chinese government. Both the road and railway networks were drastically expanded in pursuit of this goal. To study the evolution of these networks, we collected 14 infrastructure atlases covering the entire Chinese transport network for the years 2000-2013.[11] All atlases were digitized in high-resolution portable document format (PDF). After digitization, we geo-referenced all maps using ArcGIS. Following the official Chinese system, we classified the roads into three different categories: highways (including the National Trunk Highway System built between 1992

of amenity diffusion (see, among others, Ahlfeldt et al., 2015).

[11] We list the sources of all atlases in Table 6 of Appendix G.

and 2007), provincial-level roads, and prefecture-level roads.[12] Similarly, we classified the railways into two categories: high-speed-railway (HSR) lines and standard-railway lines. Since a significant fraction of associated transport infrastructure investments was spent on upgrading the existing network, our classification pays particular attention to such upgrading.

As a result, we obtained a novel and highly comprehensive database on the Chinese transport network (including highways, provincial-level roads, prefecture-level roads as well as high-speed and standard railways) at an annual level between 2000 and 2013. We use this database as a source to inform a regional quantitative analysis that investigates the overall economic effects and various components thereof across Chinese prefectures. In what follows, this section describes successively the different layers of this hand-collected novel dataset on the Chinese transport network.

3.1 Chinese Road Network

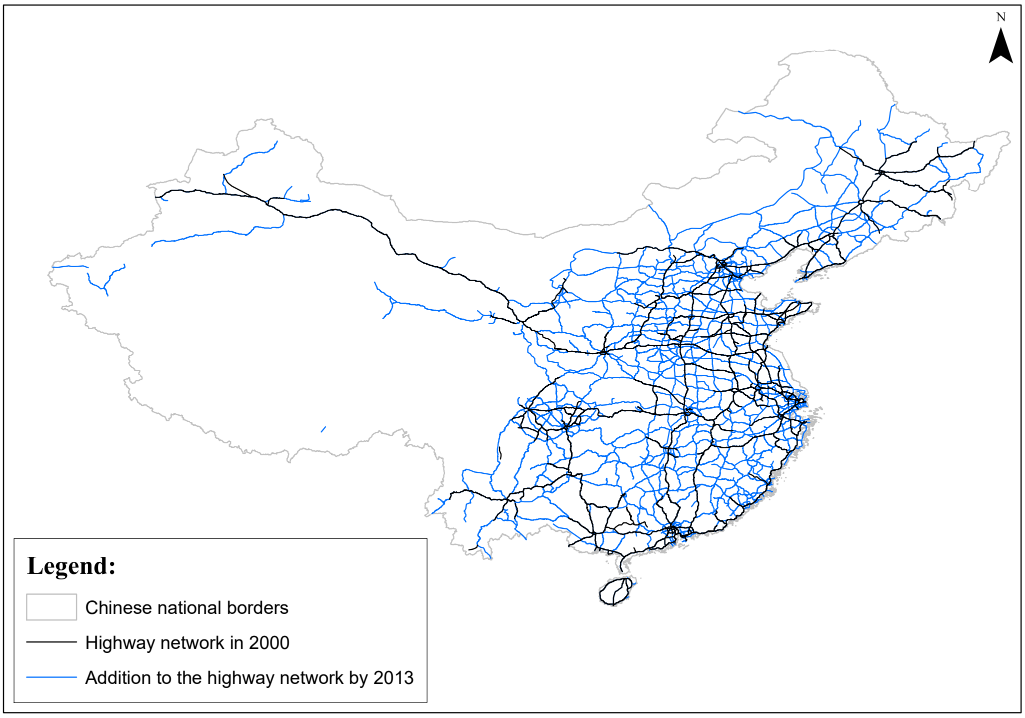
One key component of the change of the Chinese road network between 2000 and 2013 is the expansion of the highway network. The total number of kilometers of highways (including the National Trunk Highway System) grew from 50,127 kilometers in 2000 to 142,983 kilometers in 2013. By way of comparison, only this addition to the Chinese highway system during these 14 years is 20% longer than the entire US Interstate Highway System.[13] Figure 1 maps the evolution of the highway network between 2000 and 2013. As of 2013, all major population centers in China were connected by highways. Consistent with population-density patterns, most new highway segments were built in the eastern and northern parts of China.

However, looking solely at the highway network would neither capture the full extent nor the benefit of the Chinese road-network expansion. Figure 2 displays the evolution of the regional road network (i.e., the provincial-level and prefecture-level roads together), using gray lines for prefecture borders. For better display, we focus on a few prefectures in mid-eastern China here (namely, Pingdingshan, Luohe, and Zhoukou, all located in Henan Province). In 2013, Chinese provincial-level roads totaled 124,594 kilometers and prefecture-level roads amounted to 246,146 kilometers. Between 2000 and 2013, the total length of Chinese regional roads increased by 21%. While Figure 1 highlights the importance of the highway-network improvements on a nation-wide scale, Figure 2 reveals that regional roads are crucial to accurately capture bilateral connections at the prefecture level. This is especially true for short-distance travels and when it comes to providing shortest-distance access to the highway or railway systems. Prefectures are relatively

[12] In the road atlases, provincial-level roads are sometimes also referred to as national roads, and prefecture-level roads are dubbed secondary roads.

[13] In 2018, the US Interstate Highway System totaled 77,960 kilometers in length.

Figure 1: Chinese Highway Network (2000 and 2013)



small geographical units, some with a relatively small population count. Hence, to analyze the effects of transport-infrastructure improvements – even when focusing on highways or railways – it is elemental to consider regional road-network connections and their improvements.

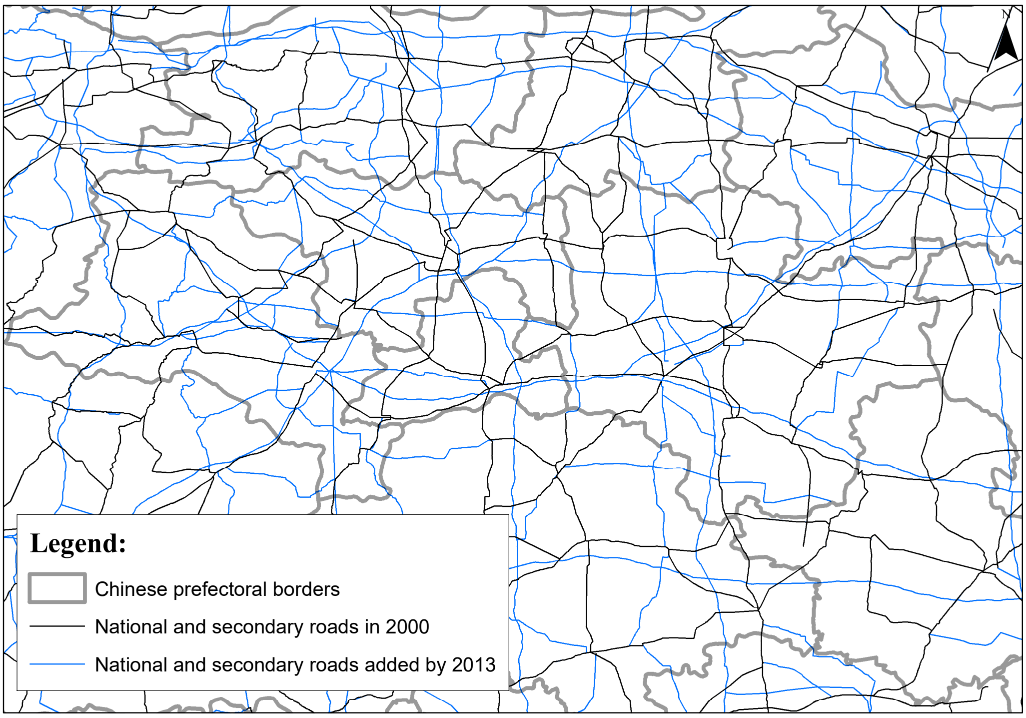
3.2 Chinese Railway Network

The construction of 12,996 kilometers of high-speed-railway (HSR) lines between 2008 and 2013 constitutes possibly the most striking change of the Chinese railway network. To put this development in perspective, it is worth noting that already in 2011, with 8,358 kilometers of length, the Chinese HSR network was the longest in the world. Figure 3 shows the location of the HSR lines in red, indicating that they are particularly dense in the coastal and more populated areas of the country. All large metropolitan centers such as Beijing, Shanghai, or Guangzhou are connected with each other through the HSR system.

Beyond the HSR network, the Chinese standard-railway network, displayed in blue in Figure 3, also grew significantly in length.[14] Due to lower fixed construction costs per kilometer relative to HSR lines, standard-railway lines connect a larger number of locations, and serve smaller mobility axes. Small peripheral prefectures in the West of the country were only connected by the railway

[14] We label as "standard" all railway lines which are not part of the HSR network.

Figure 2: Chinese Regional Road Network (2000 and 2013)



Notes: Regional roads include both provincial-level and prefecture-level roads. The figure zooms in on the prefectures of Pingdingshan, Luohe, and Zhoukou, all in Henan Province.

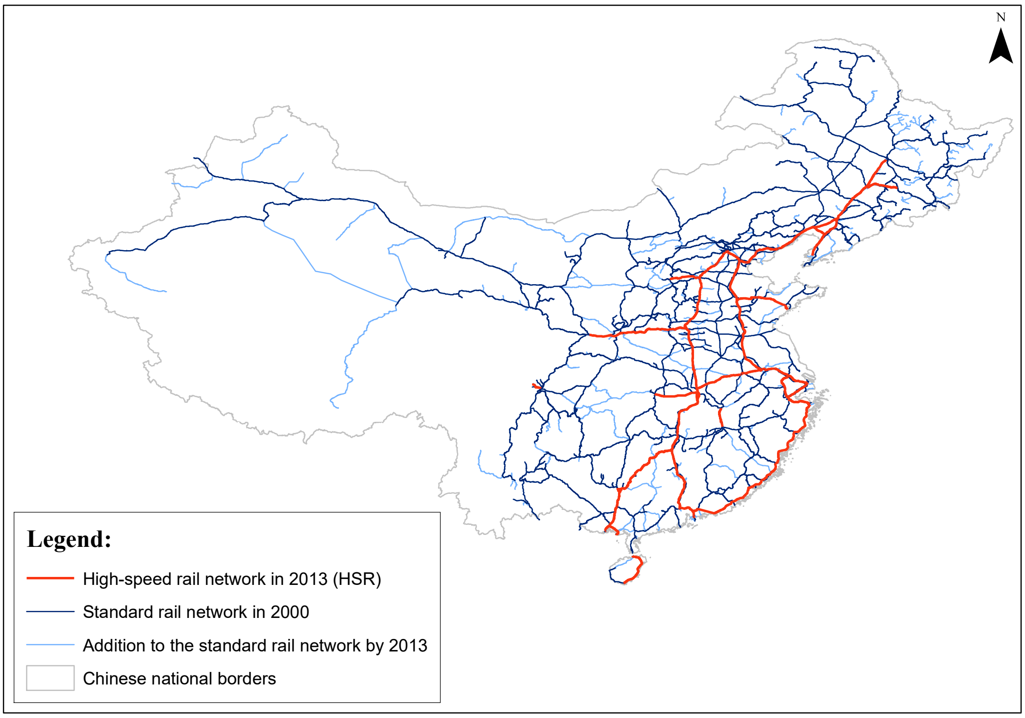
extension after 2000. This also implied that the growth of the standard-railway network appears more evenly distributed across the country than the growth of the HSR network. In total, 20,141 kilometers of new standard-railway lines were added to the network between 2000 and 2013.

3.3 Travel Times on the Network Layers

When geo-referencing the network, we registered attributes of the various connection legs on each network layer (highway, province-level, and prefecture-level roads; HSR and standard railways) for the purpose of goods and person transport. For instance, only people but no goods can travel on the HSR system, while goods and people can travel on other layers of the transport network. As prefectures are sometimes – but not always – connected directly to all layers of the network and due to different prevailing speed limits by layer and use, the average highest-possible speeds at which people and goods travel between prefectures differ.

For the purpose of measuring inter-prefectural distances, we define the location of every prefecture i by the main city center as measured by the highest density within a prefecture based

Figure 3: CHINESE RAILWAY NETWORK (2000 AND 2013)



on data from the Socioeconomic Data and Application Center (SEDAC) for the year 2000.[15] Inter-prefectural travel time is obtained by assuming that people and goods travel at the speed limits for roads and average recorded speed for railways.[16] To obtain the final bilateral travel time between two locations i and j for people, d^R_{ijt}, and goods, d^C_{ijt}, in any given year, we combine the travel time by road and rail.[17] Following descriptive evidence on Chinese travelers' mode choice by Wang et al. (2014), we suppose that people travel by road for distances of less than 100 kilometers, and by train for longer distances. The travel-mode choice for goods is determined solely by the minimum travel time between road and railway transport. Finally, we assume that travel times for both people and goods are symmetric between regions i and j, whereby for any year $d^R_{ijt} = d^R_{jit}$ and $d^C_{ijt} = d^C_{jit}$.

[15] More specifically, the highest-density point within a prefecture is the centroid of the 1×1 kilometer population-density cell with the highest density based on SEDAC data in the year 2000.

[16] Speed limits are 120km/h (kilometers per hour) on highways, 80km/h on provincial-level roads, and 50km/h on prefecture-level roads. Travel speeds on the railway network differ for people and goods, and they have significantly increased over time. Based on the average travel time in 2000, people traveled on standard railways at 57km/h, whereas goods moved at 32km/h. In 2013, people traveled at 250km/h on the HSR system and at 81km/h on standard-rail segments. Goods were shipped on rail at an average speed of 40km/h. See Lawrence et al. (2019) and Ma and Tang (2020) for details on travel speeds in China.

[17] The superscript R stands for *residential* and C stands for *commercial*. In order to remain consistent throughout the paper, we use the superscript R to refer to consumer/person-specific variables and the superscript C for firm/good-specific variables. This choice will become clear when we introduce the theoretical framework in Section 4.

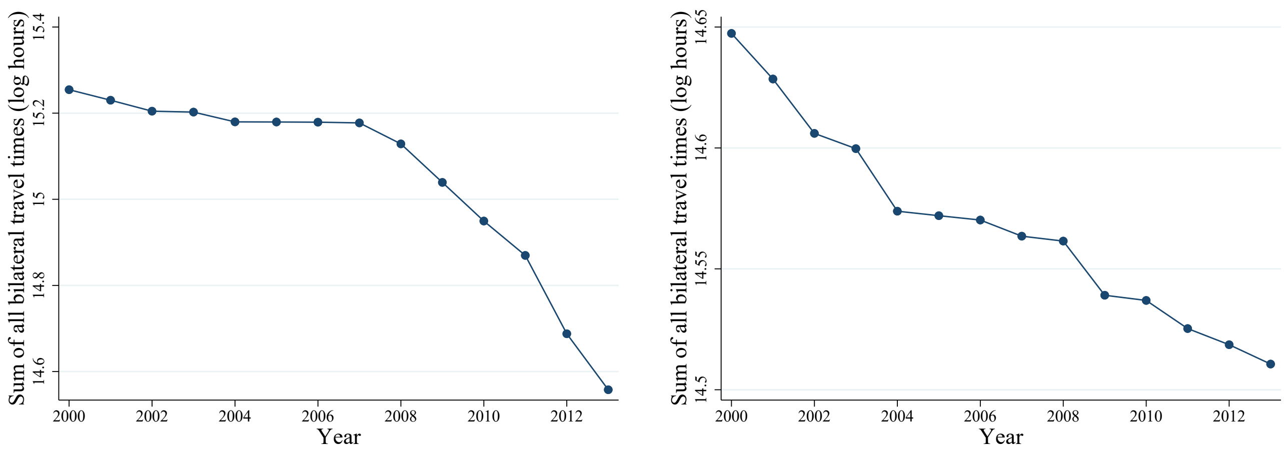
3.4 Evolution of Aggregate Travel Times in China

The large transport-infrastructure improvements documented above naturally translate in a reduction of aggregate bilateral travel times. Figure 4 shows the evolution of the (log) sum of bilateral travel times in hours between all pairs of 330 considered Chinese prefectures for each year between 2000 and 2013. Note that what we plot there is an *unweighted* sum of travel times.[18] Formally, Figure 4 displays $\log(\sum_{i=1}^{330}\sum_{j=1}^{330} d_{ijt}^{R})$ and $\log(\sum_{i=1}^{330}\sum_{j=1}^{330} d_{ijt}^{C})$ for the transport of people and goods, respectively, for every year $t = 2000, ..., 2013$.

Figure 4 portrays the differences for the transportation of people versus goods on the network when imposing the aforementioned rules about speeds and mode choice. Overall, the introduction of the HSR system from 2008 onwards had a big impact on travel times for people; see Panel (a) in Figure 4. This is represented by the almost kinked schedule. At the same time, we observe a much more linear decline in the sum of bilateral travel times throughout the sample period for goods; see Panel (b) in Figure 4.

Figure 4: Evolution of (Unweighted) Travel Times (2000 – 2013)

(a) Aggregate Travel Times for People, (Log Hours) (b) Aggregate Travel Times for Goods, (Log Hours)



Notes: Aggregate travel times are defined as the (log) sum of bilateral travel times in hours between all pairs of 330 considered Chinese prefectures for each year: $\log(\sum_{i=1}^{330}\sum_{j=1}^{330} d_{ijt}^{v})$, for $v = \{R, C\}$. Travel is done at the preferred mode and at statutory speed limits for roads and average recorded speeds for railways.

[18] Clearly, effective sums of travel times depend on residence and mobility choices which are disregarded in Figure 4.

4 Theoretical Framework

In this section, we outline a multi-region general-equilibrium model that is amenable to studying the economic effects of transport infrastructure improvements in China within the last two decades. Locations differ according to (time-varying) amenity and productivity fundamentals. These fundamentals evolve according to a dynamic amenity- and technology-diffusion process that crucially depends on each location's integration in the transport network. Locations interact in product markets through (costly) trade in goods and in labor markets through (costly) migration. Trade and migration frictions are directly affected by the transport network, which implies feedback effects in prices and consumption patterns through changes in trade volumes and individual mobility. In what follows, we describe the elements of the theoretical framework in detail.

4.1 Setup

Consider a world composed of a finite number of pairwise-connected regions $i \in S = \{1, \ldots, s\}$. Time is discrete and indexed by t, which refers to years. Available land in each location is allocated to residential (R) and commercial (C) use and measured by the land densities $H_i^R, H_i^C > 0$, respectively. We assume that income from land is owned by absentee landlords, who receive rents on commercial and residential land and consume only outside of the considered world.[19] The world economy is populated by a total of $\bar{L}$ individuals who are endowed with one unit of labor each, which they supply inelastically. After deciding on their residence location, individuals produce in the location in which they live and consume.

4.2 Individual Preferences

Each period, the utility of an individual ω residing in i at period t is defined over the consumption of a bundle of differentiated products (C_{it}), residential land use (H_i^R),[20] local endogenous amenities (a_{it}), an idiosyncratic utility shock (ε_{it}^U), and bilateral moving costs (κ_{ijt}). Formally, individual utility takes the following Cobb-Douglas form:

$$U_{ijt}(\omega) = \left(\frac{C_{it}(\omega)}{\alpha}\right)^{\alpha} \left(\frac{H_i^R(\omega)}{1-\alpha}\right)^{1-\alpha} \frac{a_{it}}{\kappa_{ijt}} \quad \varepsilon_{it}^U(\omega), \tag{1}$$

[19]This assumption ensures that neither the individuals' nor firms' location decisions are affected by distributional assumptions regarding landlords.

[20]We do not use a time index on H_i^R to indicate that it will be fixed in the data.

where $\alpha \in (0,1)$ denotes the Cobb-Douglas share of goods consumption in utility. Endogenous amenities, a_{it}, are determined by a dynamic process of the form

$$a_{it} = \exp\left(\sum_{j\in S} \mathbb{W}^R_{ijt-1} \log \frac{L_{jt-1}}{H^R_j}\right)^{-\lambda} \varepsilon^R_{it}, \tag{2}$$

where (L_{jt-1}/H^R_j) is the density of individuals over residential land in j and year $t-1$, $\mathbb{W}^R_{ijt-1} = (0,1)$ is a weighting scalar whose value rises with the transport-network-connectedness between i and j and which adheres to $\sum_{j\in S} \mathbb{W}^R_{ijt-1} \leq 1$ for all i and t. Endogenous amenities decline with the people's-transport-network-weighted residential population density in logs, $\left(\sum_{j\in S} \mathbb{W}^R_{ijt-1} \log \frac{L_{jt-1}}{H^R_j}\right)$, as long as the congestion-externality parameter $\lambda > 0$. The definition of the weighting scalar $\mathbb{W}^R_{ijt-1}$ will be discussed in detail in Section 5.4. The parameter ε^R_{it} is a location- and time-specific amenity shifter.

