High-speed rail and the spatial distribution of economic activity: Evidence from Japan's Shinkansen*

Hans R.A. Koster† Kazunobu Hayakawa‡ Takatoshi Tabuchi§
Jacques-François Thisse¶

May 1, 2023

Abstract

We investigate the effect of high-speed rail (HSR) on welfare and the spatial distribution of economic activity. We develop a spatial quantitative general equilibrium model that incorporates trade between final firms producing goods and intermediate firms providing services, as well as mode choice and commuting and residential location choices. The model is estimated for Japan's Shinkansen, i.e., the first HSR ever built. The Shinkansen had a sizable positive welfare effect that is considerably larger than the welfare effect of highways. Local effects on employment, rents and wages can be substantial. These results show that HSR plays a non-negligible role in facilitating the provision of business-to-business services.

Keywords: high-speed rail; employment; population; input-output linkages; commuting.

JEL classification: D04; H43; R42.

*We are grateful to Nobuaki Hamaguchi, Keisuke Kondo, Antonio Russo, Yukiko Saito, Yoichi Sekiguchi, Philip Ushchev, Yue Yu, and participants to the RIETI Workshop in Tokyo, the 6th Urban Economics Workshop in Barcelona, the 11th European Meeting of te Urban Economics Association in London, and a MUFG Bank seminar, as well as a research seminar in Amsterdam. This study uses data on firm networks from Tokyo Shoko Research Ltd (TSR), which is kindly provided by the RIETI.

†Department of Spatial Economics, Vrije Universiteit Amsterdam, Tinbergen Institute, and CEPR. E-mail: h.koster@vu.nl.

‡Institute of Developing Economies (Japan). E-mail: kazunobu_hayakawa@ide-gsm.org.

§Faculty of Global Management, Chuo University (Japan). E-mail: ttabuchi@tamacc.chuo-u.ac.jp.

¶Institute of Developing Economies (Japan) and CEPR. E-mail: jacques.thisse@uclouvain.be.

1 Introduction

The economic and social consequences of investments in transport infrastructure generate heated academic and policy debates because they typically involve costly investments that are supposed to yield high payoffs. Japan is known for having worldwide some of the best transport infrastructure and the high-speed railway network, and the Shinkansen, is considered as its jewel. We will see that the Shinkansen has increased gross welfare by about 20% and this effect is considerably larger than the welfare effect of highways through reducing the costs of transporting business-to-business services. Because high-speed rail (HSR) plays a more important role than roads in facilitating the provision of these services, the welfare effects of highways are considerably smaller. However, this does necessarily mean that planned Shinkansen lines (such as a Maglev train connecting Tokyo and Nagoya whose construction cost is about $50 billion) are worth undertaking.

It is well known that modern economies are heavily reliant on the provision of services, rather than on the production of goods. Actually, the cost of transporting goods has diminished in the last century (Glaeser and Kohlhase, 2004; Redding and Turner, 2015). At the same time, transportation costs of people are still considerably high. Many studies have shown that commuting costs are important and can explain why workers aim to live close to their workplace (Su, 2022). Still, what is largely overlooked is that the provision of business-to-business services also requires movement of people. Actually, the value of time for business travelers is at least twice as high as the value of time of commuters (Abrantes and Wardman, 2011).

A transport mode that facilitate business-to-business travel is *high-speed rail* (HSR). High-speed trains usually run at speeds exceeding 250 km/h and are a competitor to the airplane on medium-distance travel (Behrens and Pels, 2012).[1] We focus on Japan's high-speed rail: the *Shinkansen*, which was supposed to promote economic growth and development outside Tokyo (Sato, 2015). For four decades from its opening in 1964, the Shinkansen was the only HSR service outside of Europe and still is considered to be one of the most successful implementations of an HSR, being efficient, punctual, and frequent. The Tokaido Shinkansen connecting Tokyo, Nagoya, and Osaka is one of the world's busiest high-speed rail line carrying over 150 million passengers each year. In 2010, the share of train travel is 43.7% for trips between 300 and 500km, while it reaches almost 70% for trips between 500 and 700km (*Ministry of Land, Infrastructure, Transport and Tourism*, 2019). The Shinkansen is heavily used for business travel. Out of 160

[1] Within 10 years China has developed the most extensive HSR network, which is now about 35 thousand km and still expanding. In Europe, there are concrete plans to open HSR lines between London and Manchester in the United Kingdom and between Warsaw and Tallinn in the Baltic. Further, the Spanish government has an ambitious plan to expand the HSR network to 7 thousand km (it is now about 3.2 thousand km). The U.S. currently has one HSR under construction between Los Angeles and San Francisco and has plans to upgrade the existing Northeast Corridor line to operate at a higher speed.

million passengers per year, a very large share (about 65% in 2010) are technical workers and business travelers who provide services as intermediate inputs in the destination regions. Since the share of non-production workers in Japan has increased from 22% to 41% between 1952 and 2015, such a high number of professional trips strongly suggests that the Shinkansen may be considered a transport mode that has affected significantly firms' location choices (see Bernard *et al.,* 2019, for empirical evidence).

The construction of new HSR lines is extremely costly. For example, the total construction cost of the first HSR line in Japan, *i.e.*, the Tokaido Shinkansen connecting Tokyo and Osaka, was about $12.5 billion in current prices, which amounted to 1% of the Japanese GDP in 1965 (Sato, 2015). More recently, the *Ministry of Land, Infrastructure, Transport and Tourism* of Japan reported that the estimated cost of the Chuo Shinkansen, *i.e.*, a Maglev train connecting Tokyo and Nagoya that is expected to be completed in 2045, is $50 billion, *i.e.*, 1.3% of the GDP in 2021.

Given the large costs of building HSR lines, it is surprising that the question of whether and how HSR affects welfare, land rents and the overall spatial distribution of economic activity has not been satisfactorily answered. There are only a few papers that investigate the impacts of HSR, but most of the evidence is reduced-form, usually showing positive effects of HSR on employment, GDP, and land rents of central locations, while intermediate locations sometimes may lose (Zheng and Kahn, 2013; Qin, 2017; Ahlfeldt and Feddersen, 2018, Okamoto and Sato, 2021; Koster *et al.*, 2022). Standard cost-benefit analyses show mixed evidence (De Rus and Inglada, 1997; Dijkman *et al.*, 2000; Coto-Millán *et al.*, 2007).

In this paper, we investigate the effects of the Shinkansen on the spatial distribution of economic activity and welfare within Japan. To achieve our goal, we first develop a spatial quantitative model that suits well the main features of the Japanese economy. First, as the big majority of Shinkansen users are business travelers who provide business-to-business services, we must distinguish between different types of firms. We work with final firms providing goods and intermediate firms providing services to the former. Second, since the Shinkansen is intensively used by service providers, in contrast to most papers that focus on travel time by road, we allow intermediate services to be shipped by road or railway to final producers.[2] Third, in previous spatial general equilibrium models, land consumption by firms is typically ignored. As land is regarded as the scarcest resource in Japan, ignoring land consumption by firms would be problematic (Rose, 1992). More specifically, we assume that land is an input for both the final and intermediate sectors. Last, commuting times are long and more than 51 minutes one way in Tokyo metropolitan area, whereas in Osaka metropolitan area, one-way commutes are 43 minutes on average (NHK, 2015). Because public transport is an important mode in Japanese cities, we allow workers to

[2] We also consider an extension where we allow for air travel, but this does not materially affect our estimates.

commute to their workplaces by commuter trains or by car. Because the Shinkansen is hardly used by commuters because it is prohibitively expensive to use on a daily basis, we assume that workers only will take commuter trains to work.[3]

We then estimate, rather than calibrate, our model using data on municipalities in Japan. Often, spatial equilibrium models are calibrated using parameter estimates from other contexts. This can be problematic because model parameters depend on the local context. Moreover, for various parameters in our model no clear priors exist, such as for the elasticity of substitution between intermediates, or the mode-specific transport costs of intermediate services. We show that the parameters can be identified using a recursive estimation approach based on the structure of the model.

Using the estimated parameters, we study what happens to the population and employment in each municipality if we change travel times. By exploiting data on employment and travel times in 1957, we show that our model is able to back-cast the change in employment well and outperforms variables capturing predicted employment based on the Euclidean distance between locations. This underscores the importance of transport networks in explaining the current spatial employment distributions and shows that our model is able to explain long-term changes in the spatial distribution of employment.

We then consider three counterfactual experiments, *i.e.*, *(i)* there is no Shinkansen, *(ii)* all planned extensions of the Shinkansen are realized, and *(iii)* all highways are removed. The results of these counterfactuals highlight a few important outcomes. Our first experiment shows that removing the Shinkansen would decrease the average gross welfare by 20% (with the lower bound being about 3%), thus showing that *the Shinkansen has generated substantial welfare benefits.* Taking into account the construction and operation costs, it appears that the benefits of the Shinkansen far outweigh the costs. On average, municipalities that are close to a Shinkansen station have attracted employment (the effects are about 25%, which are in line with reduced-form estimations).

The second experiment shows that planned extensions yield a considerably lower welfare effects that are close to zero. This shows that the effects of investments in upgrading or extending existing lines are likely considerably lower.

Interestingly, the welfare effects of removing highways are also considerably smaller than those generated by removing the Shinkansen, which underlines the importance of business-to-business travel by train and *the pivotal role of the Shinkansen in sustaining and developing inter-firm trade in Japan's spatial economy.*

[3] The share of commuters that use the Shinkansen is just 0.05%, which is obtained from the *Statistical Survey on Railway Transport* in 2010.

Related literature. There are two major methodological issues in assessing the impact of new transport infrastructure on specific regions (Redding and Turner, 2015). The first is the chicken-and-egg problem as regions with high transport needs are likely to receive infrastructure. In this case, the construction of new transport infrastructure is endogenous. The second issue is that the effects of infrastructure on individual localities are hard to predict because it is unclear whether infrastructure will attract new activities or displace activities from other regions. As emphasized by 'classical' location theory, the relative position of a location in the transport network is key for understanding whether this location will attract new activities (Thomas, 2002; Behrens *et al.*, 2007). Location theory also stresses the importance of the size of local markets for firms' location choices (Koster *et al.*, 2022). Since a bigger pool of firms should attract consumers/workers, the size of local markets should be endogenous. Last, location fundamentals (*i.e.*, first nature) may affect firms' and workers' spatial choices in various ways. Those various difficulties explain probably why the empirical evidence on the expected benefits of large investments in transport infrastructure is mixed. In particular, it is still unclear whether and which locations benefit from being connected to the transport network.