Individuals choose a residence that maximizes their utility in (1), after learning their individual preference shock, $\varepsilon^U_{it}(\omega)$. The latter governs the dispersion of utility and, hence, absorbs unobserved preferences for location choices that cannot be explained by income and price differences. We assume that $\varepsilon^U_{it}(\omega)$ is drawn from a Fréchet distribution with shape parameter $1/\Omega$.[21]

Individuals consume residential housing, H^R_i, and a continuum of tradable varieties ρ, which are combined in a consumption bundle, C_{it}.[22] Their indirect utility is defined as

$$u_{ijt}(\omega) = \tilde{u}_{it} \frac{\varepsilon^U_{it}(\omega)}{\kappa_{ijt}} \quad \text{with} \quad \tilde{u}_{it} = \frac{a_{it} w_{it}}{P^\alpha_{it} r^{R\,1-\alpha}_{it}} = a_{it} y_{it}, \tag{3}$$

where $\tilde{u}_{it}$ is the indirect utility that an individual derives from residing in location i, which is independent of factors that determine the ex-ante choice to reside there.[23] $\tilde{u}_{it}$ is a function of local amenities (a_{it}), income (w_{it}) and the corresponding housing and consumer prices (r^R_{it} and P_{it}, respectively). We define the goods-consumption price index at location i and time t as $P_{it} = \left[\int_0^1 p_{it}(\rho)^{1-\sigma} d\rho\right]^{\frac{1}{1-\sigma}}$. For later convenience, we introduce y_{it} to denote real income per capita.

[21] The probability that individual ω in location i at time t draws an idiosyncratic preference smaller than z is given by $\Pr(\varepsilon^U_{it}(\omega) \leq z) = \exp(-z^{-(1/\Omega)})$. We assume that $\varepsilon^U_{it}(\omega)$ is i.i.d. across locations, individuals and time.

[22] We assume that $C_{it} = [\int_0^1 c_{it}(\rho)^{\frac{\sigma-1}{\sigma}} d\rho]^{\frac{\sigma}{\sigma-1}}$ according to constant-elasticity-of-substitution (CES) preferences, where $\sigma > 0$ is the elasticity of substitution.

[23] In other words, this assumes that bilateral moving costs, κ_{ijt}, and the idiosyncratic preference for that location, $\varepsilon^U_{it}(\omega)$, are irrelevant for the location-specific indirect utility $\tilde{u}_{it}$.

4.3 Technology and Production

Firms in i at t use two production factors, labor (L_{it}) and commercial land (H_i^C), to produce output units (Q_{it}) of product varieties ρ under a Cobb-Douglas technology:[24]

$$Q_{it}(\rho) = z_{it}(\rho) T_{it} L_{it}(\rho)^{\mu} H_i^C(\rho)^{1-\mu}, \tag{4}$$

where $\mu \in (0,1)$ denotes the labor share in production. Output depends on production inputs and a firm's total-factor productivity, which is determined by two components. The first component is a variety-specific exogenous productivity parameter, z_{it}, that is drawn from a Fréchet distribution with shape parameter $\theta > 0$. The second component is an endogenous technology part, T_{it}, which is determined as

$$T_{it} = \exp\left(\sum_{j\in S} \mathbb{W}^C_{ijt-1} \log \frac{L_{jt-1}}{H_j^C}\right)^{\gamma} \varepsilon^C_{it}. \tag{5}$$

where $\mathbb{W}^C_{ijt-1} = (0,1)$ is a weighting scalar whose value rises with the transport-network-connectedness between i and j and which adheres to $\sum_{j\in S} \mathbb{W}^C_{ijt-1} \leq 1$ for all i and t, as was the case with $\mathbb{W}^R_{ijt-1}$. Endogenous technology increases with the goods'-transport-network-weighted population density in logs through $\left(\sum_{j\in S} \mathbb{W}^C_{ijt-1} \log \frac{L_{jt-1}}{H_j^C}\right)$ as long as $\gamma > 0$. The latter is an agglomeration parameter that governs the technology diffusion through goods-transport networks, and ε^C_{it} is a location- and time-specific technology shifter. As for $\mathbb{W}^R_{ijt-1}$, Section 5.4 will provide details on the design of $\mathbb{W}^C_{ijt-1}$.

After learning their exogenous productivity draw, z_{it}, firms maximize their profits by choosing the level of employment and commercial land through

$$\max_{L_{it}(\rho), H_i^C(\rho)} \quad p_{it}(\rho) \;\; z_{it}(\rho)\; T_{it}\; L_{it}(\rho)^{\mu}\; H_i^C(\rho)^{(1-\mu)} - w_{it} L_{it}(\rho) - r^C_{it} H_i^C(\rho),$$

where $p_{it}(\rho)$ is the price a firm charges for a product that is sold in i at time t. The model structure implies that $p_{it}(\rho)$ is directly proportional to unit costs and inversely proportional to a location's total-factor productivity, which follows the basic price-productivity relation as outlined in Eaton and Kortum (2002). Then,

$$E[p_{it}(\rho)] = \frac{o_{it}}{T_{it} z_{it}} \quad \text{with} \quad o_{it} = \left[\frac{(1-\mu)^{\mu-1}}{\mu^{\mu}}\right] {r^C_{it}}^{1-\mu} w^{\mu}_{it} = \frac{1}{\mu} w_{it} \left(\frac{L_{it}}{H_i^C}\right)^{1-\mu}, \tag{6}$$

where o_{it} denotes the unit costs in location i at time t which firms take as given.

[24] We use a time index on L_{it} but not on H_i^C to indicate that the former will change from period to period, while the latter will be fixed in the data.

4.4 Trade

Bilateral sales from location i to location j are subject to iceberg-transport costs, $\zeta_{ijt} \geq 1$. The price of a good produced in i and consumed in j is $p_{ijt} = p_{it}\zeta_{ijt} = \frac{o_{it}\zeta_{ijt}}{T_{it}z_{it}}$. Then, bilateral trade shares take on the well-known gravity form as in Eaton and Kortum (2002):

$$\pi^C_{ijt} = \frac{T_{it}[o_{it}\zeta_{ijt}]^{-\theta}}{\sum\limits_{k\in S} T_{kt}[o_{kt}\zeta_{kjt}]^{-\theta}}, \ \forall i,j \in N, \tag{7}$$

where π^C_{ijt} is the share of expenditures in j on i. The price index in location j at time t is defined as

$$P_{jt} = \bar{p}\left[\sum_{k\in S} T_{kt}[o_{kt}\zeta_{kjt}]^{-\theta}\right]^{-\frac{1}{\theta}}, \tag{8}$$

where $\bar{p}$ is a normalizing constant.

4.5 Migration

Bilateral migration from location i to location j is subject to migration costs $\kappa_{ijt} \geq 1$. The indirect utility in (3) is monotonically increasing with the idiosyncratic Fréchet-distributed preference shock $\varepsilon^U_{it}(\omega)$. Then, the indirect utility for an individual residing in i at t is itself Fréchet-distributed.[25] Analogously to trade shares, the share of individuals coming from j and choosing to reside in i can be expressed in the well-known gravity form as

$$\pi^R_{ijt} = \frac{(a_{it}w_{it})^{1/\Omega}(\kappa_{ijt}P^{\alpha}_{it}r^{R\,1-\alpha}_{it})^{-1/\Omega}}{\sum\limits_{k\in S}(a_{kt}w_{kt})^{1/\Omega}(\kappa_{kjt}P^{\alpha}_{kt}r^{R\,1-\alpha}_{kt})^{-1/\Omega}} = \frac{\tilde{u}^{1/\Omega}_{it}\kappa^{-1/\Omega}_{ijt}}{\sum\limits_{k\in S}\tilde{u}^{1/\Omega}_{kt}\kappa^{-1/\Omega}_{kjt}}. \tag{9}$$

Summing π^R_{ijt} across migration locations j for a given residence, we obtain the probability that individuals reside in location i:

$$\pi^R_{it} = \frac{\sum\limits_{j\in S} L_{ijt}}{\bar{L}} = \frac{L_{it}}{\bar{L}} = \frac{\tilde{u}^{1/\Omega}_{it}\sum\limits_{j\in S}\kappa^{-1/\Omega}_{ijt}}{\sum\limits_{j\in S}\sum\limits_{k\in S}\tilde{u}^{1/\Omega}_{kt}\kappa^{-1/\Omega}_{kjt}}. \tag{10}$$

4.6 Equilibrium

Profits and utility are maximized within each period, as neither individuals nor firms are forward-looking. This implies that both firms and individuals are uninformed about the endogenous

[25] Given that, utility is distributed as $G_{it}(u) = \exp(-\Phi_{it}u^{-1/\Omega})$, where $\Phi_{ijt} = (a_{it}w_{it})^{1/\Omega}(\kappa_{ijt}P^{\alpha}_{it}r^{R\,1-\alpha}_{it})^{-1/\Omega}$.

dynamic process that governs the amenity and technology diffusion. Decisions on the improvements of the transport network, which directly impact the amenity and technology diffusion as well as trade and migration costs, are taken from an absentee central planner and, hence, are not anticipated.

An equilibrium requires that goods and factor markets clear in each period. More formally, for any population and land vectors with elements $\{L_{it},\, H_i^R,\, H_i^C\}$ for all $\{i, t\}$, any trade- and migration-cost vectors with elements $\{\kappa_{ijt},\, \zeta_{ijt}\}$, as well as any transport-network-weighting scalars $\{\mathbb{W}_{ijt}^R,\, \mathbb{W}_{ijt}^C\}$ for all $\{i, j, t\}$, a competitive equilibrium is a set of endogenous vectors with elements $\{L_{it},\, w_{it},\, u_{it},\, P_{it},\, a_{it},\, T_{it},\, r_{it}^R,\, r_{it}^C\}$ for all $\{i, t\}$, such that for all locations i in each time period t we have:

1. **Goods-market Clearing**: Total revenues in a location i are equal to the value of all purchases from it. Thus,
$$w_{it} L_{it} = \sum_{j \in S} \pi_{ijt}^C w_{jt} L_{jt}. \tag{11}$$

2. **Population Accounting**: The number of residents in all locations is equal to the total population $\sum_{i \in S} L_{it} = \bar{L}_t$, which we normalize to be constant. In view of equation (10), this implies
$$L_{it} = \sum_{j \in S} \pi_{ijt}^R L_{jt}. \tag{12}$$

3. **Land-market Clearing**: The residential and commercial land markets in each location i are in equilibrium, so that land is assigned to the highest bidder. Hence, equilibrium land rents are given as
$$r_{it}^R = [1 - \alpha] w_{it} \frac{L_{it}}{H_i^R} \quad \text{and} \quad r_{it}^C = \left[\frac{1 - \mu}{\mu}\right] w_{it} \frac{L_{it}}{H_i^C}. \tag{13}$$

We can manipulate the system of equations presented above and reduce it to one that determines the equilibrium distribution of wages, w_{it}, and employment levels, L_{it}, in all locations. Equilibrium wages are determined using the goods-market-clearing condition (11) and equilibrium employment is based on migration shares in equation (10).[26]

To guarantee that the solution to the system of equations exists, is stable and unique, we establish two crucial conditions:[27]

[26] Details are presented in the Section C of the Appendix.

[27] Section D and E of the Appendix present a detailed derivation of the uniqueness and stability conditions.

Uniqueness There exists a unique (up-to-scale) solution of equilibrium wages and equilibrium employment, if

$$\frac{\alpha}{\theta} \leq (1-\mu\alpha) + \Omega. \tag{14}$$

This condition is intuitive. It states that agglomeration forces expressed in local production externalities ($\frac{\alpha}{\theta}$) may not dominate the two dispersion forces: land-availability restrictions $(1-\mu\alpha)$ and the variance of local preferences (Ω). This result is in line with the literature on quantitative spatial general-equilibrium models (see, Allen and Arkolakis, 2014; Desmet et al., 2018; Allen and Donaldson, 2018).

Stability An equilibrium requires that both the amenity diffusion in (2) as well as the technology diffusion in (5) reach an equilibrium level in the long-run upon a shock in infrastructure. For either process, this is the case whenever (i) the infrastructure-accessibility weights $\mathbb{W}^R_{ijt}$ and $\mathbb{W}^C_{ijt}$ are properly normalized so that their sum for each prefecture and time period is bounded, and (ii) the adjustment-cost parameters λ and γ are finite. To ensure the former, we normalize the entries of both $\mathbb{W}^R_{ijt}$ and $\mathbb{W}^C_{ijt}$ such that the maximum of their row sums equal unity, i.e., $\max_i \sum_{j\in S} \mathbb{W}^{(v)}_{ijt} = 1$ for $v \in \{C, R\}$. We adopt a maximum-row-sum normalization rather than the more standard row-sum normalization as it retains differences in absolute centrality of the locations (see Kelejian and Prucha, 2010).[28]

5 Calibration

In 2000, the baseline year of data to which the model is calibrated, China was composed of 348 prefecture-level regions. We use information on the administrative boundaries of Chinese prefectures for that year from the China Data Center at University of Michigan.[29] Throughout the analysis, we hold prefecture-level borders constant as of the year 2000 and measure all model variables accordingly. In the analysis, we consider a China of 330 prefectures. We arrived at this number as follows. First, we dropped altogether 12 prefectures which include ones that are not part of what is often called *mainland China* (i.e., Hongkong and Macao), prefectures on Hainan Island (as there is no road connection to mainland China), and all prefectures in the Tibet Autonomous Region (mainly for data restrictions). Of the 336 remaining prefectures we merged altogether 11 into 5 so as to obtain greater-city quasi-prefectures within Beijing (originally

[28]We outline the stability condition for the amenity- and technology-diffusion processes in further detail in Section E of the Appendix.

[29]Note that the Chinese Data Center's website has been closed and is no longer accessible since September 2018. We acquired the data on Chinese administrative boundaries in November 2016.

two prefectures), Baoshan (originally two prefectures), Chongqing (originally three prefectures), Shanghai (originally two prefectures), and Tianjin (originally two prefectures).

While the focus of this study is on Chinese prefectures, we model China as a large open economy, given the importance of its foreign trade. However, there is no need for going into detail with regard to the Rest of the World (RoW) for the present purpose. We aggregate the activity outside of the 330 individual prefectures into one large RoW. Hence, the total number of regions covered is 331 in this paper. The population in the RoW in a year (e.g., in 2000) is just the world population net of the 330 prefectures' population. To determine the per-capita wage in the RoW, we use the following data moments. First, we compute the population-weighted average nominal wage in China as covered (across the 330 prefectures) in a year (e.g., 2000). Second, we compute the ratio of the OECD to the Chinese price level given OECD data.[30] We then multiply this ratio with the Chinese per-capita average wage and thereby back out the per-capita wage level for the RoW, assuming that the RoW is largely the OECD.

To compute the quantitative multi-region long-run general equilibrium for any baseline calibration, we need the parameters contained in the equations above and summarized in Table 1. Apart from parameters that are common to all regions in the model, these are residential and commercial land endowments $\{H_i^R, H_i^C\}$, distributions of amenities and technology, $\{T_{it}, a_{it}\}$ to estimate the respective diffusion parameters $\{\gamma, \lambda\}$, network transmission (or spillover) weights captured by the scalars $\{\mathbb{W}_{ijt-1}^R, \mathbb{W}_{ijt-1}^C\}$ measured in 2000 (and, in the counterfactual in 2013), as well as interregional trade- and migration-cost parameters $\{\kappa_{ijt}, \zeta_{ijt}\}$, also measured in 2000 (and in 2013). Table 1 lists the parameters and gives a brief explanation of how they are assigned and chosen. Three of the parameter values are based on previous work, namely $\{\sigma, \Omega, \theta\}$, while all others are either estimated or calibrated by the authors. In what follows, we briefly describe our choices for the adopted parameter values in more detail.

5.1 Elasticity-of-substitution Parameter σ

We assume a value for the elasticity of substitution between varieties of $\sigma = 4$ in the model. This value is motivated by earlier work. For instance, Bernard et al. (2003) report a value of $\sigma = 3.8$, and Broda and Weinstein (2006) an average one of $\sigma = 4$ (based on price and expenditure data measured at the three-digit trade-classification product level).

[30] Source: https://data.oecd.org/price/price-level-indices.htm. We could have used nominal wage ratios instead (since price ratios to some extent also reflect productivity ratios), but the difference in price and wage ratios is marginal.

Table 1: Calibration Overview

Parameters common to all locations		
1. Preferences & Evolution of Amenities		
$\sigma = 4$	Elasticity of substitution.	Bernard et al. (2003)
$\alpha = 0.850$	Consumption share in utility (non-land share).	Own computation
$\lambda = 0.133$	Elasticity of end. amenities w.r.t. past pop. density.	Own estimation
$\Omega = 0.5$	Elasticity of migration flows w.r.t. income.	Monte et al. (2018)
2. Technology & Evolution of Productivity		
$\mu = 0.800$	Labor share in production (non-land share).	Own computation
$\theta = 4$	Trade elasticity and dispersion of technology.	Simonovska and Waugh (2014)
$\gamma = 0.306$	Elasticity of technology w.r.t. past pop. density.	Own estimation
3. Migration- & Trade-cost Elasticities		
$\phi^R = 0.020$	Elasticity of travel time to migration costs.	Own estimation
$\phi^C = 0.040$	Elasticity of travel time to trade costs.	Own estimation
Location-specific parameters		
1. Land Endowments		
H_i^R	Residential land mass in location i.	Geofabrik, OSM
H_i^C	Commercial land mass in location i.	Geofabrik, OSM
2. Initial Distributions		
T_{it}	Initial technology distribution.	Own computation
a_{it}	Initial amenity distribution.	Own computation
3. Network Parameters & Weighting Scalars		
d_{ijt}^R	Bilateral travel time for people	Network Analyst, ArcGIS
d_{ijt}^C	Bilateral travel time for goods	Network Analyst, ArcGIS
$\mathbb{W}_{ijt}^R$	Weighting scalar governing amenity diffusion.	Own computation
$\mathbb{W}_{ijt}^C$	Weighting scalar governing technology diffusion.	Own computation
4. Migration & Trade Costs		
κ_{ijt}	Migration costs	Own estimation
ζ_{ijt}	Trade costs	Own estimation

Note: OSM refers to Open Street Map.

5.2 Land Endowments (H_i^R, H_i^C)

The total land area of each prefecture is measured using ArcGIS. However, as our theoretical framework distinguishes between residential and commercial land markets, we need additional information on land use in each prefecture. This information is provided by the Open Street Map (OSM) GeoFabrik Database.[31] Land-use data for the RoW are obtained by assuming that the patterns observed in the EU28 countries are the same for all of the RoW. For this, we use data from the European Union's Buildings Database.[32]

5.3 Cobb-Douglas Shares in Preferences and Production (α, μ)

We compute the non-land share in utility (α) and the non-land share in production (μ) using data on employment, land use as well as residential and commercial land rents for a total of 296 Chinese prefectures.[33] Rearranging the expressions for equilibrium land rents in equation (13) yields

$$\alpha = 1 - \left[\frac{r_{it}^R H_i^R}{w_{it} L_{it}}\right] \quad \text{and} \quad \mu = \left[1 + \left[\frac{r_{it}^C H_i^C}{w_{it} L_{it}}\right]\right]^{-1}.$$

Using the aforementioned data in the latter expressions, we obtain parameter estimates of $\widehat{\alpha} = 0.850$ and $\widehat{\mu} = 0.800$ as listed in Table 1. These results are in line with estimates from the China Household Income Project (CHIP) Database in 2013, which reports an average share of housing expenditure of $(1-\alpha)$= 0.17. Regarding the value of μ, the literature shows that China's labor share in value-added has continuously declined over the past decades. Bai and Qian (2010) document this evolution and find a labor share of 0.63 in the year 2000. While our value is considerably higher with $\mu = 0.8$, it has to be noted that in the given model context it represents not only the labor share but also the non-land share in production. For that reason, it is reasonable to assume that μ is larger than the observed labor-cost share in value-added in Bai and Qian (2010) .

5.4 Infrastructure Data and Network (Spillover) Weights (d_{ijt}^R, d_{ijt}^C, $\mathbb{W}_{ijt}^R$, $\mathbb{W}_{ijt}^C$)

As outlined in Section 3, we use hand-collected and digitized information on the size, average speed and exact location of all road (highways, provincial-level roads, and prefecture-level) and

[31] The GeoFabrik Database is updated every day and we extracted the information on November 11, 2018. Hence, residential and commercial land endowments (H_i^R, H_i^C) are measured in 2018 and assumed constant at those levels throughout the analysis. The data are available at `https://download.geofabrik.de/asia/china.html`.

[32] See `https://ec.europa.eu/energy/en/eu-buildings-database`.