Still, the literature devoted to the economic impacts of transport infrastructures has grown fast. In two thorough papers, Donaldson and Hornbeck (2016) and Donaldson (2018) highlight the positive effect of the development of railroads in the U.S. from 1870 to 1890 and in colonial India (1870 to 1930), respectively. They use a standard model of trade and show that railroads affect agricultural production and land values. Berger and Enflo (2017) analyze the effects of 150 years of railways on urban growth in Sweden. They find that the connection to the railway resulted in a strong increase in population in the first 20 years, but railways were much less effective in spurring population growth in the long run. Banerjee *et al.* (2020) measure the positive, but modest, impact of roads on the economic growth of Chinese cities. As for the effects of large transport infrastructure on peripheral areas, Faber (2014) and Baum-Snow *et al.* (2017, 2020) find that highways have led to a relative reduction in GDP among unconnected peripheral Chinese counties.

Despite the large amounts of money flowing to the construction of HSRs, the effects of large-scale railway investments are understudied. There are a few exceptions. Zheng and Kahn (2013) argue that China's HSR facilitates suburbanization and market integration. Qin (2013) finds that non-urban counties on the upgraded railway lines experienced reductions in GDP per capita following the upgrade. Ahlfeldt and Feddersen (2018), however, provide evidence that access to an HSR leads to an increase in GDP by 8.5% in three counties between Cologne and Frankfurt (Germany). These papers adopt a reduced-form framework, which is probably the reason for the seemingly contradictory results. By decreasing passenger travel time between headquarters and affiliates, the development of the HSR network in France has allowed management functions to be concentrated in headquarters (Charnoz *et al.,* 2018). This points toward the specialization of centers and peripheries. By proposing a spatial quantitative model where we include travel times by rail and by road while investigating the effects on central and peripheral areas, we aim to reconcile these findings within a unifying framework.

A related paper by Bernard *et al.* (2019) proposes a model of buyer-seller relationship formation and shows that lower search and outsourcing costs lead to more buyer-seller links. Using an extension of the Shinkansen opened in 2004, they find that a reduction in travel costs by HSR has large effects through inter-firm trade. In contrast to them, our model is not only concerned with linkages between firms in different locations, but also with the location of both employment *and* population. We therefore consider commuting flows between locations, as they constitute a large part of traffic between close locations and influence the local employment elasticities with respect to infrastructure shocks. Furthermore, we also include a land market and allow for mode choice for firms and workers.

Our paper is also related to Monte *et al.* (2018). However, we differ from them in several important respects. First, we include trade in intermediate services, like Ethier (1982), and (final) consumption goods, like Krugman (1980). Second, we assume that land is consumed by workers, as well as by firms in the intermediate and final sectors. We estimate that about 7% of the expenditures of final firms are on land, which is non-negligible. Third, we allow for modal choice in commuting and business trips. Finally, we estimate the business travel time elasticities as well as the elasticity of substitution rather than choosing somewhat arbitrary values.

The remainder of the paper is organized as follows. In Section 2, we provide a short historical survey of the development of the Shinkansen network; we also discuss our data sources and present reduced-form effects of Shinkansen stations on employment density. In Section 3, we present our model. In Section 4, we explain the estimation of the model parameters, and report and discuss the estimation results. Section 5 investigates the performance of our model through three counterfactual experiments, while Section 6 concludes.

2 Data and context

2.1 The Shinkansen network and highways

High-speed rail. The *Tokaido* local railway line opened in 1889 and is the first railway that connects Tokyo, Nagoya, Osaka, and Kobe. The *Sanyo* local line, the second railway built in Japan, connects Kobe with the northern tip of Kyushu. In the 1930s, transport capacity along the local Tokaido-Sanyo railway line almost reached its limit due to the demand increase in transport to Korea and China. This situation called for a large capacity increase in this corridor.

Consequently, the first plan for a 'Shinkansen' (meaning: 'new trunk line') was approved by the Imperial Diet in 1940 and extended in 1942. This Shinkansen was supposed to run between Tokyo and Shimonoseki at a speed of 200 km/h. This HSR was based on new railway tracks, but the initial plan suggested stops at several stations located along the Tokaido and Sanyo local lines. Specifically, 18 municipalities were selected for the construction of stations. Some land acquisitions and the construction of a few tunnels were completed in the early 1940s.

Owing to the high economic growth of Japan in the 1950s, there was a renewed call for the construction of a Shinkansen. The construction plan of the Tokaido Shinkansen was approved in 1959 and construction was completed in 1964. The first Shinkansen line enabled travel times between Tokyo and Shin-Osaka of 3 hours and 10 minutes. The maximum speed of operation gradually increased over time to 220 km/h in 1986. A drastic gain in speed was achieved in 1992 by the introduction of a new train, called the *Nozomi*, which connected Tokyo and Shin-Osaka in 2.5 hours at an average speed of 270 km/h. The frequency of operation also increased over time.

Soon after 1964, the Shinkansen was extended to connect other cities. The Sanyo Shinkansen connected Shin-Osaka to Hakata in Fukuoka prefecture (at the northern tip of Kyushu). The Shinkansen plan from 1942 served again as a reference in the construction of the Sanyo Shinkansen. The services between Shin-Osaka and Hakata started operating in 1975. As a result, the traveling time between Tokyo and Hakata was reduced to just seven hours.

Other Shinkansen lines were constructed in the hope of boosting rural areas in Japan, based on a plan that was approved in 1972 (Sato, 2015), including the Tohoku Shinkansen and Joetsu Shinkansen in 1982, the Hokuriku Shinkansen in 1997, the Kyushu Shinkansen in 2004, as well as the Hokkaido Shinkansen in 2010, which will be extended to Sapporo in 2030.[4]

Figure 1, which provides a map of the railway network in Japan, shows that the Shinkansen network covers most of Japan, except Hokkaido where an extension to Sapporo is planned. Trains travel the fastest on the Tokaido-Sanyo line, which connects Japan's largest cities.

Figures 2A and 2B show the average travel time by railway across Japanese municipalities in 1957 and 2013. One observes that the average travel time has been reduced from 27 hours to just over 11 hours, *i.e.*, a reduction of 60%. Most of the travel time reductions are in and around Tokyo where there is a strong concentration of population.

[4] Along the Tohoku Shinkansen, there are two "mini-Shinkansens," which have a maximum operating speed of only 130 km/h. The first one is the Akita Shinkansen, which connects Morioka and Akita in Akita prefecture; it was completed in 1997. The second one is the Yamagata Shinkansen, which was completed in 1999. It operates between Fukushima in Fukushima prefecture and Shinjo in Yamagata prefecture. The crucial difference between a mini-Shinkansen and a local railway is that the former can operate on Shinkansen tracks with a maximum speed of more than 300 km/h. Since the Tohoku Shinkansen has stops at Morioka and Fukushima, these mini-Shinkansens can operate directly from Tokyo station by using the Tohoku Shinkansen tracks.

Figure 1 — Overview of Japan's Railway Network

In 2017, the total distance of domestic passenger transport in Japan amounted to approximately 605 billion passenger kilometers, with railway transport accounting for 72.3% of the transport distance, while motor vehicles only account for 11.3% (and air travel for 16.4%). Moreover, 50% of the trips by motor vehicles are mostly for commuting purposes. For medium- and long-distance travel, the train share is higher: in 2010, the share of trips by train is 43.7% between 300 and 500km and 70% between 500 and 700km. Hence, *for long-distance travel, the train is by far the most preferred transport mode.* The Shinkansen tends to be used for attending business meetings rather than for commuting. For example, according to the 2017 JR Tokai Media Guide, business users account for the largest share (67%) in the Tokaido Shinkansen users, followed by private travelers for sightseeing (12%). Only 0.9% of travelers in a Shinkansen train use the Tokaido Shinkansen for commuting. Moreover, the overall share of commuters that use the Shinkansen is just 0.05% (see *Statistical Survey on Railway Transport* in 2010).

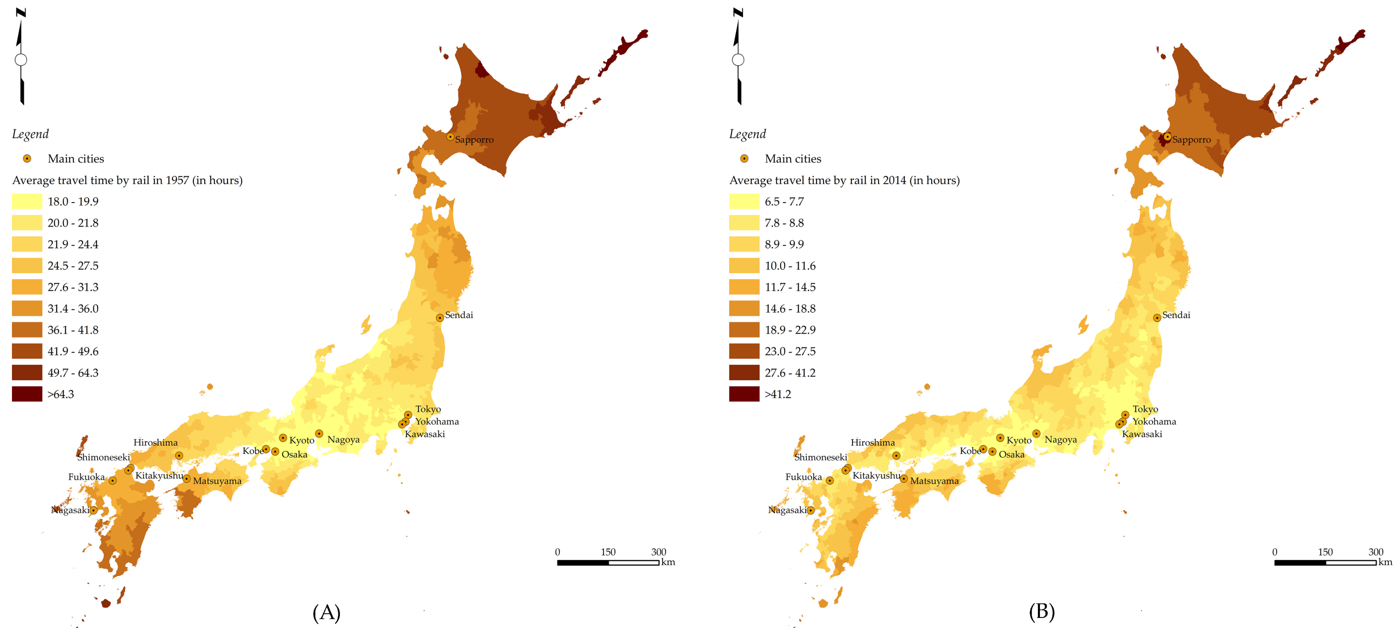


Figure 2 — Overview of Japan's Railway Network

Regarding the shipment of goods, the picture is quite different. In 2017, 91.5% of the domestic freight transportation is made by trucks, 7.5% is by sea, but only 0.9% is by rail (Ministry of Land, Infrastructure and Tourism, 2019). Therefore, railways in Japan are almost exclusively used for the transportation of people.

Highways. Express highways were planned to be built in 1943 but were not constructed during the war. After the war, the share of paved roads was only 1.2% of the road network. As the Japanese economy grew substantially in the subsequent decades, the number of cars and trucks increased rapidly. The need for roads for freight and passenger transport increased substantially. The first highway in Japan was completed in 1963 linking Osaka and Kyoto. In 1965, Nagoya and Osaka were connected, while in 1969, the highway between Tokyo and Nagoya was completed. In the 1970s, the highway networks were expanded to more peripheral areas, including Hokkaido, the Tohoku region, and Kyushu. According to the *Road Statistics Annual Report* from the *Ministry of Land, Infrastructure and Transport,* the length of highways totaled around a thousand km in 1973, around 5 thousand km in 1992, and around 9 thousand km in 2016. The average travel time by road has been reduced by about 55% since 1957.

Like Shinkansen trains, highways are hardly used for commuting because of sizable tolls. For example, according to the survey for 2012 conducted by *MyVoice Communications, Inc.*, only 2.5% of highway users are commuters.

2.2 Data sources

Our analysis will be undertaken at the municipality level. There are 1,719 municipalities in Japan. Since the boundaries of the municipalities have been revised several times, we have amalgamated historic municipalities to match those in 2014.[5] In the analysis, we only kept municipalities on Honshu, Hokkaido, Shikoku, and Kyushu, which we refer to as Mainland Japan. We thus consider 1,658 municipalities whose average (median) population is almost 75 thousand (26 thousand).[6]

In contrast to the literature that uses dummy variables to describe (improvements in) accessibility, in our quantitative model we will adopt a *network approach*. That is, we use detailed information on the railway and highway networks to calculate the travel time between any two locations in a specific year. The data on the railway network is from the *National Land Numerical Information*. For each railway line, we know the opening date so that we can construct the railway network each year for which we have data. From the *JTB Timetable* and the *JR Timetable* (*Kotsu Shinbunsha*), we obtain information on the average speeds on railways in Japan over the years. In 1957 the average speed was only 38 km/h, while it increased to 60 km/h in 2015.[7] For the Shinkansen lines, the average speed is 130 km/h, while for the fast Tokaido-Sanyo line it increased to 250 km/h after 2000. For each year, we calculate the distance from each municipality centroid to the nearest railway station. We assume that the speed to travel to the nearest station is 1/4 of the average speed on the railway network.[8] In this way, we may assess the time people need to get to the station by car, other public transport, or bicycle. Then, for each municipality pair we calculate the travel time over the network.[9]

We also have data on the highway network since the 1960s from the *National Land Numerical Information*. For the highways we obtain average speeds from the *Road Traffic Census*. We also use the information on the underlying road network from 2015. Unfortunately, we do not have time-series data for roads other than highways. Hence, we assume that the road network has not changed during our study period (*i.e.*, from 2000 onwards). Indeed, according to the *Road Statistics Annual Report* in Japan, the total length of highways increased by 28% from 2000 to

[5] We redraw municipal boundaries using the information provided in `https://uub.jp/upd/` and `http://toshidata.web.fc2.com/dantai_code.html`.

[6] The information about each municipality's location (*i.e.*, longitude, latitude, and shapes) is provided by the *National Land Numerical Information*. The data on geographical area are drawn from the *Census of Population* in 2015.

[7] These speeds are computed by dividing the route distance by the actual time when leaving Tokyo station for Shimonoseki station by local trains.

[8] The value 1/4 is arbitrary. We have played around with different values, but this makes very little difference in the results.

[9] For each municipality pair, we also calculate the Euclidean travel time, defined by the Euclidean distance multiplied by the travel speed (which is again 1/4 of the average speed on the railway network). For each municipality pair, we then take the minimum of railway travel time and Euclidean travel time. In this way, municipalities that are close to each other, but which do not have a rail connection, will not be separated by an unrealistically long travel time.

2015, while the increase in non-highways (*e.g.*, national roads) was only 4%.

Importantly, we make a distinction between the travel time by commuters and by business travelers. As discussed earlier, highways are hardly used by commuters because of the high tolls. Likewise, the Shinkansen is hardly used by commuters because of the longer distances between stations and expensive tickets. Hence, we calculate the travel time *for commuting* for each municipality pair while disregarding the Shinkansen network and highway links.

To be able to estimate the model that we will propose in Section 3, we further obtain data on *(i)* trade flows of intermediate services, *(ii)* commuting flows, *(iii)* population and employment, *(iv)* land rents, *(v)* wages, *(vi)* geographic characteristics, and *(vii)* historical data.

(i) We obtain yearly data on production networks between 2007 and 2017 from *Tokyo Shoko Research Ltd* (*TSR*). *TSR* provides information on credit reports of firms in 18 sectors on potential suppliers *and* customers. Each firm provides a list of the 24 most important suppliers and customers by decreasing order each year between 2007 and 2017. However, note that for a few large firms we observe many more input suppliers. The reason is that for many small suppliers a large firm is likely to be one of their most important customers. The database contains information on more than one million firms, which is a representative sample of the population of firms in Japan (Bernard *et al.*, 2019). We can identify the location of each firm at the municipality level. The *TSR* data is at the firm level, rather than the establishment level, which means that we only know the location of the firm's headquarters. Hence, in the main specifications we focus on single-plant firms, which means that we keep 33% of the input-output linkages. We will show that our results are robust to the inclusion of multi-plant firms.

Unfortunately, the *TSR* data does not provide information on the *value* of trade links. Since we are interested in the expenditure share, S_{ji}, by final firms located in i on intermediate services produced in j, we use input-output data at the national level for 2005, 2011, and 2015. These data provide the value of trade between the 9 different sectors provided in the *TSR* data. We then count the number of linkages between municipalities by sector-pair and normalize the number of linkages by overall industry-pair trade flows from intermediate services firms to final firms.[10] The correlation between actual linkages between municipalities and 'normalized' trade values is 0.966 for all firms and 0.941 for single-plant firms.

(ii) We gather data on commuting flows between municipalities for 2000, 2005, 2010, and 2015, which are drawn from the *Census of Population.* We focus on commutes by workers who are older than 15 years.

(iii) We obtain the municipality-level data on population from the *Census of Population* for 1955 every five years until 2005, as well as for 2008 and 2013. Municipality-level data on employment by industry are obtained from the *Establishment Census* for 1957, 1972, 1978, 1981,

[10] An illustration of this procedure is provided in Appendix A.1.

1986, and 1991, the *Establishment and Enterprise Census* for 1996, 2001, and 2006, and the *Economic Census for Business Frame* for 2009 and 2014. These censuses cover all establishments in Japan in 9 broad sectors, including manufacturing, energy production, mining, transportation, construction, finance and insurance, real estate services, and consumer services. In the model laid out in the next section, we distinguish between an intermediate sector producing services and a final sector producing goods. Following a common definition in the literature, we define intermediate services firms to be firms in finance, insurance and real estate (FIRE). The FIRE sector is a prominent part of the service industry in the United States and in other developed countries and is argued to be responsible for a large part of the economic growth in the last decades.[11]

Since the commuting data are available from 2000 onwards, we use data on employment and population between 2000 and 2015 to estimate the model. However, we will use data from 1957 to verify whether our model delivers meaningful predictions.

(iv) We gather data on assessed land prices since 1983 from the *Ministry of Land, Infrastructure, Transport, and Tourism.* For almost 26,000 plots we observe the assessed land price. We know the land use (residential, commercial, industrial, and forest). To obtain the predicted land rent in each municipality we estimate regressions with land use dummies, prefecture-by-year fixed effects, as well as municipality fixed effects. We then use the predicted land rents per m^2 for a residential plot with a median size as the observed land rent. For a few municipalities with missing data, we use the land rent observed in adjacent municipalities.

(v) Regarding wages, we use the total taxable income in a municipality divided by the number of taxpayers. Data are obtained from the *Report of Taxation Status on Municipal Taxes* (*MIAC*). We construct wage data at the municipality level using data on wages in manufacturing at the municipality level from the *Census of Manufacture* and wages for all sectors at the prefecture level from the *Monthly Labour Survey* by the *Ministry of Health, Labour, and Welfare*, as well as employment shares at the municipality level. We describe this procedure in Appendix A.2.

(vi) We further compile data on geographic characteristics. First, using land cover data between 2014 − 2016 from the *Earth Observation Research Center* (*JAXA*) we determine for each municipality the share of developed land. Following Saiz (2010), we calculate the share of developable land in each municipality, using very fine-grained 30 by 30m data on elevation and slopes, as well as detailed information on water bodies. In Appendix A.3, we outline the exact procedure to calculate the share of developable land in each municipality. Further, from *JAXA* we obtain data on the mean elevation in each municipality.