[33] For details on the data sources regarding land rents, see Section F of the Appendix.

rail (high-speed and standard) segments in China, specifically for the years 2000 and 2013. Based on this comprehensive transport-network information, we derive the travel time on the network between any two prefectures i and j in hours for each year for people, d^R_{ijt}, and goods, d^C_{ijt}, using the Network Analyst from ArcGIS.

We add, to the within China travel times information, the travel times between prefectures and the Rest of the World (RoW). The latter are defined as the travel times on the network to the nearest of the six largest Chinese sea ports – Shanghai, Shenzhen, Ningbo, Guangzhou, Qingdao, and Tianjin – plus the travel time from that port to the centroid of the population-weighted RoW countries.[34] By this token, the centrality relative to the RoW is impacted by Chinese transport-network improvements.

The cross-prefecture diffusion of amenities and technology through the network are governed by the weights $\mathbb{W}^R_{ijt-1}$ and $\mathbb{W}^C_{ijt-1}$, respectively. We normalize intra-regional travel times to unity, whereby $d^R_{iit} = 1$ and $d^C_{iit} = 1$ for all $\{it\}$. Moreover, we define the degree centrality of network V for region i at time $t-1$ as $d^V_{it-1} = \sum_{j \in S}(d^V_{ijt-1})^{-1}$. The shorter the average travel time to other prefectures of i is, the bigger is this centrality value. Similarly, we define the maximum centrality in the network at time t as $d^V_{t-1} = \max_{i \in S} d^V_{it-1}$ (i.e., the maximum sum of inverse travel times for any prefecture i at $t-1$) for $V \in \{R, C\}$. Finally, we define $\mathbb{W}^V_{ijt-1} = \frac{(d^V_{ijt-1})^{-1}}{d^V_{t-1}}$ for $V \in \{R, C\}$. Note that the scalar normalization by d^V_{t-1} is innocuous, as it merely leads to a transformation of the scalars $\{\lambda, \gamma\}$ (Kelejian and Prucha, 2010). However, what is key is that $\mathbb{W}^V_{ijt-1}$ increases in the connectivity in terms of d^V_{ijt-1} as captured by the inverse travel time of goods and people. Since with this approach the maximum row sum of $\mathbb{W}^V_{ijt-1}$ equals unity, finite scalars $\{\lambda, \gamma\}$ and a finite population density guarantee stability of the technology- and amenity-dissipation processes (see Section 4.6).

Given data on population density (L_{jt-1}/H^V_j) and transport-infrastructure weights $\mathbb{W}^V_{ijt-1}$ for $V \in \{R, C\}$ and all $\{i, j, t-1\}$ in conjunction with information on technology T_{it} and amenities a_{it} for all $\{i, t\}$ – we will outline below how to obtain measures thereof – one could estimate the parameters γ and λ when treating ε^V_{it} as a residual.

5.5 Instrumental Variables

One key concern when analyzing transport networks is that connections between places are not random. For transport networks, central planners decide on the location of new transport segments (road or railway) to maximize some objective function (e.g., connecting large centers, providing

[34]For the travel time from any Chinese port to the centroid of the population-weighted RoW we assume an average travel speed of 25km/h.

access to formerly isolated regions, etc.), subject to some initial distribution of the population and features of a country's geography. The explicit modeling of this objective function is outside of the focus of the present paper. However, for the identification of parameters governing migration and trade costs, κ_{ijt} and ζ_{ijt}, as well as the diffusion parameters for technology and amenities, γ and λ, we still have to take into account the mentioned endogeneity. In order to estimate the impact of travel time on mobility and trade costs, as well as on amenity and productivity diffusion, we propose three IV strategies either of which aims at avoiding the bias associated with endogenous road-/railway-infrastructure placements in the estimation of some of the structural parameters. We provide details of these strategies in the remainder of this subsection.

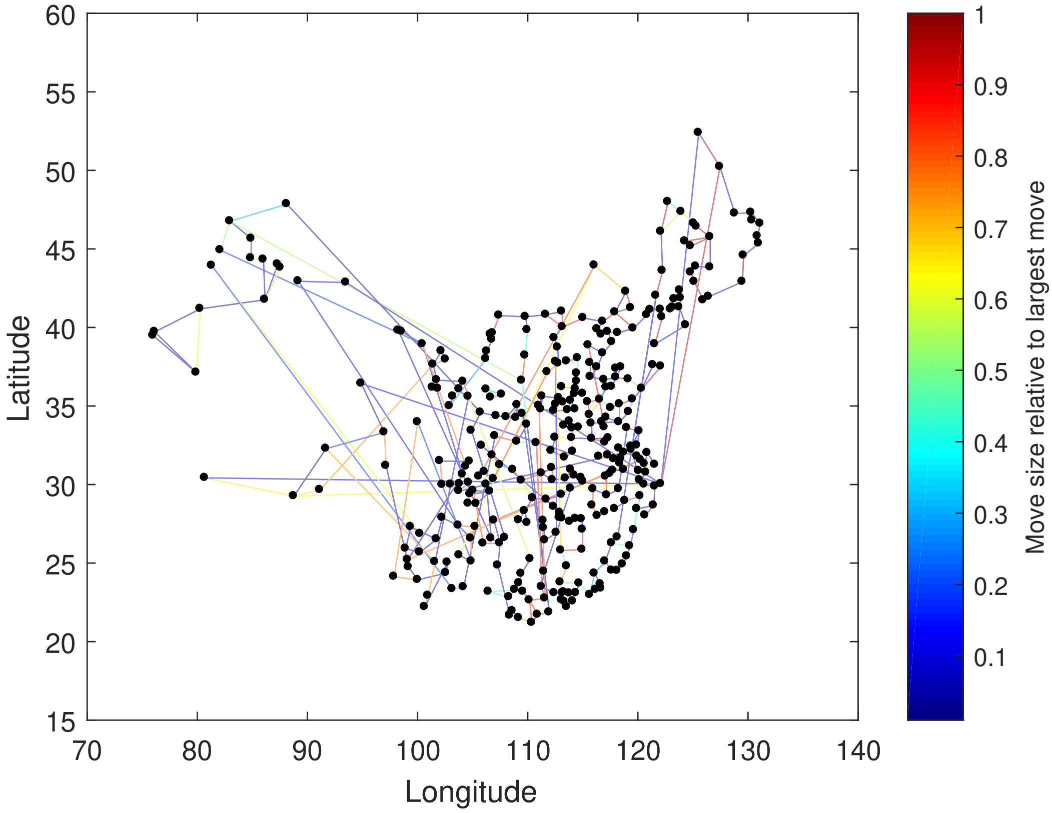
The (symmetric) matrix of travel times between the 331 spatial units considered (330 prefectures and one RoW) consists of two components of interest: one inter-prefectural and one between the prefectures and the RoW. In this subsection, we focus on the component pertaining to inter-prefectural transport. The reason is that travel times between the prefectures and the RoW themselves consist of two components: one pertaining to travel times between the prefectures to the six ports (which belong in prefectures) and one to the travel time between the ports and the RoW. The latter does not change in its prediction depending on a particular IV used (nor in any considered counterfactual equilibrium). We hold this component constant throughout the analysis. However, inter-prefectural travel-time predictions depend on the IV in use. By the latter token, also the predicted travel times of those prefectures which do not host one of the six major ports in China depend on the IV, and they change in counterfactual equilibrium. Towards mapping the information generated by one of the IV approaches into travel times we refer the reader to Appendix B, in particular, to Step 4.

Geography-based Optimal Transport Network

As a first strategy, we propose instrumenting travel time in the observed network by the solution of the classical Monge-Kantorovich optimal-transport problem as outlined in Monge (1781) and Kantorovitch (1958). The corresponding algorithm is about transporting some mass in the most efficient way between a given set of locations. In our case, the latter are the Chinese prefectures and the (weighted centroid of the) RoW. One could think of a natural candidate for the mass to be transported as population, in particular, when measured at some historical point in time prior to the sample period. However, population at a place is endogenous in the proposed model and there is a risk that population density is jointly generated with the transport network by some deep parameters. Therefore, we decided to use a prefecture's and the weighted RoW's longitude as the mass to be transported, as this is fundamentally fixed. The use of longitude as the transported mass nicely captures the increase in population density as one moves eastwards from China's West, i.e., towards the coast. We provide a detailed description of this approach in Appendix B.

In short, we derive the optimal transport network in three steps. First, using a grid of 4,000,000 equally-sized and -spaced cells over China (of approx. 3km^2 each), we predict the cost of building a transport segment in each cell using observed geographical and year-2007 network data.[35] Second, using Dijkstra's (1959) algorithm, we compute a path of minimum costs between any two Chinese prefectures. In a final step, we select among all cost-minimizing paths the ones that solve the Monge-Kantorovich optimal-transport problem using longitude as the local mass to be transported. Figure 5 displays the resulting network. For later usage in the IV estimation, we additionally compute the degree centrality of the optimal network for each prefecture as $\sum_{j \in S} d_{ijt-1}^{-1}$, where d_{ijt-1}^{-1} is the inverse of the Monge-Kantorovich network distance.

Figure 5: Monge-Kantorovich Longitude-based Optimal Transport Network



This approach has several merits. First, it predicts significantly more connections between the nodes (here, prefectures and the RoW) in a network than a minimum-spanning-tree approach would do.[36] This leads to a higher predictive power in explaining the observed network for the

[35]We choose the year-2007 network data for the cost prediction as this is the year in the middle of the sample period 2000-2013.

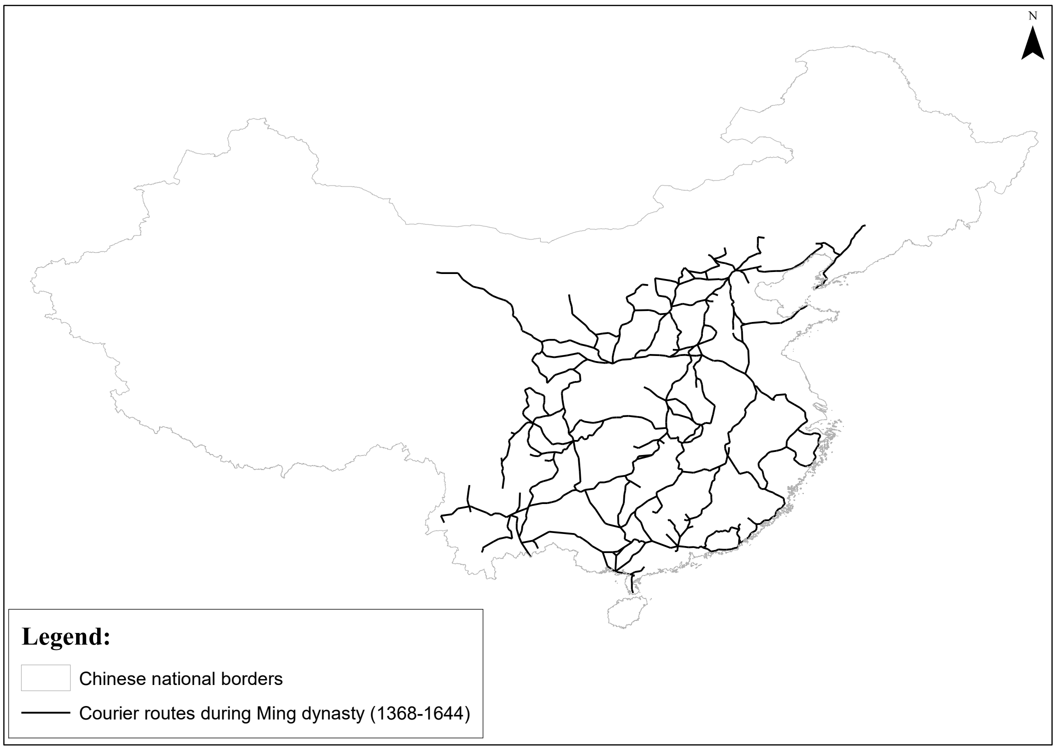
[36]The minimum-spanning-tree approach would minimize the number of connections between the 331 regions under the constraint that all regions can be reached and that no cycles emerge.

Monge-Kantorovich predicted network than for a minimum-spanning-tree network. The latter is important for the first-stage process in a two-stage IV model. Second, it allows to predict an intensity-of-connection usage and, hence, a connection typology (here, of highways, provincial-level, prefecture-level roads; and of high-speed and standard-railway lines), which would not be the case for a standard minimum-spanning-tree network.

Historical Transport Network

As a second IV strategy, we instrument the travel time on the observed network using the travel time on the network of ancient courier routes used during the Ming dynasty (Berman and Zhang, 2017). Figure 6 displays the placement of the courier routes. The Ming dynasty was the dynasty ruling China from 1368 to 1644. Its territory spanned from Beijing in the North to the Pearl River Basin in the South, and from the coast to Tibet in the West. The courier routes where built over time during the dynasty's ruling time, and they covered the entirety of the Ming territory. We compute the travel distance between each location on the network using ArcGIS network analyst. Prefectures are assumed to be connected in this network if located within 50 kilometers of the nearest courier route. Again for later usage in the IV estimation, we compute the network centrality of each prefecture on the ancient courier routes. Figure 6 illustrates this network.

Figure 6: Courier Routes During the Ming Dynasty (1368-1644)



Note that this approach is different from the previous as well as the subsequent one in that it does not utilize any information on travel speeds. Regarding distances between prefectures and the RoW we use the combination of distances between each prefecture and the hosting prefectures of one of the six major ports in conjunction with great-circle distances between the latter and the centroid of the RoW. In that, this IV differs fundamentally from the other ones by being based on geographical distances alone.

Historical-population-based Optimal Transport Network

Finally, we propose a third instrumentation strategy for the travel time on the observed transport network. Instead of using longitude as a location mass to be transported as with the first approach, we define a location's mass using its population count based on the 1953 Population Census.[37] Relative to the former two, this approach is mixed in that it combines both natural (geographical) and historical information to create the IV.[38] Additionally, we compute the network centrality of each prefecture using the respective optimal transport network using 1953 local population data. As with the first approach, degree centrality is based on the sum of inverse distances of the respective network for each prefecture. Figure 7 displays the resulting network.

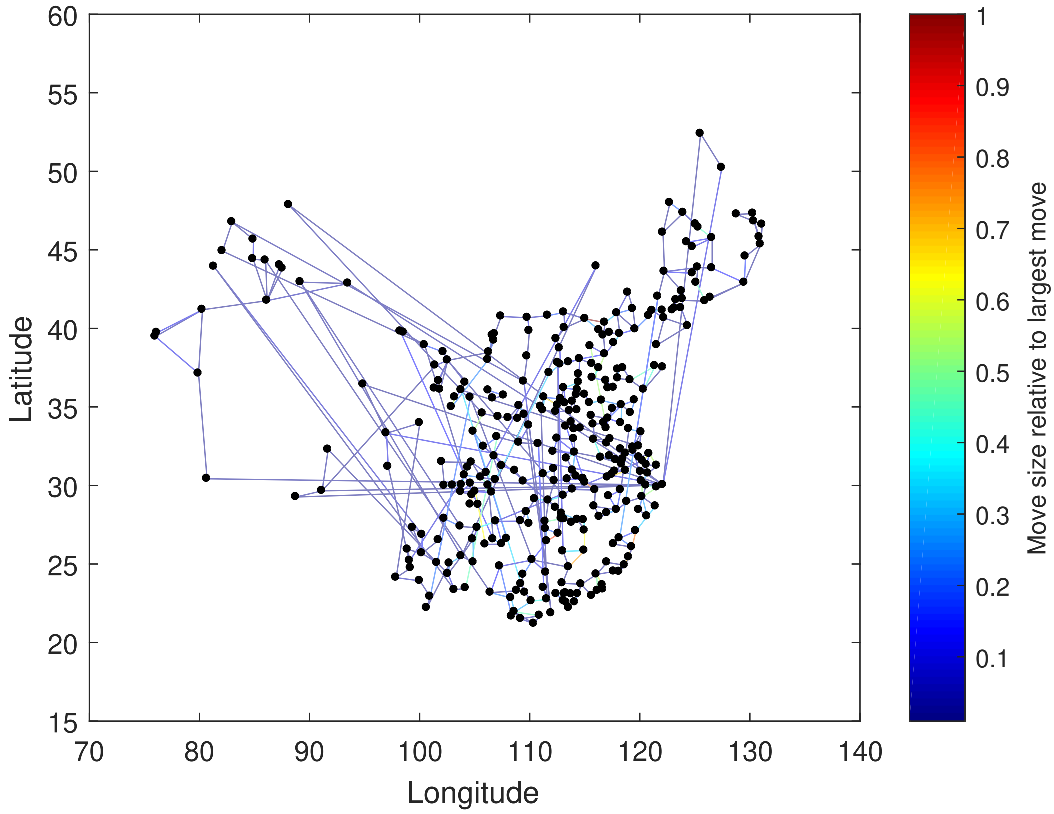
5.6 Migration-cost Parameters (Ω, κ_{ijt}, ϕ^R)

We follow Desmet et al. (2018) and Allen and Donaldson (2018) and assume a common parameter about the Fréchet dispersion of stochastic migration costs, Ω. The latter pertains to the heterogeneity in individuals' tastes for location choices. More formally, Ω describes the inverse of the elasticity of migration flows to location-specific utility levels (net of bilateral migration costs), which becomes apparent in equation (10). We assume a value of Ω on the basis of previous estimates that have been derived in the within-country context, i.e., without formal migration restrictions at country borders. We follow Desmet et al. (2018), who set a value of $\Omega = 0.5$, based on studies by Diamond (2016), Ortega and Peri (2016), Monte et al. (2018), and Fajgelbaum et al. (2019). This value is informed by data of the European Union. Two remarks are in order with regard to that choice. First of all, assuming a common value for Ω should appear less problematic

[37] The 1953 Population Census for China is available in pdf format and contains information on population counts at the provincial and county levels. We digitized the data using optical character recognition (OCR) methods. Then, we aggregated all county-level population data to the prefecture level using prefecture borders as of the year 2000.

[38] As indicated when outlining the first IV approach, we consider this approach inferior to the one where longitude is transported, as contemporaneous and historical population density are both endogenous so that IVs based on historical population counts might be an inferior IV and remedy of endogeneity of contemporaneous population counts in empirical models. In particular, this concern might be more important for population data within a century's time. Clearly, also the courier routes in the Ming dynasty as used in the second approach connected relatively densely populated places at the time. However, we would deem that concern to be less serious as the underlying population that mattered for this IV was the one of several centuries ago.

Figure 7: Monge-Kantorovich Historical-population-based Optimal Transport Network (Using Local Population in 1953)



here than in global regional studies because the regional entities considered in this study are – with the exception of one large RoW – all located within China. Second, the residence choice in China is to some extent impeded by the household registration system called *hukou*. The latter aims at reducing an excess population pressure (congestion) in large urban centers and Special Economic Zones (SEZs), while at the same time increasing the share of the population living in urban centers (see, Egger et al., 2017). Between the 1970s and the mid-2000s, the *hukou* system has undergone numerous reforms towards greater flexibility. In any case, beyond the *hukou* system, there are (almost) no cultural and virtually no language barriers for internal migrants in China, unlike for internal migrants in the European Union.

As is customary in the quantitative economic-geography literature, we model mobility frictions κ_{ijt} as an exponential function of bilateral travel times:

$$\kappa_{ijt} = \exp(\phi^R d^R_{ijt}), \tag{15}$$

where ϕ^R governs the translation of bilateral person-travel times, d^R_{ijt}, into migration frictions. Accordingly, there is a semi-log gravity equation for mobility flows to destination (residence) j

from origin i in terms of travel time between i and j in year t:

$$\log \pi_{ijt}^R = -\xi d_{ijt}^R + a_{it} + b_{jt} + \epsilon_{ijt}, \tag{16}$$

where b_{jt} is a destination- (residence-)region-time fixed effect capturing all corresponding characteristics, a_{it} is an origin-region-time fixed effect, and ϵ_{ijt} is a residual. Note that b_{jt} subsumes the denominator of equation (9) in logs, as it is constant across origin regions i within $\{jt\}$. The parameter ξ is the semi-elasticity of migration flows with respect to travel time. It is defined as $\xi = \phi^R/\Omega$, where $\Omega = 0.5$ is the assumed heterogeneity parameter associated with the Fréchet-distributed shock on individuals' utility.