[11] We emphasize that we have considered a wide range of alternative definitions and ways to determine the intermediate sector (*e.g.* using the share of intermediates produced for other sectors, or whether the share of intermediates exceeds 50%). This appears not to materially influence our results.

Table 1 — Descriptive statistics for commuting data

	(1) mean	(2) sd	(3) min	(4) max
Number of commuters	479.7	7,884	0	1083738
Travel time by train *(min)*	74.2831	32.1005	0	257
Travel time by road *(min)*	107.2520	48.6603	0	394
Euclidian distance *(log)*	46.0534	21.2675	0.0088	125.4
Location were connected via the network in 1942	0.0954	0.2937	0	1
Locations were connected via the network in 725	0.1246	0.3303	0	1
On different sides of the east-west border	0.0287	0.1669	0	1
On same island	0.9931	0.0825	0	1
Year of observation	2,008	5.5883	2,000	2,015

Notes: We exclude pairs that are further than 120 minutes apart by both train and car. The number of observations is $499,882$. Travel times for commuters are calculated excluding highways and Shinkansen links.

(vii) Eventually, for some robustness checks of the model's parameters, we will rely on historical data on infrastructure networks and population, going back to the 8^{th} century. We describe the compilation of these data in Appendix A.4.

2.3 Descriptives

In Table 1, we report descriptive statistics for the commuting data for areas that are within 4 hours traveling from each other. In total, we observe about 60 million commuters each year on mainland Japan. Observe that 83% of the one-way commutes are less than 30 minutes long, while 97% are less than an hour (we show a more detailed histogram in Appendix A.5). The average travel time over all OD-pairs by train is then 155 minutes. However, if we weight by the number of commuters on each link, it is only 12 minutes. Similarly, the average travel time by road is 211 minutes, but the weighted travel time is only 17 minutes.

Table 2 shows descriptive statistics for the *TSR* data. There are $17,268,734$ input-output linkages between firms. We normalize these linkages to trade value between sectors. Note that 67% of the trade value is within 1 hour traveling by train, 77% within 2 hours, and 90% within 4 hours. Unsurprisingly, the 'spatial decay' of firms' trade thus seems to be lower than for commuting (we show a more detailed histogram in Appendix A.4). The travel time by train is on average almost 10 hours, but if we weight it by the trade value, it is almost 90 minutes. Similarly, the average travel time by road is almost 15 hours, but only about 2.5 hours if we weight it by the trade value.

Table 3 reports descriptive statistics for the employment and population data for 1957 and 2014. The average population density is $1,073$ per km^2, while it is about half this value for employment density. Population and employment are highly correlated: the correlation between

Table 2 — Descriptive statistics for TSR data

	(1) mean	(2) sd	(3) min	(4) max
Total value of inputs	15.342	786.8	0	505,599
Travel time by train *(min)*	593.9	492.7253	0	3,322
Travel time by road *(min)*	891.6	681.5396	0	7,283
Euclidean distance *(log)*	556.3	375.2458	0	1,959
Location were connected via the network in 1942	0.0653	0.2471	0	1
Locations were connected via the network in 725	0.0904	0.2868	0	1
Locations are connected via a Shinkansen station <10km	0.0536	0.2253	0	1
On different sides of the east-west border	0.4673	0.4989	0	1
On same island	0.7032	0.4793	0	1
Year of observation	2,010	3	2,007	2,013

Notes: The number of observations is $5,821,051$. The east-west border is defined as per Wrona (2018).

Table 3 — Descriptive statistics for municipality data, 1957-2014

	(1) mean	(2) sd	(3) min	(4) max
Employment (per km^2)	502.3	2,963	0.323	83,446
Population (per km^2)	948.4	2,306	1.509	30,639
Share of employment in intermediate sector	0.0459	0.0347	0.00268	0.500
Value of land *(in ě per m^2)*	10.64	0.850	7.928	15.14
Wage *(in ě per month)*	302,085	78,874	141,210	553,985
Average travel time by train *(in m)*	641.7	468.7	209.7	7,168
Average travel time by road *(in m)*	875.7	525.1	348.7	3,541
Shinkansen station	0.0699	0.255	0	1
Railway station <10km	0.814	0.389	0	1
Highway <10km	0.392	0.488	0	1
Shinkansen station <10km in 1972 plan	0.117	0.321	0	1
Shinkansen station <10km in 1942 plan	0.0331	0.179	0	1
Distance to coastline *(km)*	22.43	18.50	0.0183	85.45
Mean elevation *(m)*	284.9	259.3	0	1,139
Total area size *(km^2)*	220.3	250.9	1.50e-06	2,180
Year of observation	1,990	16.53	1,957	2,014

Notes: The number of observations is $18,238$. The data on the share of intermediate employment is since 1978, while the data on land values is available since 1981.

log population density and log employment density is 0.985. The share of employment in the intermediate services sector (as defined by employment in FIRE) is 4.6%. The correlation of log employment density with log land values and log wages is respectively 0.702 and 0.658. Hence, denser areas are more expensive but they offer higher wages. The correlation between log land values and log wages is 0.694.

The population-weighted average travel time by train is 10 hours, while the average travel time by road is just over 14.5 hours. Note that 81% of the municipalities are within 10km of a railway station, while just 7% of the municipalities are within 10km of a Shinkansen station. Most of Japan is close to the sea as the average distance to the coast is just over 22km.

In the theoretical framework we hypothesize that trade of intermediate services imply business trips to facilitate face-to-face interactions. Using prefecture-level data on travel flows, we show in Appendix A.6 that the elasticity of travel flows with respect to the value of intermediate services is about $0.7 - 0.8$ for train travel, while it is close to zero for travel by road. This confirms that for trade of intermediate services business trips are required for which the train is the obvious travel mode.

2.4 Reduced-form evidence

2.4.1 Econometric framework

Before we introduce our structural model, we aim to provide reduced-form evidence on the impact of Shinkansen stations on employment density, using municipality-level data between 1957 and 2014. Reduced-form effects of transport infrastructure are hard to identify because the assignment of HSR lines is not random and is related to the level of economic activities of a location. Although it is not the main aim of this paper to provide clear-cut reduced-form evidence, we still think it is insightful to sign the effects and to get an idea of the order of magnitude of the various effects. Let us denote $\overline{T}_{\mathcal{R},jy}$ to be the average travel time of the Japanese population by rail from location j to all other locations in year y. Further, let $M_{jy}/\mathcal{A}_j$ be the employment density in municipality j in year y. We then estimate:

$$\left\{\log \overline{T}_{\mathcal{R},jy}, \log \frac{M_{jy}}{\mathcal{A}_j}\right\} = \beta_0 + \beta_1 \mathcal{S}_{jy} + \beta_2 \mathcal{X}_{jy} + \lambda_j + \lambda_y + \epsilon_{jy}, \tag{1}$$

where β_0, β_1, β_2 are parameters to be estimated, λ_j are municipality fixed effects, λ_y are year fixed effects; $\mathcal{S}_{jy}$ is a dummy that equals one when the municipality (centroid) is within 10km of a Shinkansen station in year y, while $\mathcal{X}_{jy}$ are control variables.

A remaining concern is endogeneity. Although municipality fixed effects control for unobserved time-invariant characteristics of a location, new infrastructure may be correlated to unobserved trends. For example, infrastructure may be placed in areas where economic development is to be expected. Redding and Turner (2015) consider three approaches to mitigate endogeneity concerns: *(i)* apply an inconsequential unit approach; *(ii)* use planned route instrumental variables; or *(iii)* use historical route instrumental variables.

We choose a similar approach as in Koster *et al.* (2022), who use planned routes, but we do not apply the inconsequential unit approach here because we are interested in the overall effect of improved accessibility, rather than just on intermediate places. To construct our instruments, we consider two types of plans. The first is a plan that was approved by the cabinet around 1972, although similar versions of these plans were around for longer (Sargent, 1973). We refer to this plan as the '1972 plan' and define a dummy indicating whether a municipality has a stop in the corresponding plan. The second is the 1942 plan to link Tokyo to Beijing via a fast railway line. The idea was to increase the transport of passengers and goods between Japan, Korea, and China. Maps of the plans are reported in Appendix A.7. The validity of the instrument rests on the assumption that the station areas in the plans were not chosen because of expected employment growth of these places. We think this is a reasonable assumption as it was simply the goal to connect Japan's main cities. Further, the instrument addresses any issues with non-random timing.

2.4.2 Results

We report the results for accessibility in Panel A of Table 4, while effects on employment density are reported in Panel B of Table 4.

In column (1) we do not include municipality fixed effects and show that infrastructure is indeed assigned to the densest places. In Panel A a Shinkansen station area changes travel times by $(\exp(-0.449)-1)\times 100\% = -36\%$. Employment density is on average almost 10 times higher than in areas without a Shinkansen station. Municipality fixed effects address this endogeneity issue to a large extent. Column (2) shows that average travel times are 17.8% lower, so that a Shinkansen station strongly reduces travel times, while employment density increases by 11%. In column (3) we control for the proximity to other infrastructure: 'ordinary' railway stations and highways. This hardly changes the impact on average travel times, but the impact on employment density is somewhat lower: a Shinkansen station is now associated with an increase in employment density of 6%. In Appendix B.1 we report event studies to show that the effect on mean travel times adjusted a few years before a Shinkansen station is opened, presumably because other railway connections were improved before the opening of the new Shinkansen station. For employment density we do not find pre-trends. After the opening of a station, the employment density increases by about 5% but the effect increases to about 25% after 40 years.

Columns (4) and (5) in Table 4 focus on the IV estimates. With the 1972 plan, the Kleibergen-Paap F-statistic is 540, while for the 1942 plan, the value is almost 100, which is considerably higher than the rule-of-thumb value of 10. In Appendix B.2 we report the full first-stage results. The second-stage results for the impact of the Shinkansen on average travel times are similar to what we showed earlier, with coefficients ranging from -0.347 to -0.257. In line with previous studies, the IV approach displays stronger effects of being connected on employment density. The opening of a Shinkansen station is now associated with an increase in employment density of 40%,

Table 4 — Reduced-form estimates: accessibility and employment density

Panel A: Accessibility	*Dependent variable: the log of average railway travel time*				
	No spatial fixed effects	+ Municipality fixed effects	+ Infrastructure controls	Instrument: 1972 plan×year	Instrument: 1942 plan×year
	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS
Shinkansen station <10km	-0.4485***	-0.1949***	-0.1778***	-0.2566***	-0.3473***
	(0.0067)	(0.0058)	(0.0057)	(0.0171)	(0.0298)
Railway station <10km			-0.1011***	-0.0996***	-0.0978***
			(0.0076)	(0.0076)	(0.0076)
Highway <10km			-0.0511***	-0.0467***	-0.0417***
			(0.0031)	(0.0031)	(0.0034)
Municipality fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	18,678	18,678	18,678	18,678	18,678
R^2	0.3724	0.9754	0.9765		
Kleibergen-Paap F-statistic				539.8	97.36

Panel B: Employment density	*Dependent variable: the log of employment density*				
	No spatial fixed effects	+ Municipality fixed effects	+ Infrastructure controls	Instrument: 1971 plan×year	Instrument: 1942 plan×year
	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS
Shinkansen station <10km	2.3321***	0.1052***	0.0585**	0.3391***	0.4067***
	(0.0525)	(0.0231)	(0.0232)	(0.0580)	(0.1286)
Railway station <10km			0.0509***	0.0455***	0.0442***
			(0.0158)	(0.0158)	(0.0160)
Highway <10km			0.1667***	0.1512***	0.1474***
			(0.0104)	(0.0109)	(0.0122)
Municipality fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	18,678	18,678	18,678	18,678	18,678
R^2	0.1323	0.9758	0.9764		
Kleibergen-Paap F-statistic				539.8	97.36

Notes: Accessibility is the average travel time by train to the population in mainland Japan. In column (4) we instrument Shinkansen <10km with an interaction term of whether the municipality is within 10km of a planned Shinkansen station with the year of observation. In column (5) we use a dummy indicating that a municipality is within 10km of the planned line in 1942 interacted with the year of observation. Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

which is sizable. In column (5) the impact is even somewhat stronger (*i.e.* 50%).

Hence, we find consistent evidence for, on average, strong positive effects. Still, one may argue that the impact of a Shinkansen station is likely to be heterogeneous across locations. For example, if the station is opened in an area that is already very well connected, effects may be smaller than, say, in a peripheral area that was poorly connected. We therefore repeat the analysis but allow for heterogeneity in the effect. In Appendix B.3 we show that effects are considerably smaller and even may turn negative for dense areas. Indeed, being connected to the Shinkansen reduces travel times, thus allowing firms to relocate activities in areas where land is cheaper. This result strongly suggests that a model studying the impacts of the Shinkansen network should take into account that the effect of a connection is heterogeneous and may depend on *e.g.* the place in the network, initial amenities, productivities, and employment density.[12]

3 The model

3.1 The economy

Consider an economy with a mass M of mobile workers, a finite location space $i = 1, ..., I$ with $I \geq 2$, a homogeneous final good, and a continuum of differentiated intermediate goods. There are two sectors $s = 1, 2$.

The first one produces the consumption good under perfect competition and constant returns, using land, labor, and intermediate (business-to-business) services. The second sector produces intermediate services under increasing returns and monopolistic competition, using land and labor. Workers are perfectly mobile between the two sectors. As the transport costs of goods within developed countries have tremendously decreased, the final good is costlessly tradable; this good is chosen as the numéraire. By contrast, commuting between any two different locations remains costly. Likewise, the provision of intermediate services at a distant location often involves the movement of workers whose travel costs are also fairly high. Hence, shipping intermediate services to the final sector involves a positive cost. There are two transport modes, *i.e.*, roads and highways ($m = 1$ or $\mathcal{H}$) and railways ($m = 2$ or $\mathcal{R}$) that can be used to travel between any location pair. Finally, each location i is endowed with a quantity $L_i > 0$ of land. Land is owned by immobile landlords who use their incomes to consume the final good only.

3.2 Workers

Workers choose simultaneously a residence $i = 1, ..., I$, a workplace $j = 1, ..., I$, and a transport mode $m = 1, 2$, that is, a triple ijm, as well as their housing and final good consumption. Each residential location i is endowed with amenities $A_i > 0$ and each workplace j has amenities

[12] We do not think that the finding of an average positive effect contradicts with the finding of a negative effect in Koster *et al.* (2022). This paper focuses on the so-called 'intermediate' places and identifies the effects within close distance of a planned line. In other words, in that paper peri-urban areas that receive a station are compared to similar areas that do not receive a station. These intermediate places are more often denser places, for which we also find negative effects. In Appendix B.4 we show that we can replicate the findings of Koster *et al.* (2022) if we apply their approach.

$B_j > 0$.

Each worker $\omega \in [0, M]$ is characterized by her type, which is defined by the vector of match values with the triple ijm, *i.e.*, $\mathbf{Z}(\omega) \equiv (z_{ijm}(\omega)) \in \mathbb{R}_+^{I \times I \times 2}$. The distribution of types $\mathbf{Z}(\omega)$ is the product measure of $4 \times I^2$ identical Fréchet distributions, that is,

$$G(\mathbf{Z}) = M \exp\left(- \sum_{i=1}^{I} \sum_{j=1}^{I} \sum_{m=1}^{2} z_{ijm}^{-\varepsilon} \right), \tag{2}$$

where $\varepsilon > 0$ accounts for the dispersion of idiosyncratic tastes.

Amenities and transport modes have the nature of horizontally differentiated goods, implying that workers are heterogeneous in the quality of their match with a triple ijm. An ω-worker who lives at i, works at j, and uses mode m has a utility given by

$$U_{ijm}(\omega) = \frac{1}{\alpha^\alpha (1-\alpha)^{1-\alpha}} z_{ijm}(\omega) A_i B_j h_i^\alpha y_i^{1-\alpha},$$

where h_i is the amount of land used for housing and y_i the quantity of the final good consumed in location i. Commuting involves an iceberg cost $t_{ijm} > 1$ between i and j using mode m, which is the same for all $\omega \in [0, M]$.

Workers in location i work in the final or in the intermediate sector. Since workers are perfectly mobile between sectors, the wages paid in the two sectors are the same in equilibrium. Free entry implies that intermediate firms make zero profits. Since the final sector operates under constant returns and perfect competition, equilibrium profits are also equal to zero. As the aggregate land rent goes to absentee landlords, a worker's income is equal to her wage. The budget constraint of a worker who has chosen the triple ijm is then given by

$$\frac{w_j}{t_{ijm}} = R_i h_i + y_i,$$

where w_j is the wage paid in location j and R_i the land rent at i.

Maximizing utility subject to the budget constraint yields the demand for housing at i given by

$$h_i = \frac{\alpha w_j}{t_{ijm} R_i},$$

so that the indirect utility of an ω-worker is as follows:

$$V_{ijm}(\omega) = z_{ijm}(\omega) A_i B_j \frac{w_j}{t_{ijm} R_i^\alpha}. \tag{3}$$

Workers make mutually exclusive choices among a finite number of indivisible alternatives, *i.e.*, they choose a triple ijm. Then, using (2) and (3), the share n_{ijm} of workers who choose the triple ijm equals

$$n_{ijm} = \frac{(A_i B_j w_j / t_{ijm} R_i^{\alpha})^{\varepsilon}}{\sum_{k=1}^{I} \sum_{l=1}^{I} \sum_{m=1}^{2} (A_k B_l w_l / t_{klm} R_k^{\alpha})^{\varepsilon}}, \tag{4}$$

where the equality stems from the Fréchet distribution assumption. Other things being equal, a location i endowed with more amenities host more residents regardless of their workplaces. Likewise, a workplace j with higher amenities attract more workers regardless of their residences. By contrast, fewer workers choose a triple ijm when commuting cost t_{ijm} are higher.

Let M_i (N_i) be the mass of workers (residents) in i. Since workers are free to choose where to live through the choice of the pair ij, the population N_i in location i is endogenous. Furthermore, since workers commute, the population N_i generally differs from the volume of employment M_i in i, which is also endogenous.

Conditional on the residential location i, the share $n_{ij|i}$ of workers who take a job at j is given by the commuting equation:

$$n_{ij|i} \equiv \frac{\sum_{m=1}^{2} (B_j w_j / t_{ijm})^{\varepsilon}}{\sum_{k=1}^{I} \sum_{m=1}^{2} (B_k w_k / t_{ikm})^{\varepsilon}}. \tag{5}$$

In other words, the share $n_{ij|j}$ depends on the wage w_j and amenities B_j at location j, on the transport mode and commuting cost between i and j (bilateral resistance), as well as the wages, amenities and commuting costs to all workplaces (multilateral resistance). Using (5), employment at location j is the sum across all locations i of the i-residents who commute to j:

$$M_j = M_{j1} + M_{j2} = \sum_{i=1}^{I} \frac{\sum_{m=1}^{2} (B_j w_j / t_{ijm})^{\varepsilon}}{\sum_{k=1}^{I} \sum_{m=1}^{2} (B_k w_k / t_{ikm})^{\varepsilon}} N_i, \tag{6}$$

where M_{js} is the employment at location j in sector s.

Likewise, conditional on the workplace in j, the share $\bar{n}_{ij|j}$ of workers who live in i is given by the following equation:

$$\bar{n}_{ij|j} \equiv \frac{\sum_{m=1}^{2} (A_i / (t_{ijm} R_i^{\alpha}))^{\varepsilon}}{\sum_{k=1}^{I} \sum_{m=1}^{2} (A_k / (t_{kjm} R_k^{\alpha}))^{\varepsilon}}. \tag{7}$$

In other words, the share $\bar{n}_{ij|j}$ depends on the land rent R_i and amenities A_i at location i, on the transport mode and commuting cost between i and j (bilateral resistance). Using (7), the population N_i in i is equal to the sum across all locations j of the i-workers who commute to j:

$$N_i = \sum_{j=1}^{I} \bar{n}_{ij|j} M_j = \sum_{j=1}^{I} \frac{\sum_{m=1}^{2} (A_i / (t_{ijm} R_i^{\alpha}))^{\varepsilon}}{\sum_{k=1}^{I} \sum_{m=1}^{2} (A_k / (t_{kjm} R_k^{\alpha}))^{\varepsilon}} M_j. \tag{8}$$

The gravity equations (6) and (8) describe the residential and workplace choices made by workers through commuting and migration flows across locations.