Table 2: Gravity Estimation Results of Bilateral Migration

	(1)	(2)	(3)
	(a) 2SLS		
Instrument	Optimal Netw. (Long.)	Historical	Optimal Netw. (Pop. 1953)
Travel time (min)	-0.052***	-0.040***	-0.352
	(0.002)	(0.002)	(0.481)
Obs.	5512	2137	5512
Weak instr. (F-statistic)	730.61	1538.22	0.42
	(b) PPML - Control Function		
Instrument	Optimal Netw. (Long.)	Historical	Optimal Netw. (Pop. 1953)
Travel time (min)	-0.041***	-0.039***	-0.039***
	(0.003)	(0.003)	(0.003)
Obs.	3688	3685	3688
t-statistic of IV (First stage)	78.19	50.77	83.29

Notes: Optimal Netw. (Long.) refers to the instrument using the longitude-based optimal transport network. *Historical* refers to the instrument using the historical transport network. *Optimal Netw. (Pop. 1953)* refers to the instrument using the historical-population-based optimal transport network. For one endogenous regressor, one instrumental variable and a maximum relative bias of 5%, the critical value for the weak instrumentation F-statistic is 16.38 (Stock and Yogo, 2005). PPML stands for Pseudo Poisson Maximum Likelihood. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

To estimate the semi-elasticities of migration flows, we measure bilateral migration shares from the Fifth National Population Census of the year 2000.[39] The data include, for each individual, their origin and current location at the prefecture level, and the year they moved, if they did. Of

[39] Population statistics from the Fifth National Population Census of the year 2000 are a random sample of the census with a coverage of 0.95 per 1,000 people. Overall, there are 1,180,111 individuals covered.

all the individuals in the data, 10.5% had a prefecture of residence different from the prefecture of origin. We aggregate all individuals to the prefecture level (for origins and destinations) to obtain π_{ijt}^{R} as the fraction of people in j originating from i, measured in year 2000. We instrument the travel time between two locations using the three alternative IVs outlined in Subsection 5.5. Table 2 reports on the parameter estimates of equation (16) using the mentioned data and strategies in either linear two-stage least-squares (2SLS) or exponential Poisson-pseudo-maximum likelihood (PPML) estimation with a control function.[40] Overall, all instrumentation strategies lead to very similar estimates. We use the estimate in Column (2) as preferred specification as it reports the highest F-statistic when testing against weak instruments. Given $\Omega = 0.5$, we obtain an estimate of $\hat{\phi}^R = 0.020$.

5.7 Trade-cost Parameters (θ, ζ_{ijt}, ϕ^C)

To calibrate the trade-cost parameters, we start by assuming a value for the trade elasticity $\theta = 4$, following the findings in Simonovska and Waugh (2014). Compared the estimated parameter in Eaton and Kortum (2002), this parameter value was obtained allowing for a small-sample bias in trade-elasticity estimation which leads to a significantly lower mean estimate for the trade elasticity.

Then we turn to the estimation of trade costs, ζ_{ijt}. We face the difficulty that Chinese prefecture-to-prefecture trade data are not available.[41] However, in contrast to many other countries, prefecture-level trade aggregates with various regional aggregates within China and with foreign countries are available. These data can then be used for parameter identification using a model-guided approach. Prefecture-level trade aggregates are recorded in the Investment Climate Survey (ICS).[42] For the present calibration exercise, we use the survey recordings on the percentage of sales of an establishment (i) within the prefecture of its location, (ii) to other prefectures within the province of its location, (iii) to the rest of China (i.e., to all other provinces together), and (iv) to foreign countries (including Taiwan, Hongkong, and Macao). The ICS also offers information on the firms' performance from their income statements, which we employ in order

[40] With IV estimation, one splits the information in one or more endogenous variables in two parts, an exogenous and a residual part, where the latter contains the endogenous component. This is done in a first stage or step. In linear models, one can either use the predicted (fitted) endogenous regressors based on first-stage parameters instead of the measured endogenous regressors or include the first-stage residuals (as ingredients of a linear control function) along with the measured endogenous variables in the second stage. These two alternatives obtain identical point estimates. In most non-linear models, the plug-in (use of fitted endogenous regressors) estimator is not consistent, whereas the control-function approach is.

[41] This is also the case in other datasets used in quantitative multi-region general-equilibrium work (see, e.g., Desmet et al., 2018).

[42] This survey is part of the 2005 World Bank Enterprise Survey, and it collects a wide range of information on 12,400 establishments in 123 prefectures in mainland China (for the year 2004).

to aggregate sales of the sampled firms to the prefecture level.

To make use of the ICS trade data in the context of our framework, it is useful to define $\mathfrak{p}^i$ as the set of all prefectures within the same province as prefecture i, and $\mathfrak{r}^i$ as the set of all Chinese prefectures not contained in (geographically, outside of the) set $\mathfrak{p}^i$. Then, let us use the indices p^i and r^i to refer to these aggregates as destination regions. According to the model, the share of aggregate purchases from firms in i in the total expenditures of the province that i is located in (denoted by p^i), and in the total expenditures of other Chinese provinces (denoted by r^i), may be written as:

$$\pi^C_{ip^it} = \frac{\sum_{j\in\mathfrak{p}^i} T_{it}(o_{it})^{-\theta}\zeta_{ijt}^{-\theta}}{\sum_{j\in\mathfrak{p}^i}\sum_{k\in S} T_{kt}(o_{kt})^{-\theta}\zeta_{kjt}^{-\theta}}, \qquad \pi^C_{ir^it} = \frac{\sum_{j\in\mathfrak{r}^i} T_{it}(o_{it})^{-\theta}\zeta_{ijt}^{-\theta}}{\sum_{j\in\mathfrak{r}^i}\sum_{k\in S} T_{kt}(o_{kt})^{-\theta}\zeta_{kjt}^{-\theta}}. \tag{17}$$

Now, consider the ratio of ratios of expenditure on two prefectures i and i' from the same and other provinces in China. For i and i' that belong to the same province, this ratio of ratios can be written as:

$$\frac{\pi^C_{ip^it}\pi^C_{i'r^it}}{\pi^C_{i'p^it}\pi^C_{ir^it}} = \frac{\sum_{j\in\mathfrak{p}^i}\zeta_{ijt}^{-\theta}\sum_{j\in\mathfrak{r}^i}\zeta_{i'jt}^{-\theta}}{\sum_{j\in\mathfrak{p}^i}\zeta_{i'jt}^{-\theta}\sum_{j\in\mathfrak{r}^i}\zeta_{ijt}^{-\theta}}. \tag{18}$$

The left-hand-side variable of equation (18) is observed in the data from the ICS. Given a functional form for ζ_{ijt} as an exponential function of observed bilateral goods-travel times, $\zeta_{ijt} = \exp(\phi^C d^C_{ijt})$, and the assumed trade elasticity value ($\theta = 4$), the right-hand-side variable of equation (18) is known up to the scalar ϕ^C. This scalar governs the mapping of travel time into goods-trade costs. We can then employ a brute-search approach to determine the value for ϕ^C that minimizes the sum of squared distances between the left-hand side and the right-hand side in equation (18).[43] In order to avoid endogeneity concerns here, we use the predicted goods-travel times based on the Monge-Kantorovich longitude-transport IV from a first stage, $\hat{d}^C_{ijt}$. Doing so obtains a parameter estimate of $\widehat{\phi}^C = 0.040$ as revealed in Figure 15 in the Appendix. Using this estimate and data on goods-travel times d^C_{ijt}, we then obtain an estimate of ζ_{ijt}.

[43] It is useful to define $\mathfrak{M}_{it}$ as the Mahalanobis distance of a prefecture i from the Chinese average in terms of vector $\hat{d}^C_{it}$. It measures how distant a prefecture i is to all others in terms of goods-travel times. Using available data for 2004, the goods-travel-time Mahalanobis distance for prefecture i is:

$$\mathfrak{M}_{i2004} = \sqrt{\hat{d}_{i2004}\Sigma^{-1}_{2004}\hat{d}'_{i2004}}, \tag{19}$$

where Σ_{2004} is the symmetric, positive-semi-definite, $s-1\times s-1$ variance-covariance matrix of travel times among the Chinese prefectures in 2004. The proposed procedure is best informed for those prefectures i and i', for which the absolute difference $|\mathfrak{M}_{it} - \mathfrak{M}_{i't}|$ takes on the maximum value within a province. Consequently, we use only those prefectures in a province for which $|\mathfrak{M}_{it} - \mathfrak{M}_{i't}|$ is maximal.

5.8 Measurement and Diffusion of Endogenous Technology (T_{it}, γ)

Solving the model for a reference year t requires knowledge on the parameter governing the technology-adjustment process, γ. In order to estimate γ using the technology-diffusion equation (5), we need to solve for region-specific endogenous technology levels, T_{it}. We do so by employing a standard contraction-mapping procedure based on the model structure. The system of equations is derived from the goods-market-clearing condition and uses parameter values for $\{\mu, \theta, \zeta_{ijt}\}$ for all regions $\{i, j, k\}$ from the earlier subsections, along with observed population data, wages, and land endowments for four consecutive years $t \in \{2001, 2002, 2003, 2004\}$.[44,45]

Based on the respective model solutions for T_{it}, we can estimate γ. Taking logs of equation (5) obtains

$$\log T_{it} = \gamma \left(\sum_{j \in S} \mathbb{W}^C_{ijt-1} \log \frac{L_{jt-1}}{H^C_j} \right) + \log \varepsilon^C_{it}, \tag{21}$$

where $\log \varepsilon^C_{it}$ is the log of the technology shifter that includes a common constant. Table 3 reports

Table 3: Endogenous Technology-parameter Estimation Results

	(1)	(2)	(3)	(4)
	Dependent variable: $\log(T_{it})$			
Instrument family	Opt. Netw. (Long.)	Historical	Opt. Netw. (Pop. 1953)	
IV ingredient for $\mathbb{W}^C_{ijt-1}$	Opt. Netw.	Courier Routes	Opt. Netw.	Opt. Netw.
IV ingredient for L_{jt-1}/H^C_j	Netw. Centr.	Pop. 1953	Netw. Centr.	Area
Network-weighted density (γ)	0.306**	9.269	0.340**	0.348**
	(0.138)	(15.60)	(0.163)	(0.173)
Obs.	1319	747	1319	1319
Weak instr. (F-statistic)	216.2	0.366	148.3	129.9

Notes: Opt. Netw. (Long.) refers to the instrument using the longitude-based optimal transport network. *Historical* refers to the instrument using the historical transport network. *Opt. Netw. (Pop. 1953)* refers to the instrument using the historical-population-based optimal transport network. *Netw. Centr.* stands for the respective network centrality measure. *Pop. 1953* refers to the observed population levels in 1953. *Area* refers to the prefecture's total area in sqkm (measured in 2000). Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For one endogenous regressor and a maximum relative bias of 5%, the critical value for the weak-instrumentation F-statistic is 16.38 (Stock and Yogo, 2005).

the results from estimating equation (21) by 2SLS for the years $t \in \{2001, 2002, 2003, 2004\}$. Each

[44] Using the goods-market-clearing condition in (11), the system of equations solving for endogenous technology levels is

$$T_{it} = \frac{w_{it}^{1+\theta} L_{it}^{1+\theta(1-\mu)} {H_i^C}^{\theta(\mu-1)}}{\sum_{j \in S} \left[w_{jt} L_{jt} \zeta_{ijt}^{-\theta} \left(\sum_{k \in S} T_{kt} L_{kt}^{-\theta(1-\mu)} w_{kt}^{-\theta} {H_k^C}^{-\theta(\mu-1)} \zeta_{kjt}^{-\theta} \right)^{-1} \right]}. \tag{20}$$

[45] Population data come from the Socioeconomic Data and Application Center (SEDAC), annual average wages per capita from the Chinese Annual Survey of Industrial Firms (CASIF), and land endowments from the OSM GeoFabrik database.

column uses a different instrumentation strategy for both components of the single endogenous regressor, $\left(\sum_{j\in S}\mathbb{W}^C_{ijt-1}\log\frac{L_{jt-1}}{H^C_j}\right)$, one for $\mathbb{W}^C_{ijt}$ and one for the population density over commercial land use (L_{jt}/H^C_j) (see the IV strategy outlined in Section 5.5).

For instance, in Column (1), we replace d^C_{ijt-1} for all $i\neq j$ by the Monge-Kantorovich optimal network where longitude is transported across prefectures. Then, we use this instead of the measured d^C_{ijt-1} as the ingredient in the otherwise identical procedure to construct weights $\mathbb{W}^C_{ijt-1}$. Moreover, we use the Monge-Kantorovich-network centrality as introduced in the first part of Subsection 5.5 to replace (L_{jt-1}/H^C_j). Finally, we use those ingredients to compute the counterpart to $\left(\sum_{j\in S}\mathbb{W}^C_{ijt-1}\log\frac{L_{jt-1}}{H^C_j}\right)$ and employ it as an instrument for the latter in estimating equation (21). In Columns (2)-(4) we proceed in a similar way except for using alternative measures as indicated in the table to replace $\mathbb{W}^C_{ijt-1}$ and (L_{jt-1}/H^C_j) to construct the IV.

The results suggest that the effect of past network-weighted population density on endogenous technology is statistically significant and stable at around 0.3 (except for Column (2)). We choose Column (1) as the preferred specification as the weak instrumentation F-statistic is highest there. This obtains a parameter value of $\hat{\gamma}=0.306$.

5.9 Measurement and Diffusion of Endogenous Amenities (a_{it}, λ)

The adopted procedure to obtain endogenous amenity levels, a_{it}, and the amenity-diffusion parameter, λ, is analogous to the one in the previous Section 5.8. First, we numerically solve for a_{it} through a contraction-mapping approach based on the model structure. The system of equations is derived from the migration shares in equation (10) and uses parameter values for $\{\alpha,\mu,\Omega,\theta,\kappa_{ijt},\zeta_{ijt},T_{it}\}$ for all regions $\{i,j,k\}$ from the earlier subsections, along with observed population data, wages, and land endowments for four consecutive years $t\in\{2001,2002,2003,2004\}$.[46,47]

With a measure for a_{it} at hand, we can estimate λ using the endogenous amenity equation

[46]The system of equations can be derived when rewriting the migration shares in (10), replacing the location-specific indirect utility with the expression in equation (3) and the price index with the expression in equation (8). Then, after some rearrangements, a_{it} can be expressed as

$$a_{it}=\frac{L_{it}^{\Omega+(1-\alpha)}\left(\bar{L}\tilde{\kappa}_{it}\right)^{-\Omega}w_{it}^{-\alpha}H_i^{R(\alpha-1)}}{\sum\limits_{j\in S}\left[\left(\sum\limits_{k\in S}\left(\frac{a_{kt}w_{kt}^{\alpha}H_k^{R(1-\alpha)}}{\kappa_{kjt}L_{kt}^{(1-\alpha)}}\right)^{\frac{1}{\Omega}}\left(\sum\limits_{k\in S}\Pi_{kt}\zeta_{kjt}^{-\theta}\right)^{\frac{\alpha}{\theta\Omega}}\right)^{-\Omega}\left[\sum\limits_{i\in S}\Pi_{it}\zeta_{ijt}^{-\theta}\right]^{\frac{\alpha}{\theta}}\right]},\tag{22}$$

where $\tilde{\kappa}_{it}=\sum\limits_{j\in S}\kappa_{ijt}^{-1/\Omega}$ and $\Pi_{kt}\equiv T_{kt}w_{kt}^{-\theta}(L_{kt}H_k^C)^{-\theta(1-\mu)}$.

[47]Observed levels of population, wages, and land endowments are taken from the same sources as described in Footnote 45.

(2). Taking logs of equation (2) obtains

$$\log a_{it} = -\lambda \left(\sum_{j \in S} \mathbb{W}^R_{ijt-1} \log \frac{L_{jt-1}}{H^R_j} \right) + \log \varepsilon^R_{it}, \tag{23}$$

where $\log \varepsilon^R_{it}$ is the log of the amenity shifter that includes a common constant.

Table 4: Endogenous Amenity-parameter Estimation Results

	(1)	(2)	(3)	(4)
	Dependent variable: $\log(a_{it})$			
Instrument family	Opt. Netw. (Long.)	Historical	Opt. Netw. (Pop. 1953)	
IV ingredient for $\mathbb{W}^R_{ijt-1}$	Opt. Netw.	Courier Routes	Opt. Netw.	Opt. Netw.
IV ingredient for L_{jt-1}/H^R_j	Netw. Centr.	Pop. 1953	Netw. Centr.	Area
Network-weighted density $(-\lambda)$	-0.0270	-0.171***	-0.133***	-0.147***
	(0.0340)	(0.0587)	(0.0403)	(0.0423)
Obs.	1319	747	1319	1319
Weak instr. (F-statistic)	260.6	48.60	184.9	168.4

Notes: Opt. Netw. (Long.) refers to the instrument using the longitude-based optimal transport network. *Historical* refers to the instrument using the historical transport network. *Opt. Netw. (Pop. 1953)* refers to the instrument using the historical-population-based optimal transport network. *Netw. Centr.* stands for the respective network centrality measure. *Pop. 1953* refers to the observed population levels in 1953. *Area* refers to the prefecture's total area in sqkm (measured in 2000). Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For one endogenous regressor and a maximum relative bias of 5%, the critical value for the weak-instrumentation F-statistic is 16.38 (Stock and Yogo, 2005).

As for endogenous technology, we estimate (23) by 2SLS for the years $t \in \{2001, 2002, 2003, 2004\}$ and report the estimation results in Table 4. Again, the columns differ by the instrumentation strategy for the endogenous regressors and this strategy is analogous to the one used in the previous subsection (see also the IV strategy outlined in Section 5.5). The results suggest that the effect of the past network-weighted population density on the endogenous amenity is statistically significant and stable at around 0.15 (except for Column (1)). We choose Column (3) as the preferred specification as the weak-instrumentation F-statistic is highest there. This obtains a parameter value of $\hat{\lambda} = 0.133$.

6 Benchmark Equilibrium and Counterfactual Analysis

In this section, we investigate the economic effects of the (road and railway) transport-infrastructure improvements that took place in China between 2000 and 2013. The remainder of this section is organized as follows. We first present the main aspects of the initial state (Section 6.1), i.e., the level of trade and migration costs which are consistent with the transport infrastructure of the year 2000. We report on some outcome variables – population, nominal wages relative to Beijing's, amenities, and technology across Chinese prefectures – in the year 2000. In Section 6.2, we describe the distribution of key long-run-equilibrium values across Chinese prefectures associated with the baseline calibration. These values include population, nominal wages relative to Beijing's, amenities, and technology, and they are obtained when letting the model run based on the initial conditions and transport-infrastructure levels of the year 2000. In Section 6.3, we discuss the overall long-run effects on population and real income in the model when changing the transport infrastructure in 2000 to the level of 2013. We then decompose the changes of the long-run equilibrium effects by channel and by road and railway layer. We first focus on the decomposition along four channels in Section 6.4 into: (1) increased accessibility of amenities, (2) increased technology diffusion, (3) reduced migration frictions, and (4) reduced trade frictions. Finally, we decompose the economic effects of transport-infrastructure improvements by infrastructure type: (1) highway-network improvements, (2) regional road-network improvements, and (3) railway-network improvements in Section 6.5.

6.1 Changes in Trade and Migration Costs and Initial Model State

In Figure 8, we provide prefecture-level maps to highlight the implicit trade and migration costs associated with the road and railway infrastructure in the initial period in 2000 as well as their evolution between 2000 and 2013.[48] In the top two panels of Figure 8, we illustrate the summed levels of trade costs (left panel) and migration costs (right panel) across prefectures. A larger value implies larger trade and migration costs.[49] These summarize the status quo of relative aggregate remoteness across the prefectures as of the year 2000. The two panels suggest that, not surprisingly, the domestic transport network was particularly dense – and associated trade and mobility costs were low – in what may be called central-eastern China (i.e., the blue areas). This is particularly true for the greater region around Beijing and the provinces surrounding it. Moreover, frictions were particularly low in the south in the bigger area of the Pearl River Basin around Guangdong province.

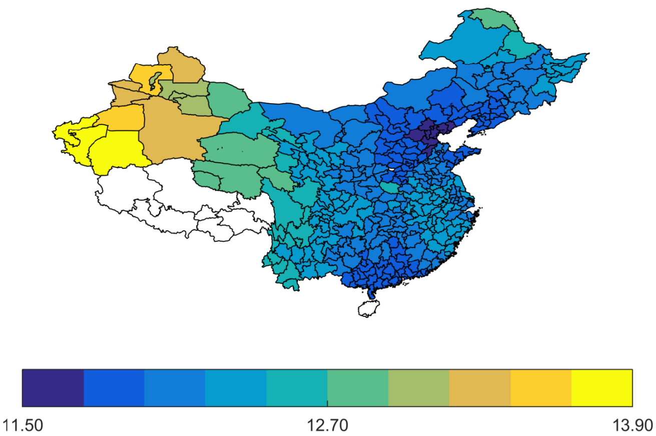
In the bottom two panels of Figure 8, we report on changes in these remoteness variables due to the transport-infrastructure changes between the years 2000 and 2013. The figure suggests that large reductions in trade and migration costs happened especially in peripheral areas which are remote from the coast. The western and northern prefectures (along the Russian and Mongolian border) experienced a particularly important reduction in trade and migration costs. Moreover, a

[48]The prefecture-level maps use the delineation of prefectures as of the year 2000.