The demand for land stemming from workers located at i is given by

$$L_{ic} = \sum_{j=1}^{I} \sum_{m=1}^{2} \frac{\alpha w_j}{t_{ijm} R_i} \bar{n}_{ij|j} M_j. \tag{9}$$

National labor market clearing implies that the total mass of residents is equal to the total mass of workers:

$$\sum_{j=1}^{I} M_j = \sum_{i=1}^{I} N_i = M. \tag{10}$$

Summing (6) across j and (8) across i show that (10) always holds.

Finally, because the consumption good is the numeraire, we follow the literature and define gross welfare as the sum of individual indirect utilities:

$$W = \Gamma \left(\frac{\varepsilon - 1}{\varepsilon} \right)^{\varepsilon} \left[\sum_{i=1}^{I} \sum_{j=1}^{I} \sum_{m=1}^{2} \left(A_i B_j w_j / \left(t_{ijm} R_i^{\alpha} \right) \right)^{\varepsilon} \right]^{\frac{1}{\varepsilon}},$$

where $\Gamma(\cdot)$ is the gamma function.

3.3 Production

The final sector. The production technology of the consumption good is the same across locations and the output in location i given by

$$Y_i = E_{i1} M_{i1}^{\beta} L_{i1}^{\gamma} Q_i^{1-\beta-\gamma},$$

where M_{i1} and L_{i1} are, respectively, the amount of labor and land consumed by this sector at location i and

$$Q_i \equiv \left[\sum_{j=1}^{I} \sum_{m=1}^{2} \int_{\Omega_j} \left(x_{jim}(\nu) \right)^{\frac{\sigma-1}{\sigma}} \mathrm{d}\nu \right]^{\frac{\sigma}{\sigma-1}} \tag{11}$$

is the CES bundle of services supplied by the intermediate sector while $\beta, \gamma \in (0,1)$ and $\beta + \gamma < 1$. In the bracketed term, $x_{ji}(\nu)$ is a continuous measure of intermediate services ν produced in j and used in i, while $\sigma > 1$ the elasticity of substitution between any two intermediate services. In (11), Ω_j is the set of intermediate services produced in location j. Denote by X_{ji} the final sector's demand in location i for any intermediate service delivered from location j.

Profits in the final sector are given by

$$\Pi_i = Y_i - w_i M_{i1} - R_i L_{i1} - \sum_{j=1}^{I} \sum_{m=1}^{2} \int_{\Omega_j} p_{jim}(\nu) x_{jim}(\nu) \, \mathrm{d}\nu,$$

where $p_{jim}(v)$ is the price paid in location i for the intermediate v delivered from location j when using the transport mode m and the wage w_i paid by the final sector in i.

Differentiating Π_i with respect to M_{i1}, L_{i1}, and $x_{jim}(\nu)$ leads to the equilibrium conditions

$$\gamma E_{i1} M_{i1}^{\beta} L_{i1}^{\gamma-1} Q_i^{1-\beta-\gamma} - R_i = 0, \tag{12}$$

$$\beta E_{i1} M_{i1}^{\beta-1} L_{i1}^{\gamma} Q_i^{1-\beta-\gamma} - w_i = 0, \tag{13}$$

and the inverse demand for intermediate v

$$p_{jim}(v) = (1-\beta-\gamma)\, E_{i1} M_{i1}^{\beta} L_{i1}^{\gamma} Q_i^{\frac{1-\beta\sigma-\gamma\sigma}{\sigma}} (x_{jim}(\nu))^{-\frac{1}{\sigma}},$$

which yields the demand $X_{ji} = X_{ji1} + X_{ji2}$ for the service ν supplied by the intermediate sector. Thus,

$$X_{jim} = (1-\beta-\gamma)^{\sigma} E_{i1}^{\sigma} M_{i1}^{\beta\sigma} L_{i1}^{\gamma\sigma} Q_i^{1-\beta\sigma-\gamma\sigma} p_{jim}^{-\sigma}, \tag{14}$$

where we drop the label ν hereafter due to symmetry.

It follows from (12) and (13) that

$$L_{i1} = \frac{\gamma M_{i1} w_i}{\beta R_i}, \tag{15}$$

while (14) implies

$$Q_i = \left(\frac{E_{i1}(1-\beta-\gamma)}{P_i}\right)^{\frac{1}{\beta+\gamma}} M_{i1}^{\frac{\beta}{\beta+\gamma}} L_{i1}^{\frac{\gamma}{\beta+\gamma}}, \tag{16}$$

where

$$P_i \equiv \left[\sum_{j=1}^{I} \sum_{m=1}^{2} \int_{\Omega_i} p_{jim}(v)^{1-\sigma}\, \mathrm{d}\nu\right]^{\frac{1}{1-\sigma}}$$

is the price index at i.

Combining (12) and (16) and plugging (15) into the resulting expression yields:

$$R_i = \gamma \left(\frac{\beta}{w_i}\right)^{\frac{\beta}{\gamma}} E_{i1}^{\frac{1}{\gamma}} \left(\frac{1-\beta-\gamma}{P_i}\right)^{\frac{1-\beta-\gamma}{\gamma}}. \tag{17}$$

Solving this expression for E_{i1} yields

$$E_{i1} = \left(\frac{P_i}{1-\beta-\gamma}\right)^{1-\beta-\gamma} \left(\frac{w_i}{\beta}\right)^{\beta} \left(\frac{R_i}{\gamma}\right)^{\gamma}, \tag{18}$$

which ensures that the marginal cost of the final sector in location i is constant.

The intermediate sector. An intermediate service is provided by a single firm and each firm supplies a single service. An intermediate firm can choose a transport mode $m = 1, 2$ to ship its intermediate service to the final sector. Since locations have different relative positions in the two transport networks, the shipping costs are specific to the mode m and the origin-destination pair ij. More specifically, traveling from j to i by using the mode m involves the iceberg cost $\tau_{jim} > 1$.

An intermediate firm in j requires a fixed number $\ell_j > 0$ of units of land to operate, and uses $1/E_{j2} > 0$ units of labor to produce one unit of service. Both E_{j2} and ℓ_j are location-specific. An intermediate firm in j maximizes its profit given by the following expression:

$$\pi_j = \sum_{i=1}^{I} \sum_{m=1}^{2} \left(p_{jim} - \tau_{jim} w_j / E_{j2} \right) X_{jim} - R_j \ell_j .$$

Profit maximization yields the equilibrium price:

$$p_{jim} = \frac{\sigma}{\sigma - 1} \frac{\tau_{jim} w_j}{E_{j2}}, \tag{19}$$

which implies that the demand for intermediate services (14) is small when the trade cost τ_{jim} is high for mode m.

Free entry and (19) imply that the equilibrium output of a firm at j is given by

$$q_j \equiv \sum_{i=1}^{I} \sum_{m=1}^{2} \tau_{jim} X_{jim} = \frac{(\sigma - 1) E_{j2} R_j \ell_j}{w_j}, \tag{20}$$

which decreases with the marginal cost, w_j/E_{j2}, and increases with the fixed cost, $R_j \ell_j$, like in the CES model monopolistic competition where labor is the only production factor. Consequently, an intermediate firm hires m_j workers:

$$m_j = q_j / E_{j2} = \frac{(\sigma - 1) R_j \ell_j}{w_j}, \tag{21}$$

which increases with the local land rent and decreases with the local wage. The mass K_j of intermediate firms at j is then obtained from the labor market clearing condition at j:

$$K_j = \frac{M_{j2}}{m_j}, \tag{22}$$

where the mass K_j of firms increases with the mass of workers employed in the intermediate sector.

Using (21) and (22), the total land demand from intermediate firms at i is thus given by

$$L_{i2} = K_i \ell_i = \frac{M_{i2}}{m_j} \ell_i = \frac{w_i M_{i2}}{(\sigma - 1) R_i}. \tag{23}$$

The price index may then be rewritten as follows:

$$P_i = \left[\sum_{j=1}^{I} \sum_{m=1}^{2} K_i \left(\frac{\sigma}{\sigma - 1} \frac{\tau_{jim} w_j}{E_{j2}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{24}$$

The zero-profit condition in the intermediate sector is as follows:

$$\sum_{i=1}^{I} \sum_{m=1}^{2} (p_{jim} - \tau_{jim} w_j / E_{j2}) X_{jim} = R_j \ell_j. \tag{25}$$

Last, using (19) and (8), it is readily verified that the expenditure share S_{ji} of the final sector in i on all intermediate services provided in j is given by the following gravity equation:

$$S_{ji} \equiv \frac{\sum_{m=1}^{2} K_j p_{jim} X_{jim}}{\sum_{k=1}^{I} \sum_{m=1}^{2} K_k p_{kim} X_{kim}} = \frac{\sum_{m=1}^{2} (w_j M_{j2} / R_j L_{jc}) (w_j \tau_{jim} / E_{j2})^{1-\sigma}}{\sum_{k=1}^{I} \sum_{m=1}^{2} (w_k M_{k2} / R_k L_{kc}) (\tau_{ikm} w_k / E_{k2})^{1-\sigma}}. \tag{26}$$

3.4 The spatial equilibrium

Equilibrium conditions. Since the land market clearing condition at i is given by $L_{ic} + L_{i1} + L_{i2} = L_i$, (9), (15) and (23) lead to the following condition:

$$L_i = \sum_{j=1}^{I} \sum_{m=1}^{2} \frac{\alpha w_j}{t_{ijm} R_i} \frac{(A_i / (t_{ijm} R_i^{\alpha}))^{\varepsilon} (M_{j1} + M_{j2})}{\sum_{r=1}^{I} \sum_{m=1}^{2} (A_r / (t_{rjm} R_r^{\alpha}))^{\varepsilon}} + \frac{\gamma M_{i1} w_i}{\beta R_i} + \frac{w_i M_{i2}}{(\sigma - 1) R_i}. \tag{27}$$

Substituting P_i in (24) into (17) and (25), we obtain:

$$\frac{\beta}{w_i} = E_{i1}^{-\frac{1}{\beta}} \left\{ \frac{1}{1 - \beta - \gamma} \left[\frac{w_i M_{i2}}{(\sigma - 1) \ell_i R_i} \sum_{j=1}^{I} \sum_{m=1}^{2} \left(\frac{\sigma}{\sigma - 1} \frac{\tau_{jim} w_j}{E_{j2}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \right\}^{\frac{1-\beta-\gamma}{\beta}} \left(\frac{R_i}{\gamma} \right)^{\frac{\gamma}{\beta}} \tag{28}$$

and

$$\begin{aligned} R_j \ell_j \quad = \quad & \frac{[E_{j1} (1 - \beta - \gamma)]^{\frac{1}{\beta+\gamma}}}{\sigma} \left(\frac{\sigma}{\sigma - 1} \frac{w_j}{E_{j2}} \right)^{1-\sigma} \\ & \times \sum_{i=1}^{I} \sum_{m=1}^{2} \tau_{jim}^{1-\sigma} M_{i1}^{\frac{\beta}{\beta+\gamma}} L_{i1}^{\frac{\gamma}{\beta+\gamma}} \left[\frac{w_i M_{i2}}{(\sigma - 1) \ell_i R_i} \sum_{j=1}^{I} \sum_{m=1}^{2} \left(\frac{\sigma}{\sigma - 1} \frac{\tau_{jim} w_j}{E_{j2}} \right)^{1-\sigma} \right]^{-\frac{1}{1-\sigma} \frac{1-\beta\sigma-\gamma\sigma}{\beta+\gamma}} \end{aligned} \tag{29}$$

The workers living in i, working in j and using mode m to commute spend a share $1 - \alpha$ of their income w_j/t_{ijm} on the final good. It then follows from (8) that their total expenditure on this good is equal to

$$(1-\alpha)\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{m=1}^{2}\frac{w_j}{t_{ijm}}\frac{\left(A_i/\left(t_{ijm}R_i^{\alpha}\right)\right)^{\varepsilon}}{\sum_{r=1}^{I}\sum_{m=1}^{2}(A_r/\left(t_{rjm}R_r^{\alpha}\right))^{\varepsilon}}M_j,$$

while landlords spend their entire income ($\Sigma_i R_i L_i$) on the final good.

The Walras Law implies that the global market for the final good clears:

$$(1-\alpha)\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{m=1}^{2}\frac{w_j}{t_{ijm}}\frac{\left(A_i/\left(t_{ijm}R_i^{\alpha}\right)\right)^{\varepsilon}}{\sum_{r=1}^{I}\sum_{m=1}^{2}(A_r/\left(t_{rjm}R_r^{\alpha}\right))^{\varepsilon}}M_j+\sum_{i=1}^{I}R_iL_i=\sum_{i=1}^{I}Y_i.$$

Interior spatial equilibrium. The *spatial equilibrium* is given by the equilibrium wages at each location, $(w_1^*, ..., w_I^*) \in \mathbb{R}_+^I$, the equilibrium land rent at each location, $(R_1^*, ..., R_I^*) \in \mathbb{R}_+^I$, the equilibrium population at each location, $(N_1^*, ..., N_I^*) \in \mathbb{R}_+^I$ with $\Sigma_i N_i^* = N$, the equilibrium employment level at each location and each sector, $(M_{1s}^*, ..., M_{Is}^*) \in \mathbb{R}_+^I$ with $\Sigma_s \Sigma_i M_{is}^* = M$, which solve the $5I$ equations (6), (8), (27), (28), and (29). This solution is such that workers maximize their utilities under their budget constraints, intermediate firms maximize their own profits, the final sector maximizes profit, and markets clear at non-negative and finite prices.

The existence of an equilibrium follows from Lemmas S1 and S2 in Ahlfeldt *et al.* (2015). Assume that there is a location i such that $R_i^* = 0$. Since there exists a location j such that $w_j^* > 0$, the Inada conditions imply that workers who choose ijm have an infinite utility because the land rent is zero. Since the support of the Fréchet distribution is unbounded from above, we have $N_i^* > 0$. If $w_j^* > 0$, location j hosts a positive share of residents, as well as a positive share of the final and intermediate sectors because L_{ic} is positive from (9), L_{i1} is positive from (15), and L_{i2} is positive from (23).

It remains to consider M_{j1}^* and M_{j2}^*. If $M_{j2}^* = 0$, then the intermediate service is not produced at j ($q_j^* = 0$). Since $R_j^* > 0$, it then follows from (20) that w_j^* goes to infinity, a contradiction. Since $M_{j2}^* > 0$, P_j^* is positive and finite by (24) while w_j^* is positive and finite from (17). Data are such that M_{j1} is positive for all j. Since w_j^*, R_j^* and P_j^* are positive and finite while (18) implies that the marginal production cost of the final sector is equal to the price of the consumption good ($p = 1$) at all locations j, we consider the case where $M_{j1}^* > 0$. In other words, *we focus on interior equilibria.*

Showing uniqueness turns out to be a hard task. For example, we cannot apply the approach followed by Ahlfeldt *et al.* (2015), Allen *et al.* (2015) and Monte *et al.* (2018). Given that both the final and intermediate sectors, as well as consumers, use a variable amount of land, we end up with a land market clearing condition (27) which is not amenable to a multiplicative equation. Furthermore, it is well known that new economic geography models with several locations and land consumption are subject to multiplicity of equilibria (Takayama *et al.*, 2020). Given the main purpose of our paper, we have chosen not to embark in an in-depth research to prove uniqueness. Instead, we have solved a large number of cases for plausible values of the parameters. In the next sections, we consider only interior equilibria and check whether the so-obtained equilibrium is unique.

3.5 A model extension – agglomeration economies

Quantitative spatial models typically allow for agglomeration economies in the services sector (Ahlfeldt *et al.*, 2015). In the baseline version of our model, we allow for agglomeration economies through input-output linkages between final and intermediate firms, which makes it attractive for firms to locate close together (Krugman and Venables, 1995; Ellison *et al.*, 2010). Still, one may argue that we do not allow for agglomeration economies among intermediate services firms. Hence, in an extension of the structural model, we will allow the TFP of intermediate firms to depend on employment density of intermediate firms located in j. More specifically, we will assume that

$$E_{j2} \equiv e_{2j} \left(\frac{M_{j2}}{L_i} \right)^{\lambda}. \tag{30}$$

It is well documented that workers' productivity increases with the intermediate employment density M_{j2}/L_j. The literature on agglomeration economies suggests that λ ranges between $0.02 - 0.07$ (Rosenthal and Strange, 2004; Combes and Gobillon, 2015). The constant e_{2j} stands for location-specific fundamentals that affect intermediate productivity in j.

4 Structural estimation

4.1 Estimation and identification

To start, we determine two parameters based on Japanese data, that is, the share of income spent on land α and the labor share β in the final sector. Fortunately, there exist clear priors on what the values for α and β are. In line with data from Japan's *Statistics Bureau* on the expenditure share of housing in Japan in 2014, we choose $\alpha = 0.23$. Based on the *Economic Census for Business Activity* in 2016, we set $\beta = 0.5$.

The estimation of the remaining parameters consists of five steps. First, using commuting data, we recover commuting travel time elasticities for trains and cars ($\varkappa_{\mathcal{R}}$ and $\varkappa_{\mathcal{H}}$). Second, using data on production networks, we estimate a gravity equation to obtain the mode-specific travel time elasticities ($\vartheta_{\mathcal{R}}$ and $\vartheta_{\mathcal{H}}$) for intermediate services. Third, using data on wages and a

Bartik-instrument, we identify the degree of worker heterogeneity (ε). Fourth, using data on land use and the land market equilibrium condition (27), we estimate the elasticity of substitution (σ) and the land share in the final sector on land (γ). Finally, given the number of intermediate and final firms and the estimated parameters, we recover the corresponding TFPs (E_{j1} and E_{j2}).

4.1.1 Step 1: A gravity model for commuting trips

Unfortunately, there is no data on commuting flows by travel mode. Hence, we use aggregate commuting flows by municipality pairs. The commuting gravity equation (4) leads to the first equation we estimate:

$$\log n_{ijy} = \nu_0 + \nu_{iy} + \nu_{jy} + \log(t_{\mathcal{R},ijy} + t_{\mathcal{H},ijy}) + \epsilon_{ijy}, \tag{31}$$

where $t_{\mathcal{R},ijy} \equiv e^{\varkappa_{\mathcal{R}} \widetilde{T}_{\mathcal{R},ijy}}$ and $t_{\mathcal{H},ijy} \equiv e^{\varkappa_{\mathcal{H}} \widetilde{T}_{\mathcal{H},ijy}}$. Recall that the share of commuters using the Shinkansen is just 0.05% and that the share of commuters using highways is 2.5%, so $\widetilde{T}_{\mathcal{R},ijy}$ and $\widetilde{T}_{\mathcal{H},ijy}$ are travel times by train and car in which *HSRs and highways are excluded*. The variables ν_{iy} and ν_{jy} are residence-by-year and workplace-by-year fixed effects capturing $(A_{iy}/R_{iy})^{\varepsilon}$ and $(B_{jy}w_{jy})^{\varepsilon}$; $\varkappa_{\mathcal{R}} \equiv -\kappa_{\mathcal{R}}\varepsilon$ and $\varkappa_{\mathcal{H}} \equiv -\kappa_{\mathcal{H}}\varepsilon$ are parameters to be estimated; and ν_0 is a normalization constant that equals 1 in the theoretical model. Following Larch *et al.* (2019), we can sum up the total number of commuters between i and j in year y and estimate the above specification by Poisson Pseudo-Maximum Likelihood (PPML).