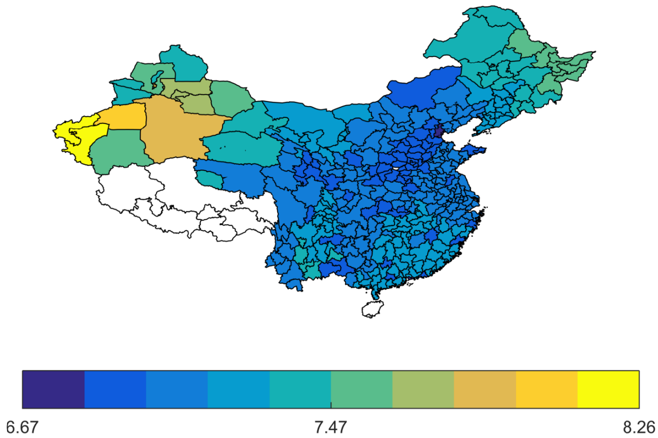
[49]In the figure we assign to each prefecture the average value within the decile it belongs to. This is done to avoid the domination of extreme values in illustration.

Figure 8: PREFECTURE-LEVEL TRADE AND MIGRATION COSTS

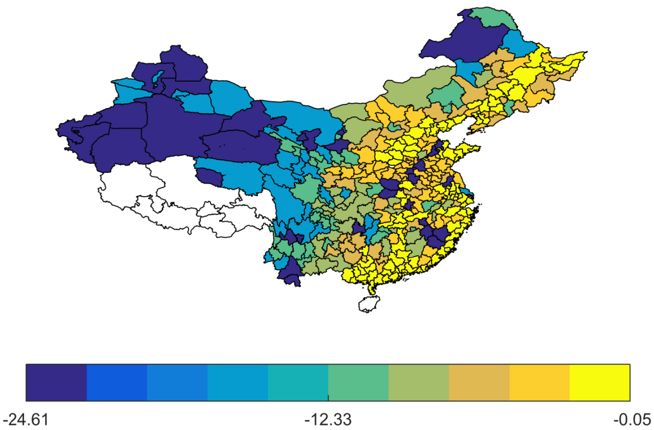
(a) Trade Costs ($\sum_{j \in S} \zeta_{ijt}$),
Log-levels in 2000



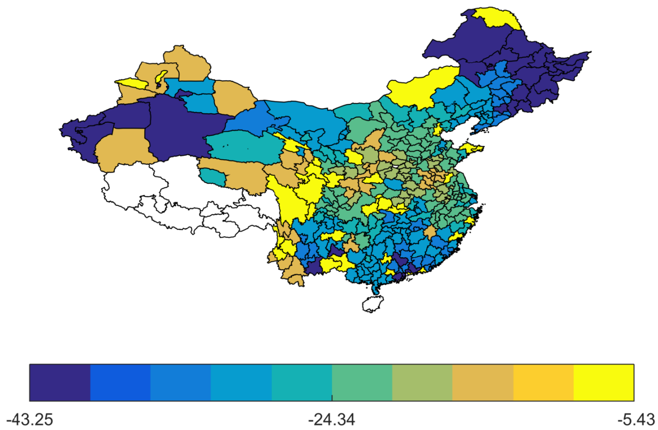
(b) Migration Costs ($\sum_{j \in S} \kappa_{ijt}$),
Log-levels in 2000



(c) Trade Costs,
Changes in 2000-2013



(d) Migration Costs,
Changes in 2000-2013



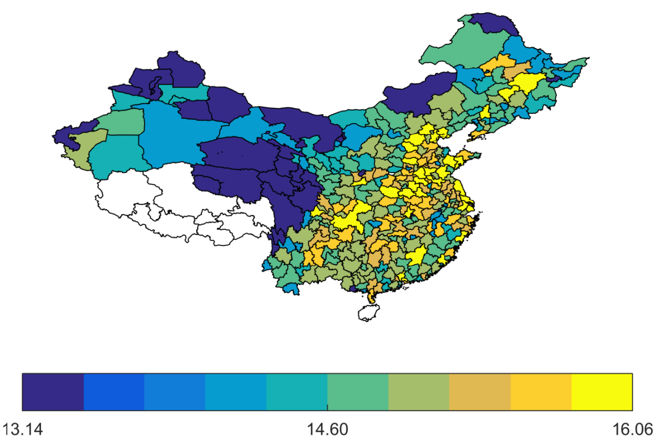
Notes: White prefectures are omitted for network accessibility reasons, as indicated in Section 5. Location-specific migration and trade costs are defined as the sum of associated frictions over all destinations.

few prefectures in central-eastern China without direct access to the coast experienced an increase in connectedness. This last observation is especially true for the transport of goods (Figure 8c).

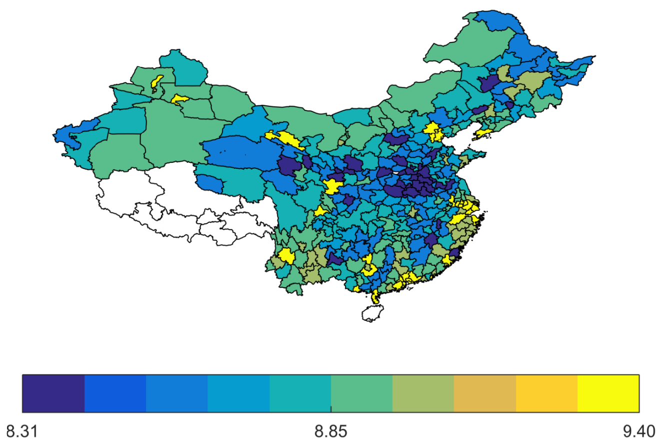
In Figure 9, we report on the *observed* distribution of population, normalized wages, location-specific amenities and location-specific technology in 2000. Normalized wages are expressed relative to the ones in Beijing. Clearly, the two maps illustrate that the population was dense in central-eastern China, especially, in the greater region around Beijing and bordering provinces. Wages were high along the coast and in few central prefectures. Initial location-specific amenities were particularly high in remote and less connected prefectures, whereas initial location-specific technology was high along the coast and in large urban centers such as Beijing or Shanghai.

Figure 9: Key Variables in the Initial State

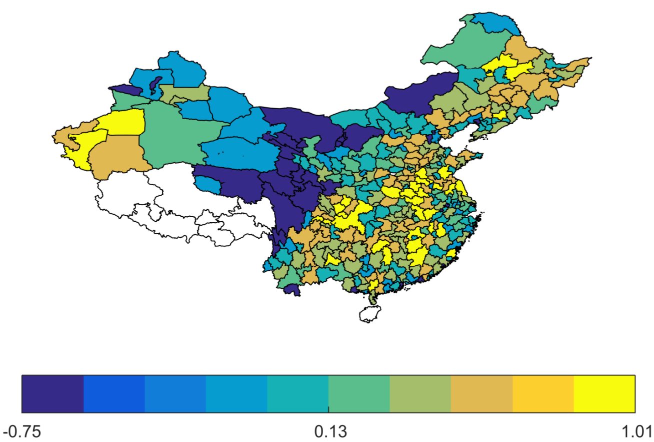
(a) Initial Population, Log-levels in 2000



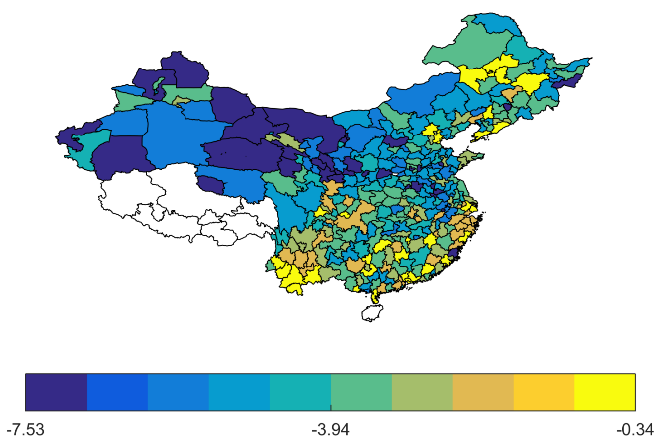
(b) Initial Normalized Wages, Log-levels in 2000



(c) Initial Location-specific Amenity Log-levels in 2000



(d) Initial Location-specific Technology Log-levels in 2000



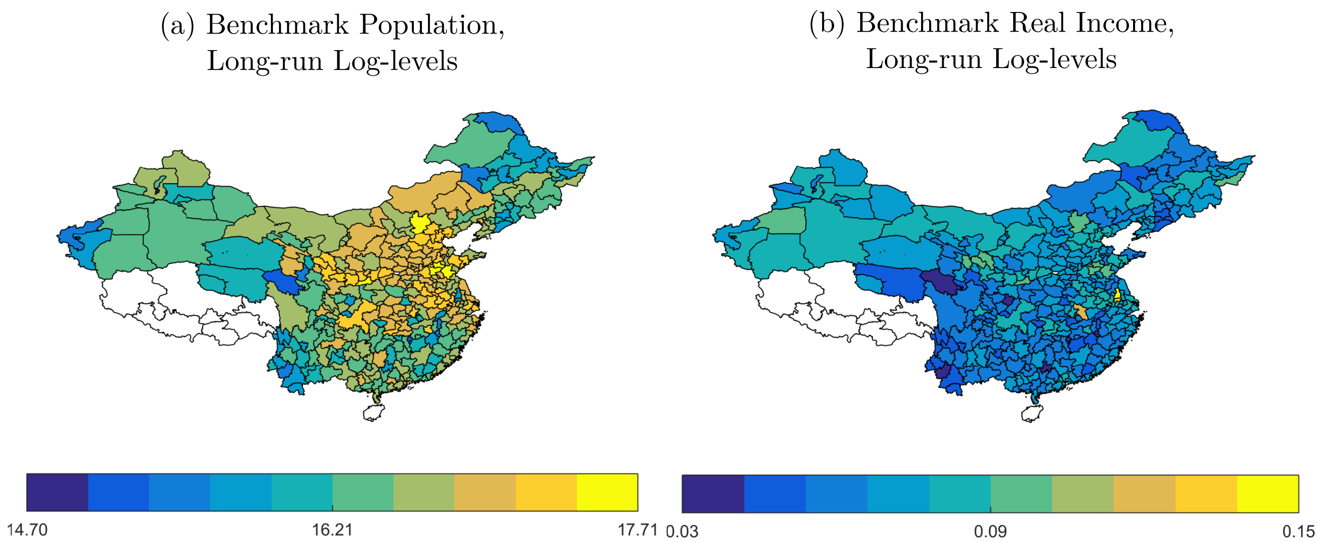
Note: White prefectures are omitted for network accessibility reasons, as indicated in Section 5.

6.2 Simulating Baseline-equilibrium Outcomes

For the counterfactual analysis, we start with simulating the model forward to obtain long-run vectors of endogenous variables of particular interest based on the initial conditions established for the year 2000 and keeping the road and railway infrastructure constant at that year's level. Population and real income summarize the key effects at play in our framework, and all other effects can be derived from these.

The simulation runs forward for P time periods until it reaches a long-run general equilibrium. How many time periods it needs to reach this equilibrium depends on the differences $(L_{it-1} - L_{it})$ and $(w_{it-1} - w_{it})$ for all $\{i, t\}$ in the forward-iteration path. We assume that as soon as the maximum absolute difference for either equilibrium outcome across regions i is lower than $1e-03$, the system has reached an equilibrium state $\{L_i^*, w_i^*\}$. Knowing the equilibrium distribution of employment and wages, the one of technology and amenity levels, T_i^* and a_i^*, is also determined. One can then solve for the location-specific equilibrium utility $\tilde{u}_i^*$ through a contraction mapping using (10). This determines the equilibrium real income $y_i^* = \tilde{u}_i^*/a_i^* = \frac{w_i^*}{P_i^{\alpha *} r_{it}^{R\,1-\alpha *}}$. With the mentioned stopping criteria, we reach the baseline long-run equilibrium after $P = 8$ periods. We illustrate the long-run baseline-equilibrium values of the mentioned outcomes by way of Figure 10. In contrast to the observed distribution in 2000, Figure 10 suggests that there is a decentralization pattern for population, especially in favor of peripheral prefectures. Real income appears to converge nationally along the adjustment path as peripheral prefectures reduce the gap to the large urban centers such as Beijing and Shanghai.

Figure 10: KEY VARIABLES IN THE LONG-RUN EQUILIBRIUM CONSISTENT WITH YEAR-2000 TRANSPORT INFRASTRUCTURE



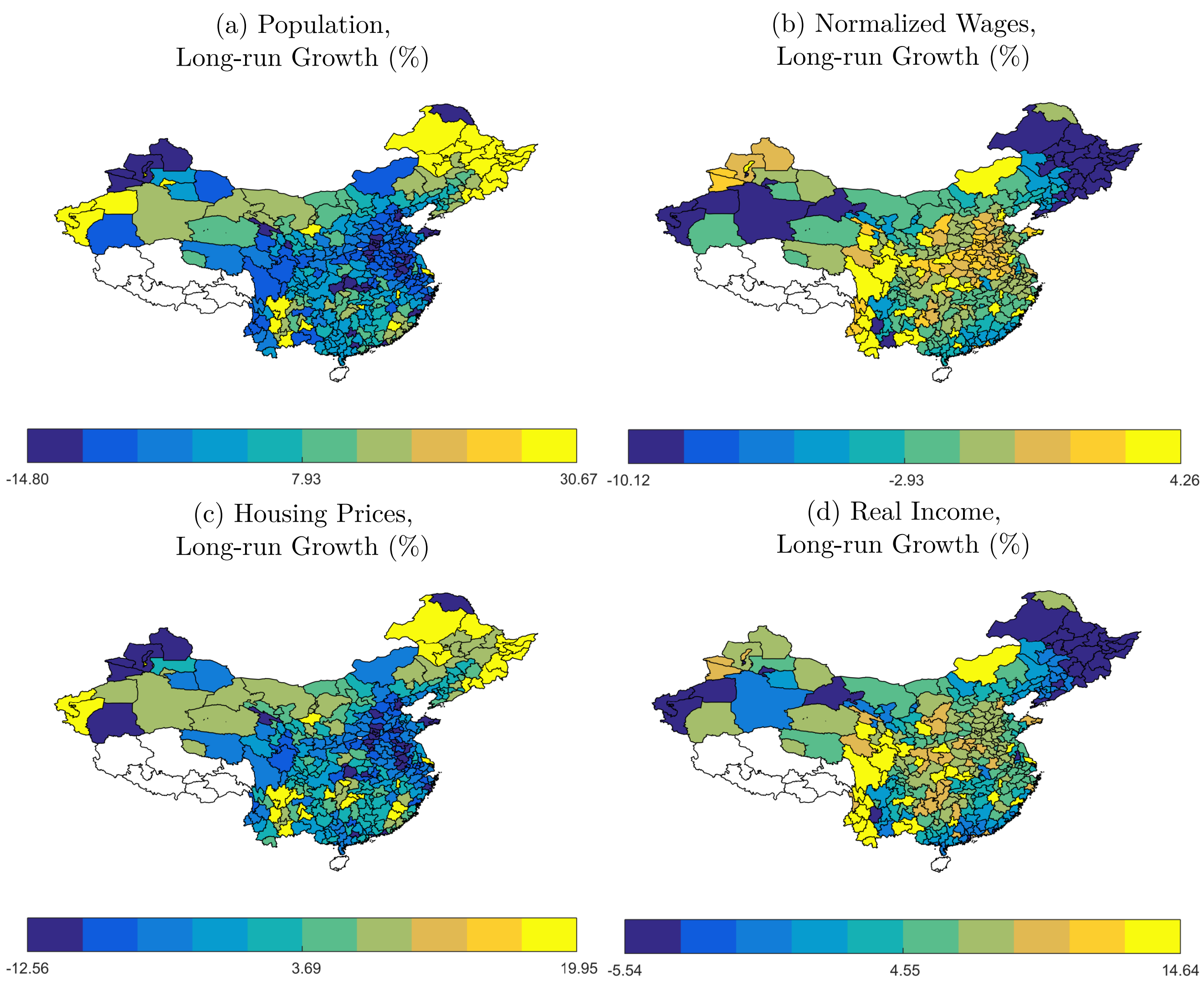
Note: White prefectures are omitted for network accessibility reasons, as indicated in Section 5.

6.3 Long-run Economic Effects of a Counterfactual Change in the Transport Infrastructure Between Years 2000 and 2013

To assess the overall economic effects of China's road- and railway-network improvements between 2000 and 2013, we consider the long-run equilibrium with the transport network using year-2013 versus year-2000 data. We illustrate the overall changes in the outcome variables of particular interest – population, normalized wages, housing prices, and real incomes – in Figure 11 by way of maps for all covered prefectures of mainland China.

All changes of variables between the counterfactual and baseline long-run equilibrium are expressed as growth rates in percent. Hence, larger values imply larger *relative* changes. Figure

Figure 11: Changes in Long-run Equilibrium Variables (2000-2013)



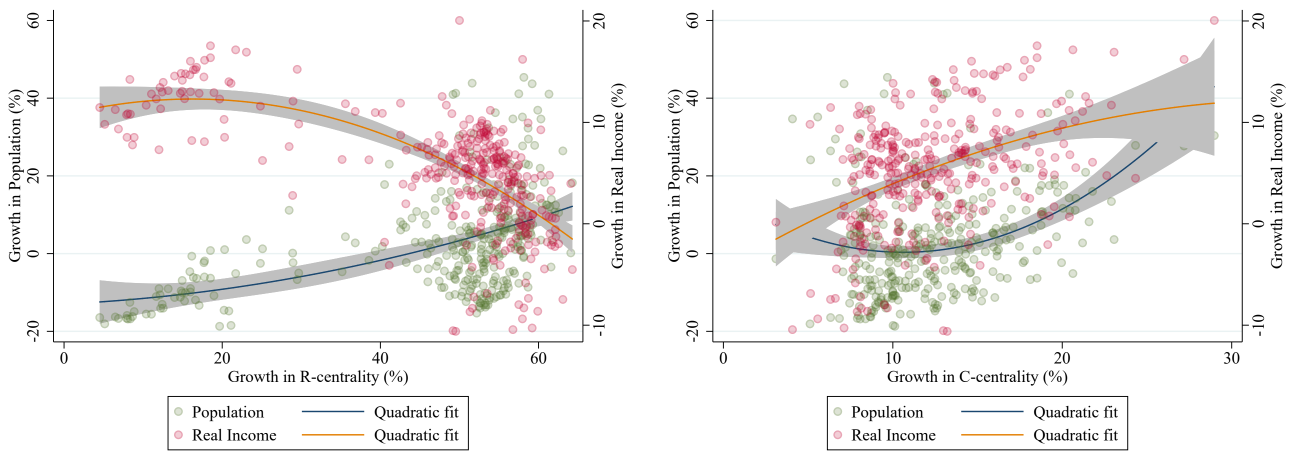
Note: White prefectures are omitted for network accessibility reasons, as outlined in Section 5.

11 suggests that peripheral prefectures (along the Russian border and in the far West) particularly benefited in terms of population. Their population growth comes at the expense of coastal or near-coastal prefectures. The largest reductions in total migration costs in Figure 8 appear to explain the largest population growth. Local population increases translate into smaller normalized wage growth (due to the increase in labor supply) and larger housing-price growth (due to congestion in the local housing market). In turn, real income grows relatively more in places with a smaller population growth due to a reduction in local congestion. Real income growth is particularly large for central prefectures on the longitude line south of Mongolia (i.e., 105° East). These prefectures benefited from important gains in connectedness, while enjoying a strategic geographical location (being a mid-way between China's large centers and more peripheral regions in the far West).

Figure 12 shows how population and real income grew as a function of transport-centrality improvements. We label centrality in the network for people's mobility as "R-centrality" and centrality in the network for goods transport as "C-centrality". Centrality is measured in a straightforward manner by taking, for every prefecture, the sum of inverse travel times to all prefectures (using unity for the intra-prefectural travel time) as defined in Subsection 5.4. Formally, R- and C-centrality growth is equal to the respective percentage changes in centrality.

Figure 12: Long-run Equilibrium Growth in Population and Real Income

(a) R-centrality (b) C-centrality



Notes: 95% confidence intervals are displayed.
Growth in R-centrality and C-centrality is measured as: $\frac{\sum_j (1/d^v_{ij2013}) - \sum_j (1/d^v_{ij2000})}{\sum_j (1/d^v_{ij2000})}$, for $v = \{R, C\}$.

Studying growth in R-centrality, Figure 12a reveals the same patterns as Figure 11. A larger centrality (or connectivity) growth leads to a larger population growth (from -10% up to 9%), and a relatively smaller increase in real income (from 10% to about -2%). Note, however, that the increase in population only concerns the top-half of the prefectures in terms of connectivity growth. Individuals relocate from the bottom-half and the RoW (i.e., initially large centers) towards prefectures with improved centrality (i.e., initially small and mid-sized prefectures). Growth in C-centrality, however, leads to increases in both population and real income as revealed in Figure 12b. It affects primarily the local production, and, hence, leads to smaller adverse congestion effects on the labor or housing markets.

Do the overall effects of the transport network improvements lead prefectures to converge in the long run in terms of the considered outcome variables? To compare convergence between the baseline and the counterfactual appropriately, given normalizations, we look at the difference in the Herfindhal-Hirschmann Index (HHI) between the counterfactual and the baseline long-run equilibria.[50] A negative value implies a smaller concentration in the counterfactual than in the

[50] Formally, denoting the counterfactual (baseline) variables with a subscript ^{c} (^{b}) and for a generic variable, V,

baseline scenario and, hence, convergence. Overall, Table 5 shows that the transport-network improvements between 2000 and 2013 appear to have led prefectures to converge in population levels and real income. Convergence in population is particularly important. These findings are in line with the patterns observed in Figure 12.

Table 5: Long-run Regional Convergence

	(1)	(2)
	Population	Real Income
$HHI \times 10^4$	-1.1798	-0.1148

Notes: A negative value indicates convergence. The HHI is scaled for better readability of the coefficient estimates.

6.4 Decomposing the Long-run Equilibrium Effects by Channel

The present framework allows for impact of transport infrastructure along four channels: (i) increased accessibility of amenities, (ii) increased technology diffusion, (iii) reduced migration frictions, and (iv) reduced trade frictions. Each channel affects the long-run equilibrium in its own way. For instance, if prefectures were symmetric in terms of all exogenous fundamentals, an equalization of mobility would lead individuals to locate such that prefectures were equally large. This would also maximize trade flows. Conversely, making migration costs extremely asymmetric would lead individuals to relatively concentrate in the prefecture with the lowest cost. This would reduce the demand and need for trade, as fewer people elsewhere would need to be served with goods from the agglomeration. Similar arguments can be made for amenities and technology. For instance, making amenities and technology elsewhere better accessible reduces the importance of workers to locate, where local amenities and expected local productivity draws are the best. Hence, all of these channels affect the attractiveness of locations on their own, and they matter for long-run-equilibrium outcome across prefectures in an interdependent way.

To disentangle the relative importance of each channel through China's transport-network expansion between 2000 and 2013, we run four different counterfactual exercises. In each of them, we replace the 2000 transport network by the 2013 network as before, but now *only through the channel under investigation*. Hence, while the change in transport infrastructure alters the accessibility of amenities, the diffusion of technology, migration costs of individuals, and transport

the HHI index is defined as:

$$HHI^V = \sum_{i \in S} \Big(\frac{V_i^c}{\sum_i V_i^c}\Big)^2 - \sum_{i \in S} \Big(\frac{V_i^b}{\sum_i V_i^b}\Big)^2. \qquad (24)$$

costs of goods simultaneously, we will consider only one of those changes at a time in this subsection. This will give us an insight into the qualitative as well as the quantitative differences of the effects through the four channels. Clearly, since the model structure is nonlinear, the sum of the individual effects across the four channels will principally add up to more (with a positive correlation of the channel-specific effects) or less (with a negative correlation) than the total. However, the generic insights are not fundamentally affected by this feature.

Channel 1: Diffusion of Amenities ($\mathbb{W}^R_{ij,2000} \rightarrow \mathbb{W}^R_{ij,2013}$) Allowing for amenity diffusion along the improved transport network leads to population and real income effects at odds with the overall transport effects, as shown in Figure 13a. As centrality grows, the importance of locating in a given location as opposed to a nearby one decreases. Hence, increased amenity diffusion leads to a negative association between connectivity and population growth. Conversely, as population locates relatively more nearby, local congestion decreases and real income increases. Either of these effects increases with a higher R-centrality.

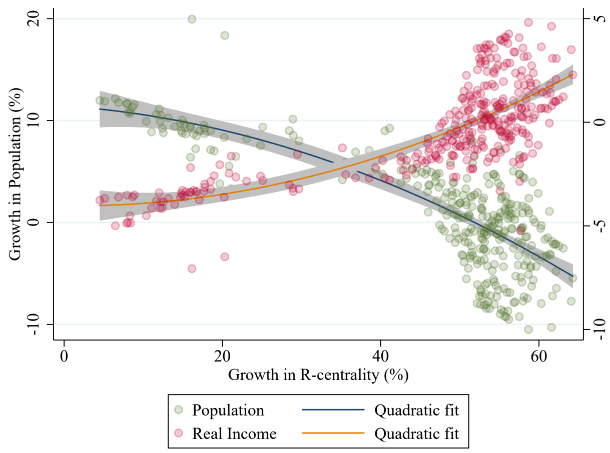
Channel 2: Diffusion of Technology ($\mathbb{W}^C_{ij,2000} \rightarrow \mathbb{W}^C_{ij,2013}$) Figure 13b shows a very different pattern when allowing for technology to diffuse along the improved transport network. Under this channel, both population and real income grow with centrality. Keeping in mind that prefectures experiencing the largest growth in C-centrality are initially more peripheral, it is not surprising that increased technology diffusion greatly benefits these prefectures as they move closer to large technological centers. As the gains in local technology affect primarily productivity, the local gains in population are relatively small (up to 4%) and the adverse effects on congestion remain minimal relative to the gains in wages. Therefore, real income increases.

Channel 3: Reduced Migration Frictions ($\kappa_{ij,2000} \rightarrow \kappa_{ij,2013}$) Allowing a reduction in migration frictions based on the improved 2013 transport network leads to effects in Figure 13c which are symmetrically opposite to the ones for amenity diffusion (Channel 1). Gains in R-centrality lead to increases in population and a reduction in real income. Reduced migration frictions make it possible to locate in growing centers of more peripheral places rather than overcrowded large centers. As population increases, real income declines due to standard congestion effects. The places of origin of migrants, i.e., peripheral or large historical centers, gain particularly in real income (up to 13%).

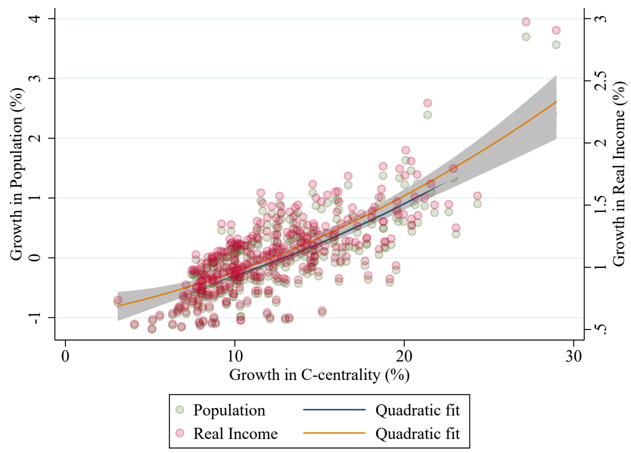
Channel 4: Reduced Goods-trade Frictions ($\zeta_{ij,2000} \rightarrow \zeta_{ij,2013}$) Finally, we isolate the impact of reduced trade frictions that are consistent with centrality changes in China between

Figure 13: Long-run Equilibrium Growth in Population and Real Income by Channel

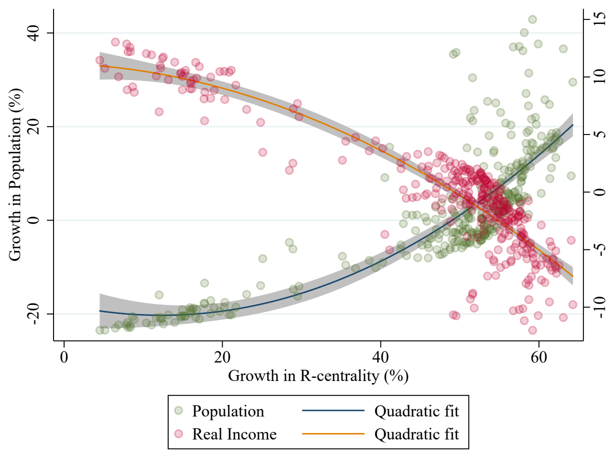
(a) Amenity Diffusion,
R-centrality



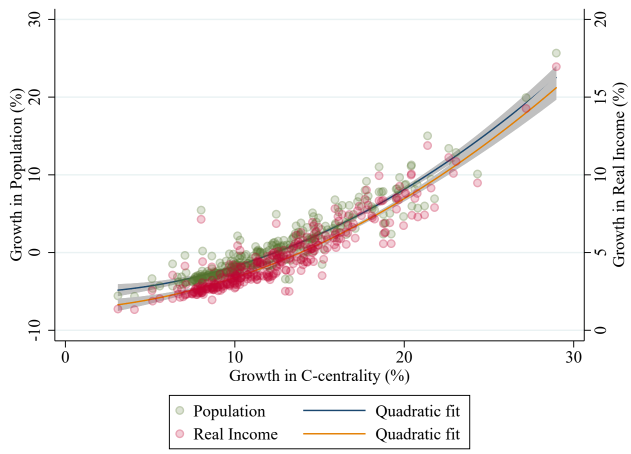
(b) Technology Diffusion,
C-centrality



(a) Reduced Migration Frictions,
R-centrality



(b) Reduced Trade Frictions,
C-centrality



Notes: 95% confidence intervals are displayed.
Growth in R-centrality and C-centrality is measured as: $\frac{\sum_j (1/d^v_{ij2013}) - \sum_j (1/d^v_{ij2000})}{\sum_j (1/d^v_{ij2000})}$, for $v = \{R, C\}$.

the years 2000 and 2013. The picture appears very similar to the effects of technology diffusion (Channel 2), but they are of larger magnitude. Figure 13d reveals that reducing trade frictions leads to large gains both in population (up to 25%) and real income (up to 17%). Local gains (in normalized wages and goods-consumption prices) outweigh the congestion costs (e.g., in housing prices) implied by the increase in population. Therefore, real income increases with C-centrality through reduced goods-transport frictions.

6.5 Decomposing the Long-run Equilibrium Effects by Infrastructure Type

As outlined above, the Chinese road network can be characterized by at least three layers, and the improvements between these layers had been heterogeneous between the years 2000 and 2013. The same is true for the railway system and its two layers (high-speed and standard railways). Highways and (especially high-speed) railways are particularly relevant to connect large centers over long distances. Regional roads connect all centers, large and small, as directly as possible, and they establish an access to the highway and the railway system. Overall, it appears reasonable to expect that each road type and the railway system affect the long-run equilibrium in their own way. In the reminder of this section we scrutinize on the heterogeneity of effects in this regard on the long-run equilibrium variables. To do so, for a generic infrastructure type v, which refers to highways, regional roads and all railways separately, we derive $\mathbb{W}^{R,\mathrm{v}}_{ij,2013}$, $\mathbb{W}^{C,\mathrm{v}}_{ij,2013}$, $\zeta^{\mathrm{v}}_{ij,2013}$ and $\kappa^{\mathrm{v}}_{ij,2013}$ by adding *only* the v-specific network improvements made between 2000 and 2013 to $\mathbb{W}^{R}_{ij,2000}$, $\mathbb{W}^{C}_{ij,2000}$, $\zeta_{ij,2000}$ and $\kappa_{ij,2000}$.

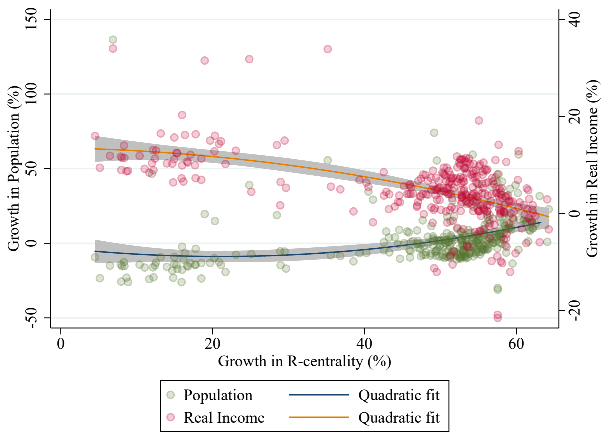
Improvements of Highways Figures 14a-b display the long-run-equilibrium values of population and real income as a function of R- and C-centrality, respectively. Both a scatter plot and the quadratic fit (in terms of centrality and centrality squared) with 95% confidence intervals are reported. Figure 14a reveals that larger highway improvements lead to a larger population growth (up to 15%). As for the overall effects, this growth in population comes at the cost of reduced growth in real income. The picture is different when looking at the effect of improved C-centrality. Growth in C-centrality due to the highway network expansion is associated with larger population and real income levels. Population changes between -15% to 35%, whereas gains in real income range from -5% to 20% in response to the shock.

Improvements of Regional Roads Results for effects due to regional-road-network expansions and improvements on the long-run equilibrium are displayed in Figures 14c-d. Qualitatively the results are similar to those under improved highways, however, the dispersion is larger with gains in population and real income up to 200% and 55%, respectively. This finding is the consequence of two effects: first, regions better connected with regional roads were smaller and characterized by a particularly low level of road infrastructure in 2000; second, the regional roads are important means of access to the highway and also the railway system. Hence, conditional on year-2000 centrality levels, regional roads have induced some of the largest long-run changes by better connecting smaller prefectures to the Chinese transport network.

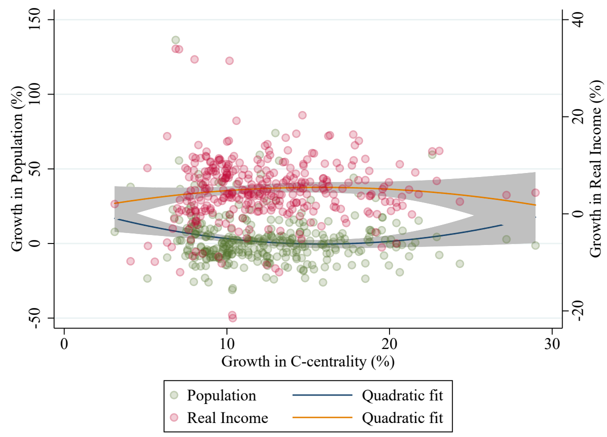
Improvements of Railways In discussing effects of changes of the railway network, we do not discern the HSR lines from other railway lines. Results are displayed in Figures 14e-f. The largest effect of the improved rail network concerned the mobility of people. Growth in R-centrality via the rail network has had a major impact on the Chinese economic geography. This owed to the increase in the size of the network (the length of tracks) as well as the speed of travel both of which facilitated longer-distance journeys, in particular, for people. The population relocates towards prefectures with the largest centrality (up to 20% in the long run) in response to this change. As a consequence, congestion in these places increases and real income declines (by about 8% in the long run). The improvement of the railway network has a smaller and less clear-cut impact on the transport of goods (see Figure 14f). For goods, transport on roads is often the faster alternative on the one hand, as they cannot travel on the HSR system. But the railway network substitutes for road transport of both people and goods, in particular, at long distances (see Figures 1-3). These opposing effects are responsible for the less-clear-cut pattern of effects on long-run outcomes.

Figure 14: Long-run Equilibrium Growth in Population and Real Income by Network Type

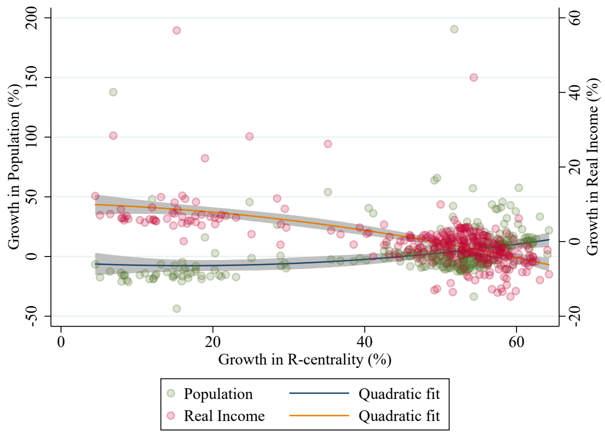
(a) Improvements of Highways, R-centrality



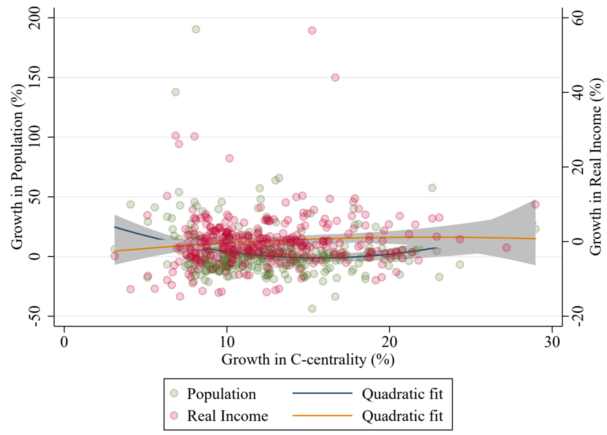
(b) Improvements of Highways, C-centrality



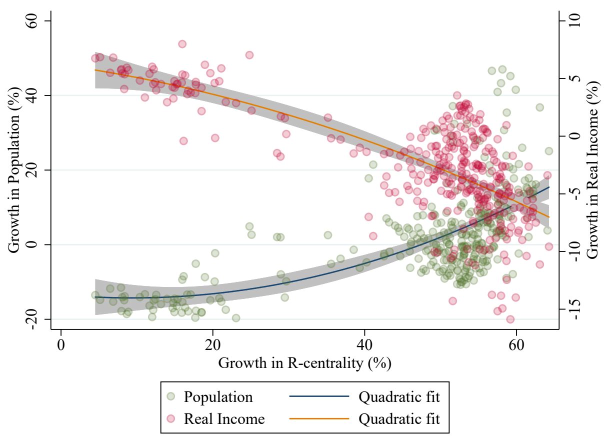
(c) Improvements of Regional Roads, R-centrality



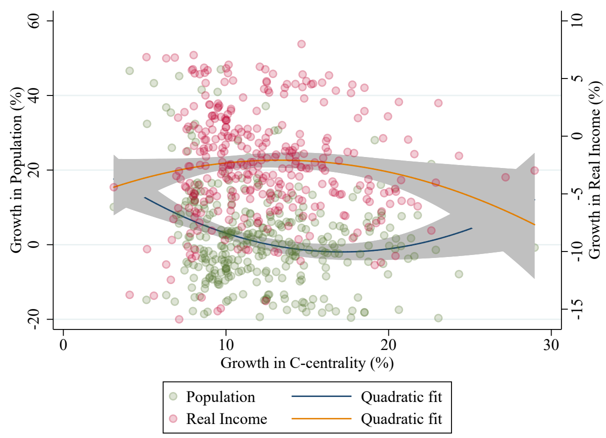
(d) Improvements of Regional Roads, C-centrality



(e) Improvements of Railways, R-centrality



(f) Improvements of Railways, C-centrality



Notes: 95% confidence intervals are displayed.

Growth in R-centrality and C-centrality is measured as: $\frac{\sum_j (1/d^v_{ij2013}) - \sum_j (1/d^v_{ij2000})}{\sum_j (1/d^v_{ij2000})}$, for $v = \{R, C\}$.

7 Conclusions

This paper contributes to the literature on the quantitative effects of transport infrastructure in open regions in four ways. First, it compiles and digitizes detailed annual data on China's transport network over the period 2000-2013. Specifically, it does so using a taxonomy which distinguishes between highways, province-level roads, prefecture-level roads, high-speed rail and standard railways. This dataset permits a detailed view on the changes in transport connectivity, e.g., in terms of shortest travel time, at a detailed regional level and it supports an attribution of the changes to specific layers of the transport network.

Second, the paper presents a quantitative regional model of 330 Chinese prefectures plus a rest of the world, where the transport network may affect regions through four direct channels. Of those, two are customary and at the heart of traditional quantitative models of regions, and these are the goods-trade-cost and the mobility-cost channels. Two other ones are not customary, and they relate to the technology-spillover and amenity-spillover effects of transport networks. By this token, we consider that technology will dissipate more easily between better-connected places, and we allow peoples' residence choices not only to depend on locally available amenities but also on ones in the well-connected neighborhood.

Third, we propose two novel and independent instrumental variable approaches to tackle the endogeneity concerns arising from non-random placement of transport connections. The first approach relies solely on topological information to solve a Monge-Kantorovich optimal transport problem. The second IV approach, instead, combines historical information from the courier routes used during the Chinese Ming Dynasty and 1953 prefecture population records. Both approaches yield similar results when estimating parameters linking travel times and moving costs (for both goods and people), and when estimating transport-related diffusion processes of productivity and amenities.

Finally, we delineate the quantitative effects of China's transport network improvements over the considered time span. Estimates of key model parameters and simulation results suggest that all four channels and all types of transport infrastructure are important in their own right. Network improvements particularly stimulate technology spillovers and the reduction in trade costs both of which foster the convergence of Chinese prefectures in terms of population density and real per capita income. The same improvements stimulate amenity spillovers – a dispersion force which increases real per capita income in better-connected places – and they reduce mobility costs – an agglomeration force which leads to increased population densities at the expense of real income in the same places. Overall, the population gains in peripheral places, associated with smaller gains in real income due to greater housing-market tightness, allow congestion to decline and real per capita income to increase in historically large and well-connected places.

References

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015): "The Economics of Density: Evidence From the Berlin Wall," *Econometrica*, 83, 2127–2189.

Alder, S. and I. Kondo (2018): "Political Distortions and Infrastructure Networks in China: A Quantitative Spatial Equilibrium Analysis," 2018 Meeting Papers 1269, Society for Economic Dynamics.

Allen, T. and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129, 1085–1140.

Allen, T. and D. Donaldson (2018): "The Geography of Path Dependence," Unpublished Manuscript.

Bai, C.-E. and Z. Qian (2010): "The factor income distribution in China: 1978-2007," *China Economic Review*, 21, 650 – 670.

Banerjee, A., E. Duflo, and N. Qian (2012): "On the Road: Access to Transportation Infrastructure and Economic Growth in China," NBER Working Papers 17897, National Bureau of Economic Research, Inc.

Barwick, P., D. Donaldson, S. Li, and Y. Lin (2018): "The Welfare Effects of Passanger Transportation Infrastructure: Evidence from China," Unpublished Manuscript.

Baum-Snow, N. (2007): "Did Highways Cause Suburbanization?" *The Quarterly Journal of Economics*, 122, 775–805.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang (2017): "Roads, Railroads, and Decentralization of Chinese Cities," *The Review of Economics and Statistics*, 99, 435–448.

Baum-Snow, N., J. V. Henderson, M. A. Turner, Q. Zhang, and L. Brandt (2018): "Does investment in national highways help or hurt hinterland city growth?" *Journal of Urban Economics*.

Baum-Snow, N. and M. Turner (2017): "Transport Infrastructure and the Decentralization of Cities in the People's Republic of China," *Asian Development Review*, 34, 25–50.

Berman, L. and W. Zhang (2017): "V6 Ming Dynasty Courier Routes and Stations," .

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93, 1268–1290.

Broda, C. and D. Weinstein (2006): "Globalization and the Gains From Variety," *The Quarterly Journal of Economics*, 121, 541–585.

Chandra, A. and E. Thompson (2000): "Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system," *Regional Science and Urban Economics*, 30, 457–490.

Coe, D. and E. Helpman (1995): "International R&D spillovers," *European Economic Review*, 39, 859–887.

Combes, P.-P. and M. Lafourcade (2005): "Transport costs: measures, determinants, and regional policy implications for France," *Journal of Economic Geography*, 5, 319–349.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018): "The Geography of Development," *Journal of Political Economy*, 126, 903–983.

Desmet, K. and E. Rossi-Hansberg (2009): "Spatial growth and industry age," *Journal of Economic Theory*, 144, 2477–2502.

——— (2014): "Spatial Development," *American Economic Review*, 104, 1211–1243.

Diamond, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106, 479–524.

Dijkstra, E. W. (1959): "A note on two problems in connexion with graphs." *Numerische Mathematik*, 1, 269–271.

Donaldson, D. (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 108, 899–934.

Donaldson, D. and R. Hornbeck (2016): "Railroads and American Economic Growth: A "Market Access" Approach," *The Quarterly Journal of Economics*, 131, 799–858.

Duranton, G., P. Morrow, and M. Turner (2014): "Roads and Trade: Evidence from the US," *Review of Economic Studies*, 81, 681–724.

Duranton, G. and M. Turner (2012): "Urban Growth and Transportation," *Review of Economic Studies*, 79, 1407–1440.

Eaton, J. and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

Egger, P., G. Loumeau, and N. Pueschel (2017): "Natural City Growth in the People's Republic of China," *Asian Development Review*, 34, 51–85.

Faber, B. (2014): "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *The Review of Economic Studies*, 81, 1046–1070.

Fajgelbaum, P. D., E. Morales, J. C. S. Serrato, and O. Zidar (2019): "State Taxes and Spatial Misallocation," *Review of Economic Studies*, 86, 333–376.

Feenstra, R. and H. L. Kee (2008): "Export variety and country productivity: Estimating the monopolistic competition model with endogenous productivity," *Journal of International Economics*, 74, 500–518.

Fracasso, A. and G. Vittucci Marzetti (2015): "International trade and R&D spillovers," *Journal of International Economics*, 96, 138–149.

Garcia-Lopez, M.-A. (2012): "Urban spatial structure, suburbanization and transportation in Barcelona," *Journal of Urban Economics*, 72, 176–190.

Kantorovitch, L. (1958): "On the Translocation of Masses," *Management Science*, 5, 1–4.

Kelejian, H. H. and I. Prucha (2010): "Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances," *Journal of Econometrics*, 157, 53–67.

Lawrence, M., R. Bullock, and Z. Liu (2019): *China's High-Speed Rail Development*, World Bank.

Ma, L. and Y. Tang (2020): "Geography, trade, and internal migration in China," *Journal of Urban Economics*, 115, 103181, cities in China.

Michaels, G. (2008): "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *The Review of Economics and Statistics*, 90, 683–701.

Monge, G. (1781): *Mémoire sur la théorie des déblais et des remblais*, De l'Imprimerie Royale.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018): "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 108, 3855–3890.

Ortega, F. and G. Peri (2016): "The effect of income and immigration policies on international migration," in *The Economics of International Migration*, World Scientific Publishing Co. Pte. Ltd., chap. 11, 333–360.

Qin, Y. (2017): "No county left behind? The distributional impact of high-speed rail upgrades in China," *Journal of Economic Geography*, 17, 489–520.

Rodriguez-Clare, A. (1996): "The role of trade in technology diffusion," Discussion Paper / Institute for Empirical Macroeconomics 114, Federal Reserve Bank of Minneapolis.

Simonovska, I. and M. E. Waugh (2014): "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 92, 34–50.

Stock, J. and M. Yogo (2005): *Testing for Weak Instruments in Linear IV Regression*, New York: Cambridge University Press, 80–108.

Villani, C. (2003): *Topics in Optimal Transportation*, Graduate studies in mathematics, American Mathematical Society.

Wang, Y., L. Li, L. Wang, A. Moore, S. Staley, and Z. Li (2014): "Modeling traveler mode choice behavior of a new high-speed rail corridor in China," *Transportation Planning and Technology*, 37, 466–483.

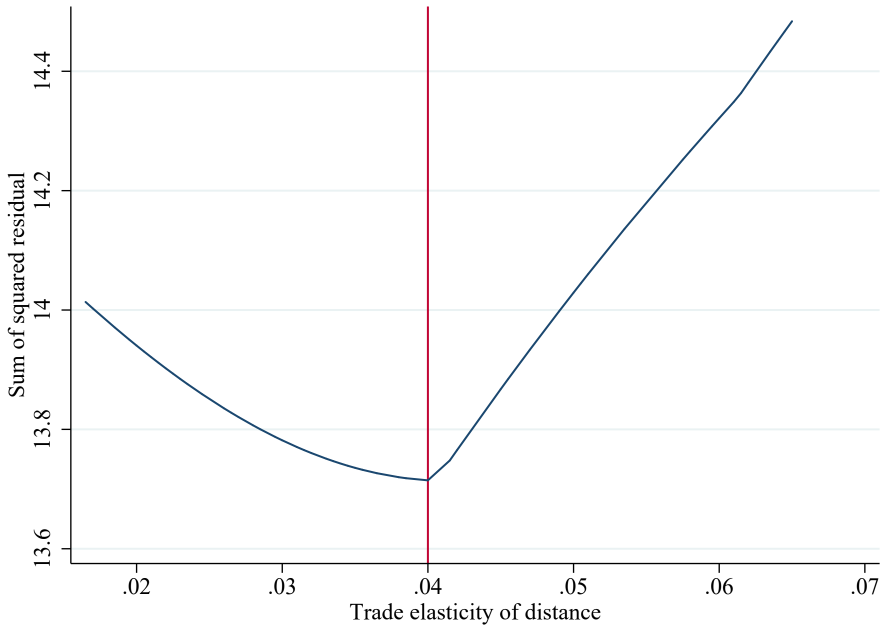
Zabreyko, P., A. Koshelev, M. Krasnosel'skii, S. Mikhlin, L. Rakovshchik, and V. Stetsenko (1975): *Integral Equations: A Reference Text*, Noordoff International Publishing Leyden.

FOR ONLINE PUBLICATION ONLY

Appendix

A Supporting Material

Figure 15: Elasticity of Travel Time to Trade Costs



B Construction of the Geography-based Optimal Transport Network

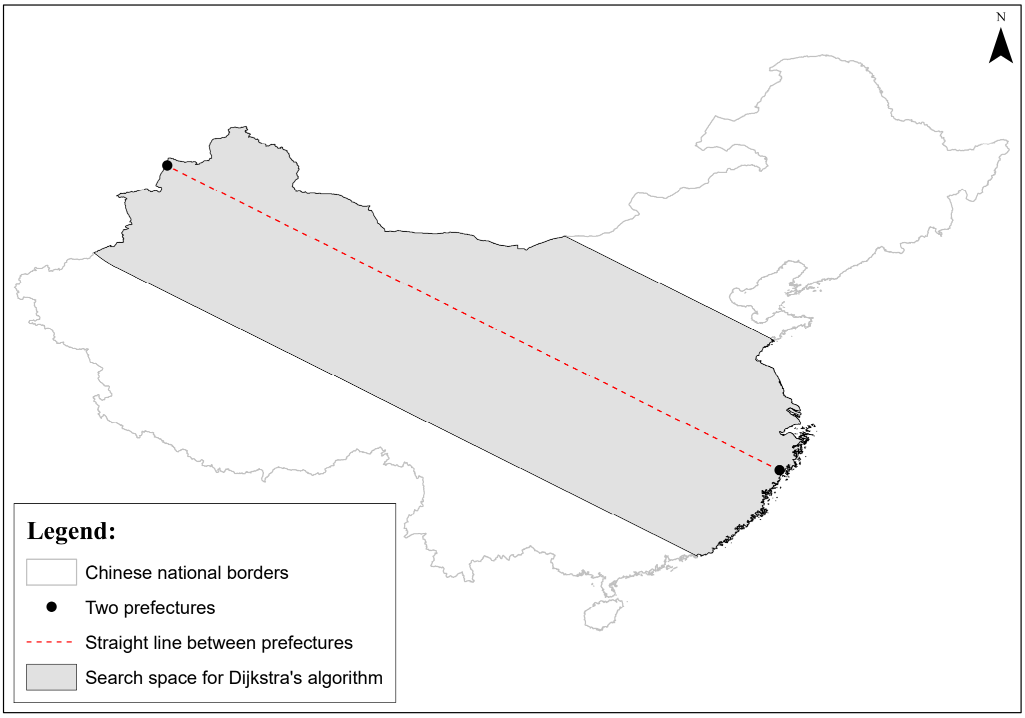
Step 1: Grid and Estimation of Construction Costs

We build a grid of 2,000×2,000 = 4,000,000 cells covering all of mainland China. This leads to cells of about 3.1×3.1 kilometers in size. To each cell, we attribute the average slope (land gradient) in the cell and the area of the cell covered by water. Note that these variables are not affected by the construction of new transport segments in 2007. We also defined an indicator variable which equals unity, if a cell does *not* contain (is not crossed by) a connection in 2007, and zero otherwise. We use the latter as a binary dependent variable in a cross-sectional probit model, where the underlying latent continuous variable should be interpreted as to reflect the costs of building a transport segment in a cell. The index which informs the propensity of construction is specified as a linear function of a constant and the two aforementioned variables (the average slope and the extent of water surface).

Step 2: Computing the Minimum Costs of Building a Connection Between Each Prefecture Pair

Based on the construction costs obtained from Step 1, we determine the minimum-cost path between any two prefectures using Dijkstra's (1959) algorithm. This algorithm searches for the cost-minimizing path by sequentially adding legs (here, cells) to the best current path and comparing it to all alternative paths. Searching the full space for each bilateral connection requires an inefficiently large amount of time as extreme routes are unlikely to be cost-minimizing. Hence, we restrict the search space to all cells within a 7° (about 750 kilometers) bandwidth around the great-circle-distance connection between any two prefectures. Figure 16 illustrates the search space for an exemplary pair of prefectures.

Figure 16: Illustration of the Search Space for Dijkstra's Algorithm



Step 3: Selecting Connections

In Step 3, we select the connections to be included in the optimal network. Importantly, to obtain a valid instrument, it is crucial that the selection process does not depend on shocks and other time-specific components in variables (specifically, population density) which are measured during the period of investigation and, in particular, in 2000, which is the benchmark year for the state of the infrastructure network (see Section 5 for more details). Solutions based on a cost-benefit analysis using contemporary data would typically not be suited. Moreover, given the relatively high density of the observed transport network and the relatively large number of prefectures, a minimum-spanning tree connecting all prefectures is likely to have low predictive power and underestimate the connectivity of prefectures during the period of investigation.

To obtain a valid instrument with a sufficiently high predictive power and a reasonable degree of predicted connectivity among prefectures, we formulate and solve the classical Monge-Kantorovich optimal-transport problem[51] proposed in Monge (1781) and Kantorovitch (1958) in the context of the Chinese transport network. Specifically, we formulate the problem as a standard linear program in which *all* mass defined here by a location's longitude must be transported to another prefecture at minimized overall moving costs. The rationale behind this choice is that China is particularly densely populated in it's East along the coastal belt. Costs between each pair of prefectures are given by the previous steps. Clearly, this leads to a denser network, when large prefectures are surrounded by smaller prefectures than when prefectures are of similar sizes. After defining, c_{ij} as the cost between prefectures i and j and h_{ij} as the mass flow between them, and $\mathbf{h}$ is the stacked vector of $s(s-1)$ elements h_{ij} with $i \neq j$, the linear program can be formulated as follows:

$$\begin{aligned} \min_{\mathbf{h}} \; & f(\mathbf{h}) = \sum_{i=1}^{s} \sum_{j \neq i} c_{ij} h_{ij} \\ \text{s.t. } & \sum_{i} h_{ij} = \sum_{j} h_{ij} \\ & h_{ij} \geq 0, \forall \{i \neq j\}. \end{aligned} \tag{25}$$

To insure that the solution of the optimal-transport problem at hand does not lead to disconnected sub-networks (clusters) within the big network of Chinese prefectures, we require each prefecture to be connected a minimum of three times. The overall optimal-transport network is illustrated in Figure 5.

The *Historical-population-based Optimal Transport Network* is constructed following the same approach, except that a prefecture's population in 1953 is used as the local mass to be transported instead of a prefecture's longitude.

Step 4: From Flows to Travel Times

From Step 3, we obtain optimal mass flows between all prefecture pairs. We then sort all $330 \times 330 = 108,900$ connections based on their optimal mass flows. Mimicking the speed limits applied to the Chinese road network, we apply a travel speed of 120km/h on the top one-third $36,300$ connections with the largest optimal mass flows, a speed of 80km/h to the next 36,300 connections,

[51] See Villani (2003), for a textbook treatment of the mathematical problem.

and a speed of 50km/h to the 36,300 connections with the smallest optimal mass flows. Applying this speed to the distance on the minimum-cost path from Step 2, we consequently obtain the final geography-based optimal transport network based on connection-specific travel times. As indicated in Subsection 5.5, we augment this information by the constant travel times between the port-hosting prefectures and the RoW. Finally, we solve for the fastest-path on this geography-based optimal transport network to derive Monge-Kantorovich fastest travel times among the 331 prefectures in the data.

Note that we pursue the same approach with the *Historical-population-based Optimal Transport Network* to map flows into travel times. With the length-of-routes approach during the Ming dynasty, we generally do not map distances into travel times but use distances instead.

C Derivation: Equilibrium Equations

This section derives the set of equations that solves equilibrium wages and employment levels in each location i.

Equilibrium Wages We start by deriving equilibrium wages. To do so, take the good-market-clearing condition (11) and substitute (7). Then,

$$w_{it}L_{it} = \sum_{j\in S} \frac{T_{it}[o_{it}\zeta_{ijt}]^{-\theta}}{\sum_{k\in S} T_{kt}[o_{kt}\zeta_{kjt}]^{-\theta}} w_{jt}L_{jt}. \tag{26}$$

Replacing unit costs (6) obtains

$$w_{it}L_{it} = T_{it}w_{it}^{-\theta}\left(\frac{L_{it}}{H_i^C}\right)^{-\theta(1-\mu)} \sum_{j\in S} \frac{\zeta_{ijt}^{-\theta} w_{jt}L_{jt}}{\sum_{k\in S} T_{kt}\left(\frac{L_{kt}}{H_k^C}\right)^{-\theta(1-\mu)} w_{kt}^{-\theta}\zeta_{kjt}^{-\theta}}. \tag{27}$$

Solving the previous equation for wages gives the first set of equilibrium equations

$$w_{it} = L_{it}^{\frac{(-\theta(1-\mu))-1}{1+\theta}} {H_i^C}^{\frac{-\theta(\mu-1)}{1+\theta}} T_{it}^{\frac{1}{1+\theta}} \left[\sum_{j\in S} w_{jt}L_{jt}\zeta_{ijt}^{-\theta}\left(\sum_{k\in S}\Pi_{kt}\zeta_{kjt}^{-\theta}\right)^{-1}\right]^{\frac{1}{1+\theta}},$$

where $\Pi_{kt} \equiv T_{kt}\left(\frac{L_{it}}{H_i^C}\right)^{-\theta(1-\mu)} w_{kt}^{-\theta}$.

Equilibrium Employment To derive equilibrium employment take migration shares in (10) and substitute location-specific utility ($\tilde{u}_{it}$). Then,

$$\frac{L_{it}}{\bar{L}} = \frac{\left[\frac{a_{it}w_{it}}{P_{it}^{\alpha}r_{it}^{R\,1-\alpha}}\right]^{1/\Omega}\tilde{\kappa}_{it}}{\sum_{j\in S}\sum_{k\in S}\left[\frac{a_{kt}w_{kt}}{P_{kt}^{\alpha}r_{kt}^{R\,1-\alpha}}\right]^{1/\Omega}\kappa_{kjt}^{-1/\Omega}}, \tag{28}$$

where $\tilde{\kappa}_{it} = \sum_{j\in S}\kappa_{ijt}^{-1/\Omega}$.

Replacing the residential land rents with (13) and the price index with (8) gives, after some rearrangements, the second set of equilibrium equations

$$L_{it} = \frac{\left(w_{it}^{\alpha}(\bar{L}\tilde{\kappa}_{it})^{\Omega}H_i^{R(1-\alpha)}a_{it}\right)^{\frac{1}{\tilde{\Omega}}}\left(\sum_{j\in S}\Pi_{it}\zeta_{ijt}^{-\theta}\right)^{\frac{\alpha}{\theta\tilde{\Omega}}}}{\left[\sum_{j\in S}\left(\sum_{k\in S}\left(\frac{a_{kt}w_{kt}^{\alpha}H_k^{R(1-\alpha)}}{\kappa_{kjt}L_{kt}^{(1-\alpha)}}\right)^{\frac{1}{\Omega}}\left(\sum_{k\in S}\Pi_{jt}\zeta_{kjt}^{-\theta}\right)^{\frac{\alpha}{\theta\Omega}}\right)\right]^{-\frac{\Omega}{\tilde{\Omega}}}}, \tag{29}$$

where $\tilde{\Omega} \equiv \Omega + (1-\alpha)$.

Note that in equilibrium, the location- and time-specific amenity and technology shifter, ε_{it}^{R} ε_{it}^{C}, cancel as their expected value is unity for every location i. Then a_{it} and T_{it} are only dependent on past variables, and hence, exogenously given in period t.

D Equilibrium: Existence and Uniqueness

The uniqueness condition in (14) can be derived along the lines of Desmet et al. (2018) (see their Section B.3). We can manipulate the system of equations that defines an equilibrium as follows.

First Set of Equations We derive a first set of equations that uses the location-specific indirect utility, $\tilde{u}_{it}$, in (3) as a basis. Take the price index in (8) and substitute the unit costs (6), commercial land rents (13), then

$$P_{it} = \psi_0\left[\sum_{j\in S}T_{jt}L_{jt}^{-\theta(1-\mu)}w_{jt}^{-\theta}H_j^{C\,-\theta(\mu-1)}\zeta_{jit}^{-\theta}\right]^{-\frac{1}{\theta}}, \tag{30}$$

where $\psi_0 = \bar{p}/\mu$ and $\bar{p} = \Gamma\left(\frac{1-\sigma}{\theta}+1\right)^{\frac{1}{1-\sigma}}$. Substituting (30) and residential land rents (13) into

(3) gives

$$\left[\frac{a_{it}}{\tilde{u}_{it}}\right]^{-\frac{\theta}{\alpha}} \left[\frac{L_{it}}{H_i^R}\right]^{-\frac{\theta(\alpha-1)}{\alpha}} w_{it}^{-\theta} \quad = \psi_1 \sum_{j\in S} T_{jt} \left[\frac{L_{jt}}{H_j^C}\right]^{\theta(1-\mu)} w_{jt}^{-\theta} \zeta_{jit}^{-\theta}, \tag{31}$$

where $\psi_1 = \psi_0^{-\theta}\left((1-\alpha)^{\alpha-1}\right)^{\theta/\alpha}$.

Second Set of Equations We derive a second set of equations that uses the equilibrium goods-market-clearing condition in (11). Insert the trade shares (7) and the price index into the goods-market-clearing condition so that

$$w_{it}L_{it} = \bar{p}^{-\theta} \sum_{j\in S} T_{it}[o_{it}\zeta_{ijt}]^{-\theta} P_{jt}^{\theta} w_{jt} L_{jt}. \tag{32}$$

Substituting unit costs (6) as well as replacing the price index with the indirect utility into (32) yields

$$\begin{aligned} &T_{it}^{-1} w_{it}^{1+\theta} H_i^{C\,\theta(\mu-1)} L_{it}^{1-(\theta(1-\mu))} \\ &\qquad = \psi_1 \sum_{j\in S} \left[\frac{a_{jt}}{\tilde{u}_{jt}}\right]^{\frac{\theta}{\alpha}} w_{jt}^{1+\theta} H_j^{R\,\frac{\theta(1-\alpha)}{\alpha}} L_{jt}^{1+\frac{\theta(\alpha-1)}{\alpha}} \zeta_{ijt}^{-\theta}. \end{aligned} \tag{33}$$

Proof We follow the uniqueness proof of Allen and Arkolakis (2014) and Desmet et al. (2018) and apply Theorem 2.19 in Zabreyko et al. (1975) for a continuous set of location. It can be shown, that the uniqueness Theorem 2.19 can be applied to the discrete case as well, which Desmet et al. (2018) uses in their empirical application.
Let us introduce the following function $\bar{f}_i$, which is the ratio of LHS's of (31) and (33):

$$\bar{f}_i = \frac{T_{it}^{-1} w_{it}^{1+\theta} H_i^{C\,\theta(\mu-1)} L_{it}^{1-(\theta(1-\mu))}}{\left[\frac{a_{it}}{\tilde{u}_{it}}\right]^{-\frac{\theta}{\alpha}} L_{it}^{-\frac{\theta(\alpha-1)}{\alpha}} w_{it}^{-\theta} H_i^{R\,-\frac{\theta(1-\alpha)}{\alpha}}}. \tag{34}$$

Equivalently, $\bar{f}_i$ also equals the ratio of the RHS's of (31) and (33) that is

$$\bar{f}_i = \frac{\int_S \left[\frac{a_{jt}}{\tilde{u}_{jt}}\right]^{\frac{\theta}{\alpha}} w_{jt}^{1+\theta} H_j^{R\,\frac{\theta(1-\alpha)}{\alpha}} L_{jt}^{1+\frac{\theta(\alpha-1)}{\alpha}} \zeta_{ijt}^{-\theta}\, dj}{\int_S T_{jt} L_{jt}^{\theta(1-\mu)} w_{jt}^{-\theta} H_j^{C\,-\theta(\mu-1)} \zeta_{jit}^{-\theta}\, dj}. \tag{35}$$

Applying symmetric trade costs, $\zeta_{ijt} = \zeta_{jit}$, we can rewrite $\bar{f}_i$ as follows

$$\bar{f}_i = \frac{\int_S \bar{f}_j^{-\beta} \bar{\bar{f}}_{ij}\, dj}{\int_S \bar{f}_j^{-(1+\beta)} \bar{\bar{f}}_{ij}\, dj}, \tag{36}$$

where

$$\bar{\bar{f}}_{ij} = \left[\frac{a_{jt}}{\tilde{u}_{jt}}\right]^{\frac{(1+\beta)\theta}{\alpha}} T_{jt}^{-\beta}\ \zeta_{ijt}^{-\theta} H_j^{C\,\beta(\theta(\mu-1))} H_j^{R\,\frac{(1+\beta)\theta(1-\alpha)}{\alpha}} w_{jt}^{1+\theta+(1+2\theta)\beta}$$
$$L_{jt}^{(1+\beta)(1+\frac{\theta(\alpha-1)}{\alpha})-\beta(\theta(1-\mu))}. \tag{37}$$

Rewrite (36) as

$$\bar{\bar{\bar{f}}}_i = \frac{\bar{f}_i^{-\beta}}{\int_S \bar{f}_j^{-\beta}\bar{\bar{f}}_{ij}\ dj} = \frac{\bar{f}_i^{-(1+\beta)}}{\int_S \bar{f}_j^{-(1+\beta)}\bar{\bar{f}}_{ij}\ dj}. \tag{38}$$

Then, changing the notation to

$$\bar{g}_i = \bar{f}_i^{-\beta} \quad \text{and} \quad \bar{\bar{g}}_i = \bar{f}_i^{-(1+\beta)}, \tag{39}$$

and rewrite both as follows

$$\bar{g}_i = \int_S \bar{\bar{\bar{f}}}_i \bar{\bar{f}}_{ij}\bar{g}_j\ dj \quad \text{and} \quad \bar{\bar{g}}_i = \int_S \bar{\bar{\bar{f}}}_i \bar{\bar{f}}_{ij}\bar{\bar{g}}_j\ dj. \tag{40}$$

Define $\bar{\bar{\bar{f}}}_i\bar{\bar{f}}_{ij}$ as kernel K_{ij}. Hence, $\bar{g}_i$ and $\bar{\bar{g}}_i$ are both solutions to the integral equation

$$x_i = \int_S K_{ij}\ x_j\ dj, \tag{41}$$

where K_{ij} can be expressed as

$$K_{ij} = \frac{T_{it}L_{it}^{\theta(1-\mu)} w_{it}^{-\theta} H_i^{C\,-\theta(\mu-1)} \zeta_{ijt}^{-\theta}}{\int_S T_{jt}L_{jt}^{\theta(1-\mu)} w_{jt}^{-\theta} H_j^{C\,-\theta(\mu-1)} \zeta_{jit}^{-\theta}\ dj}, \tag{42}$$

or

$$K_{ij} = \frac{\left[\frac{a_{it}}{\tilde{u}_{it}}\right]^{\frac{\theta}{\alpha}} w_{it}^{1+\theta} H_i^{R\,\frac{\theta(1-\alpha)}{\alpha}} L_{it}^{1+\frac{\theta(\alpha-1)}{\alpha}} \zeta_{ijt}^{-\theta}}{\int_S \left[\frac{a_{jt}}{\tilde{u}_{jt}}\right]^{\frac{\theta}{\alpha}} w_{jt}^{1+\theta} H_j^{R\,\frac{\theta(1-\alpha)}{\alpha}} L_{jt}^{1+\frac{\theta(\alpha-1)}{\alpha}} \zeta_{ijt}^{-\theta}\ dj}. \tag{43}$$

We have to ensure that K_{ij} is (i) non-negative, (ii) measurable and (iii) square-integrable. Non-negativity holds as $\bar{\bar{f}}$ and $\bar{\bar{\bar{f}}}$ are non-negative. Measurability holds because it can be shown that $\bar{\bar{f}}$ and $\bar{\bar{\bar{f}}}$ are approximately continuous everywhere. Square-integrability holds as long as population at any given location is bounded from below and above. The former is true because by construction population cannot shrink to zero unless nominal wages are zero or amenities are infinitely high. The latter is true because population at any given location cannot exceed the level of world population $\bar{L}$.

Given the properties of K_{ij}, Theorem 2.19 in Zabreyko et al. (1975) guarantees that there exists a unique (to scale) strictly positive function that satisfies the system of equations in (41). Hence,

$$\bar{g}_i = \varpi \bar{\bar{g}}_i \quad \Rightarrow \bar{f}_i^{-\beta} = \varpi \bar{f}_i^{-(1+\beta)} \quad \Rightarrow \bar{f}_i = \varpi, \tag{44}$$

where ϖ is a constant. Hence,

$$\frac{T_{it}^{-1} w_{it}^{1+\theta} H_i^{C\,\theta(\mu-1)} L_{it}^{1-(\theta(1-\mu))}}{\left[\frac{a_{it}}{\tilde{u}_{it}}\right]^{-\frac{\theta}{\alpha}} L_{it}^{-\frac{\theta(\alpha-1)}{\alpha}} w_{it}^{-\theta} H_i^{R\,\frac{-\theta(1-\alpha)}{\alpha}}} = \varpi, \tag{45}$$

and solving for w_{it} gives

$$w_{rt} = \bar{w} \left[\frac{a_{it}}{\tilde{u}_{it}}\right]^{-\frac{\theta}{\alpha(1+2\theta)}} T_{it}^{\frac{1}{1+2\theta}} H_i^{C\,\frac{-\theta(\mu-1)}{1+2\theta}} H_i^{R\,\frac{-\theta(1-\alpha)}{\alpha(1+2\theta}} L_{it}^{\frac{-\theta(\frac{\alpha-1}{\alpha})+(\theta(1-\mu))-1}{1+2\theta}}, \tag{46}$$

where $\bar{w} = \varpi^{\frac{1}{1+2\theta}}$. Substituting (46) into (31) gives

$$\begin{aligned}
&\left[\frac{a_{it}}{\tilde{u}_{it}}\right]^{-\frac{\theta(1+\theta)}{\alpha(1+2\theta)}} T_{it}^{-\frac{\theta}{1+2\theta}} H_i^{C\,\frac{\theta^2(\mu-1)\theta}{1+2\theta}} H_i^{R\,\frac{-\theta(1+\theta)(1-\alpha)}{\alpha(1+2\theta)}} \\
&L_{it}^{-\theta(\frac{\alpha-1}{\alpha})+\frac{1}{1+2\theta}\left[-\theta(1-\mu)-\theta(\frac{\alpha-1}{\alpha})-1\right]} \\
&\quad = \psi_1 \int_S \left[\frac{a_{jt}}{\tilde{u}_{jt}}\right]^{\frac{\theta^2}{\alpha(1+2\theta)}} T_{jt}^{\frac{1+\theta}{1+2\theta}} H_j^{C\,\frac{-\theta(1+\theta)(\mu-1)}{1+2\theta}} H_j^{R\,\frac{\theta^2(1-\alpha)}{\alpha(1+2\theta)}} \\
&\qquad L_{jt}^{\theta(1-\mu)-\frac{\theta}{1+2\theta}\left[\theta(1-\mu)-\theta(\frac{\alpha-1}{\alpha})-1\right]} \zeta_{ijt}^{-\theta}\, dj.
\end{aligned} \tag{47}$$

Inserting (10) into (47) gives

$$\begin{aligned}
&\bar{B}_{it}\, \hat{u}_{it}^{\frac{1}{\Omega}\left[-\theta\left[\frac{\alpha-1}{\alpha}\right]+\frac{1}{1+2\theta}\left[\theta(1-\mu)-\theta(\frac{\alpha-1}{\alpha})-1\right]\right]+\frac{\theta(1+\theta)}{\alpha(1+2\theta)}} \\
&\quad = \psi_1 \int_S \hat{u}_{jt}^{\frac{1}{\Omega}\left[\theta(1-\mu)-\frac{\theta}{1+2\theta}\left[\theta(1-\mu)-\theta(\frac{\alpha-1}{\alpha})-1\right]\right]-\frac{\theta^2}{\alpha(1+2\theta)}} \bar{\bar{B}}_{jt}\zeta_{ijt}^{-\theta}\, dj,
\end{aligned} \tag{48}$$

where

$$\begin{aligned}
\bar{B}_{it} = \;& a_{it}^{-\frac{\theta(1+\theta)}{\alpha(1+2\theta)}} T_{it}^{-\frac{\theta}{1+2\theta}} H_i^{C\,\frac{\theta^2(\mu-1)\theta}{1+2\theta}} H_i^{R-\frac{\theta(1+\theta)(1-\alpha)}{\alpha(1+2\theta)}} \\
& \tilde{\kappa}_{it}^{-\theta\left[\frac{\alpha-1}{\alpha}\right]+\frac{1}{1+2\theta}\left[\theta(1-\mu)-\theta(\frac{\alpha-1}{\alpha})-1\right]},
\end{aligned}$$

and

$$\begin{aligned}
\bar{\bar{B}}_{jt} = \;& a_{jt}^{\frac{\theta^2}{\alpha(1+2\theta)}} T_{jt}^{\frac{1+\theta}{1+2\theta}} H_j^{C\,\frac{(1+\theta)(-\theta(\mu-1))}{1+2\theta}} H_j^{R\,\frac{\theta^2(1-\alpha)}{\alpha(1+2\theta)}} \\
& \tilde{\kappa}_{jt}^{\theta(1-\mu)-\frac{\theta}{1+2\theta}\left[\theta(1-\mu)-\theta(\frac{\alpha-1}{\alpha})-1\right]},
\end{aligned}$$

and

$$\hat{u}_{it} = \tilde{u}_{it} \left[\frac{\bar{L}}{\int_S \int_S \tilde{u}_{kt}^{1/\Omega} \kappa_{kjt}^{-1/\Omega} \, dkdj} \right]^{\Omega \left[1 - \frac{\theta}{\frac{1}{\Omega}[\theta(1-\mu) - \theta(\frac{\alpha-1}{\alpha})] + \theta} \right]}. \tag{49}$$

Rewrite (48) as

$$\bar{B}_i f_i^{\tilde{\beta}_1} = \psi_1 \int_S \bar{\bar{B}}_j \zeta_{ijt}^{-\theta} f_j^{\tilde{\beta}_2} \, dj, \tag{50}$$

and apply Theorem 2.19 in Zabreyko et al. (1975), then the solution $f_{(\cdot)}$ to equation (50) exists and is unique if (a) the functions $\psi_1 \bar{B}_i^{-1}$ and $\bar{\bar{B}}_j \zeta_{ijt}^{-\theta}$ are strictly positive and continuous, and (b) $\left| \frac{\tilde{\beta}_2}{\tilde{\beta}_1} \right| \leq 1$. The latter implies

$$\frac{\frac{1}{\Omega}\left[\theta(1-\mu) - \frac{\theta}{1+2\theta}\left[\theta(1-\mu) - \theta(\frac{\alpha-1}{\alpha}) - 1\right]\right] - \frac{\theta^2}{\alpha(1+2\theta)}}{\frac{1}{\Omega}\left[-\theta\left[\frac{\alpha-1}{\alpha}\right] + \frac{1}{1+2\theta}\left[\theta(1-\mu) - \theta(\frac{\alpha-1}{\alpha}) - 1\right]\right] + \frac{\theta(1+\theta)}{\alpha(1+2\theta)}} \leq 1,$$

which after some simplification can be written as the uniqueness condition (14) that is stated in Section 4.6:

$$\frac{\alpha}{\theta} \leq (1 - \mu\alpha) + \Omega.$$

E Stability Condition For Amenity- And Productivity-dispersion Processes

Note that the form of the diffusion processes is analogous for log amenities and log productivity. Therefore, it suffices to discuss the stability of the condition for amenities, here. Let us use $\log \boldsymbol{a}^* = (\log(a_i^*))$ to refer to the 331×1 vector of log-transformed long-run values of amenities, $\boldsymbol{W}^{R*} = (\mathbb{W}_{ijt-1}^R)$ to the 331×331 matrix of measuring R-centrality in the long run, and $\log \widetilde{\boldsymbol{L}}^* = (\log(L_i^*) - \log(H_i^R))$ for the 331×1 vector of log-transformed long-run values of population density (in terms of land for housing). Moreover, note that in equilibrium, the shocks ε_i^{a*} does not change anymore. Then, the log-amenity-determining in equation (2) in equilibrium can be written as

$$\log \boldsymbol{a}^* = -\lambda \boldsymbol{W}^{R*} \log \widetilde{\boldsymbol{L}}^* + \boldsymbol{\varepsilon}^{\boldsymbol{a}*}. \tag{51}$$

Note that all elements of $\log \widetilde{\boldsymbol{L}}^*$ are finite by definition (as the world population stays constant and so does land), and the column elements of $\boldsymbol{W}^{R*}$ sum up to less than or equal to unity in each row and are all non-negative. Hence, as long as λ and all entries of $\boldsymbol{\varepsilon}^{\boldsymbol{a}*}$ are finite, which is guaranteed,[52] all entries of $\log \boldsymbol{a}^*$ are finite as well.

[52] Note that ε_{it}^a is finite in the benchmark year $t = 2000$ by definition, as a_{i2000} is finite for all regions i in the

Finite long-run productivity levels are guaranteed by analogy.

F Cobb-Douglas Parameter in Preferences and Production

We measure employment (L_{it}) by combining population data from the Socioeconomic Data and Application Center (SEDAC), time-varying labor-force participation rates from the International Labor Organization Statistics Database (ILOSTAT), and time-varying labor-force-in-population shares from the World Bank's World Development Indicators (WDI) for the years 1990-2015.[53] Hence, we obtain the employment measurement by multiplying the population count from SEDAC with the labor-force participation rate from ILOSTAT and the share of population in employment age (15-64) from WDI. Data on the average annual wage per capita (w_{it}) in RMB for the years 2000-2009 come from the Chinese Annual Survey of Industrial Firms (CASIF) provided by the National Bureau of Statistics China. As these data are geo-referenced, we can compute average wages by prefecture for each year covered.

Residential and commercial land use data (H_i^R, H_i^C) for all 330 prefectures and the RoW come from the sources described in Section 5.2. Residential land rents (r_{it}^R) in RMB per square meter for 105 prefectures between 2008-2016 are provided by the Ministry of Natural Resources in China (Department of Land Use Management). Commercial land rents (r_{it}^C) in RMB per square meter are available for 340 prefectures between 1992-2014 from the National Bureau of Statistics China. Both residential and commercial land-rent data represent purchasing prices and need to be converted to annual rental-price data. We assume that annual rental prices correspond to a hypothetical annual mortgage payment if one had to borrow the total purchasing price. The generic annual mortgage payment M is calculated using the formula $M = 12 \times \left(P \frac{\varsigma(1+\varsigma)^n}{(1+\varsigma)^n - 1} \right)$, where P is the generic total purchasing price, ς is the monthly interest rate, and n is the total tenure of the mortgage, which we assume to be $n = 240$ months (or 20 years) for all regions and years. We use data on annual interest rates for China from the World Bank database.

data. Keeping it constant at that level fulfills that ε_i^{a*}. Moreover, λ is estimated to be finite.

[53]SEDAC provides gridded population data with an output resolution of 30 arc-seconds (corresponding to approximately 1×1 kilometer at the equator) for a five-year interval between 1990-2015 for the whole world (including China). We fill in data for the missing years within the intervals using linear interpolation. The data from ILOSTAT and WDI vary by year and country and are assumed to be identical across regions within a country.

G Data Source: Transportation Atlases

Table 6: Sources of Chinese Transportation Atlases

Year*	Title (English)	Title (pinyin)	Publisher (English)	Editor
2001	New Atlas of China's transportation	Xinbian zhongguo jiaotong dituce	SinoMaps Press	
2002	China Transportation Atlas	Zhongguo jiaotong tuce	Planet Map Press	
2003	General Transportation Atlas of China	Tongyong zhongguo jiaotong dituce	Hunan Map Publishing House	Han Jianzhong
2004	Practical China Transportation Atlas	Shiyong zhongguo jiaotong dituce	Xi'an Map Publishing House	
2005	China Transportation Atlas	Zhongguo jiaotong tuce	Planet Map Press	
2006	China Transportation Atlas	Zhongguo jiaotong tuce	Planet Map Press	
2007	China Transportation Atlas	Zhongguo jiaotong dituce	Shandong Province Map Publishing House	Wang Huaibao
2008	China Transportation Atlas	Zhongguo jiaotong dituce	Shandong Province Map Publishing House	Wang Huaibao
2009	China Transportation Atlas	Zhongguo jiaotong dituce	Shandong Province Map Publishing House	Wang Huaibao
2010	New Atlas of China's transportation	Xinbian zhongguo jiaotong dituce	Fujian Province Map Press	
2011	China Transportation Atlas	Zhongguo jiaotong dituce	Shandong Province Map Publishing House	Yuan Xinfang
2012	China Transportation Atlas	Zhongguo jiaotong dituce	SinoMaps Press	Chen Zhuoning
2013	China Transportation Atlas	Zhongguo jiaotong dituce	SinoMaps Press	Chen Zhuoning
2014	General Transportation Atlas of China	Tongyong zhongguo jiaotong dituce	Hunan Map Publishing House	Jiang Gonghe

Notes: *corresponds to the year of publication. A transportation atlas that is published in a given year contains maps of the previous year.