Because we have two transport modes, the parameters of interest $\varkappa_{\mathcal{R}}$ and $\varkappa_{\mathcal{H}}$ enter non-linearly in (31). We therefore estimate (31) in two steps. First, we discretize travel times and include fixed effects for each combination of $\widetilde{T}_{\mathcal{R},ijy}$ and $\widetilde{T}_{\mathcal{H},ijy}$, as well as fixed effects ν_{iy} and ν_{jy}. We then regress the estimated travel-time fixed effects on $\log(t_{\mathcal{R},ijy} + t_{\mathcal{H},ijy})$ using non-linear least squares.[13]

In Appendix C.1 we report the detailed results of the estimation of (31) and consider various robustness checks. For example, one may be concerned about reverse causality, which would imply that our estimates would be biased downwards (*i.e.*, they would be more negative). Reverse causality would be present if areas with large commuting flows receive more infrastructure to alleviate traffic congestion. One way to address this concern is to instrument travel time by Euclidean distance. This usually does not lead to significantly different coefficients, which suggests that this is not a main issue (Dericks and Koster, 2021, Koster, 2023). Since here we consider

[13] We first estimate $\exp(n_{ijy}) = \exp(\nu_{\tilde{T}_{\mathcal{R}},\tilde{T}_{\mathcal{H}}} + \nu_{iy} + \nu_{jy})$, where $\nu_{\tilde{T}_{\mathcal{R}},\tilde{T}_{\mathcal{H}}}$ denotes a fixed effect for each combination of travel time by rail and road. In the second step, we regress $\nu_{\tilde{T}_{\mathcal{R}},\tilde{T}_{\mathcal{H}}} = \nu_0 + \log\left(e^{\varkappa_{\mathcal{R}}\widetilde{T}_{\mathcal{R},ijy}} + e^{\varkappa_{\mathcal{H}}\widetilde{T}_{\mathcal{H},ijy}}\right) + \epsilon_{iy}$, so that we obtain $\hat{\varkappa}_{\mathcal{R}}$ and $\hat{\varkappa}_{\mathcal{H}}$. In the second step, we weight by the number of location pairs that share the same combination of travel time by rail and road.

travel times by train and by car, we cannot instrument both variables by the Euclidean distance. Instead, we estimate models where we only keep location pairs that were already connected *(i)* in the plans for railway and roads infrastructure laid out during World War II, and *(ii)* by roads in 725 when Nara was the capital of Japan. By only including locations that were already directly connected in early times, we somehow alleviate the concern that two places are connected because of a high flow (Faber, 2014; Banerjee *et al.*, 2020).[14]

One may be concerned that our values for $\varkappa_{\mathcal{R}}$ and $\varkappa_{\mathcal{H}}$ are dependent on the transport infrastructure in Japan. For example, when road travel is much quicker than the train, this may affect $\varkappa_{\mathcal{R}}$ and $\varkappa_{\mathcal{H}}$ locally. We will provide some suggestive evidence that our main results are not significantly affected when the commuting time by train is substantially higher than the commuting time by road on certain links. In other words, workers who commute on links for which the train is the less attractive transport mode have similar travel time elasticities.

4.1.2 Step 2: A gravity model for trade of intermediate services

There is no data on trade flows by travel mode. Hence, we use aggregate trade flows by municipality pairs. We further estimate a gravity equation for trade between final and intermediate firms using *TSR* and input-output data. To this end, we use (26) to estimate the following equation:

$$\log S_{ji} = \upsilon_0 + \upsilon_{iy} + \upsilon_{jy} + \log(\tau_{\mathcal{R},ijy} + \tau_{\mathcal{H},ijy}) + \epsilon_{ijy}, \tag{32}$$

where S_{ji} denotes the expenditure share of the final sector in i on intermediate services produced in j. Hence, we use the normalized trade value between the final and intermediate sector for a given location pair, which follows directly from our model.

Set $\tau_{\mathcal{R},jiy} \equiv e^{\vartheta_{\mathcal{R}} T_{\mathcal{R},jiy}}$ and $\tau_{\mathcal{H},jiy} \equiv e^{\vartheta_{\mathcal{H}} T_{\mathcal{H},jiy}}$, where $\vartheta_{\mathcal{R}} \equiv (1-\sigma)\theta_{\mathcal{R}}$ and $\vartheta_{\mathcal{H}} \equiv (1-\sigma)\,\theta_{\mathcal{H}}$. The variables υ_{iy} and υ_{jy} are buyer-by-year and seller-by-year location fixed effects, respectively, which absorb the region-specific price index and scale parameters; υ_0 is a normalization constant, which equals 1 in the theoretical model. To address the issue of zero flows, we estimate the above specification by PPML. We then sum up the values of trade flows of intermediate services by location pair. Similar to the commuting gravity equation, the parameters $\vartheta_{\mathcal{R}}$ and $\vartheta_{\mathcal{H}}$ enter non-linearly because we have multiple transport modes. We therefore estimate (32) in two steps. First, we discretize travel times and include fixed effects for each combination of $T_{\mathcal{R},jiy}$ and $T_{\mathcal{H},jiy}$, as well as fixed effects ν_{iy} and ν_{jy}. We then regress the estimated fixed effects on $\log(\tau_{\mathcal{R},ijy} + \tau_{\mathcal{H},ijy})$ using non-linear least squares.

Appendix C.2 reports results of the estimation of (32) and investigates several extensions.

[14] Note that the frequency of trains between two locations may be increased when the passenger flow increases. However, frequency is not part of our measure of travel time.

First, we address omitted variable bias. The main omitted variable in (26) relates to other transport modes, *i.e.*, the possibility that workers travel by airplane. We therefore also estimate regressions where we include air ($\mathcal{A}$) travel, so we estimate an additional parameter $\vartheta_{\mathcal{A}}$. Importantly, we find that $\vartheta_{\mathcal{R}}$ and $\vartheta_{\mathcal{H}}$ are hardly affected by the inclusion of air travel. Moreover, $\vartheta_{\mathcal{A}} \approx \vartheta_{\mathcal{H}}$ and we can show that excluding air travel in what follows will not materially affect the results. Other omitted variables may relate to trade and cultural barriers that are hard to quantify but are correlated with travel times. However, because we focus on one country which is culturally homogeneous and which has one main language, we may expect cultural barriers to be very low. Likewise, there are few trade barriers within Japan. Still, we show that the travel time elasticities $\vartheta_{\mathcal{R}}$ and $\vartheta_{\mathcal{H}}$ are insensitive to the inclusion of a dummy variable indicating whether a location pair is on the east and west of Japan. This variable is based on Wrona (2018) who showed that trade between East and West Japan is considerably lower than what bilateral trade costs would imply. We also include a dummy indicating whether a location pair is on the same island, which does not affect the results.

Second, as in the commuting gravity models, reverse causality may be an issue when some final firms that are very large and have many input suppliers lobby for better infrastructure. We address reverse causality using the same approach as in Step 1 by keeping connected locations *(i)* in the high-speed railway plan designed in 1942 to link Tokyo to Beijing and the National Highway Plan designed in 1943, *(ii)* by roads in 725, or *(iii)* by the Shinkansen network.

Third, one may be worried that our results are driven by a few large firms that have many input suppliers. Indeed, Bernard *et al.* (2019) show that the distribution of in-degree links by firm is highly skewed. Our baseline regressions therefore only include single-plant firms, but we show in Appendix C.2 that including multi-plant firms hardly matters for the results.

Finally, one may once again be concerned that $\vartheta_{\mathcal{R}}$ depends on the overall provision of railways in Japan. We will therefore check whether $\vartheta_{\mathcal{R}}$ is not substantially different on links where travel time by train is relatively higher than travel time by car.

4.1.3 Step 3: Identifying workers' heterogeneity

Given estimates for $\varkappa_{\mathcal{R}}$ and $\varkappa_{\mathcal{H}}$, and using the commuting gravity equation (4), this allows us to recover the so-called 'transformed' wages that would prevail in workplace j in year y:

$$M_{jy} - \sum_{i=1}^{I} \sum_{m=1}^{2} \frac{\tilde{w}_{jy}/t_{ijmy}^{\varepsilon}}{\sum_{k=1}^{I} \sum_{m=1}^{2} \tilde{w}_{ky}/t_{ikmy}^{\varepsilon}} N_{iy} = 0,$$

where $\tilde{w}_{jy} \equiv (B_{jy} w_{jy})^{\varepsilon}$ denote the transformed wages that are the actual wages weighted by the workplace amenities B_{jy}.

We are now equipped to determine the degree ε of workers' heterogeneity. Using temporal variation in the transformed wages $\tilde{w}_{jy}$ in municipality j located in prefecture z in year y and data on observed wages for different years, we can measure workers' heterogeneity as follows:

$$\log \tilde{w}_{jy} = \varepsilon \log w_{jy} + \mu_j + \mu_{j\in z,y} + \epsilon_{jy}, \tag{33}$$

where workplace amenities B_{jy} are absorbed by municipality-specific and prefecture-year-specific fixed effects μ_{jy} and $\mu_{j\in z,y}$, respectively. In other words, (33) is a linear regression of the estimated transformed wages on the wages observed in the data and fixed effects.

Including fixed effects may not be sufficient to address endogeneity because time-varying workplace amenities are potentially correlated to wages (Ahlfeldt *et al.*, 2020). A new Shinkansen line, for example, not only affects productivity but also improves access to recreational amenities.[15] Therefore, we adopt a Bartik-style shift-share instrument to instrument for wages w_{jy}. We use the employment shares in each municipality in 1981 and predict employment from 2001 onwards using the national employment growth in each of the 10 sectors. The idea is that national shocks to employment in different sectors (*e.g.*, trade liberalization) are unrelated to local changes in amenities, so that we identify a causal estimate of ε. In Appendix C.3 we report the results of the estimation procedure to obtain ε, including various robustness checks.

4.1.4 Step 4: Recovering the elasticity of substitution and land expenses

The remaining parameters for which we do not have clear priors are the elasticity of substitution between intermediates (σ) and the expenditure share by final firms spent on land (γ). We first use data from *JAXA* to obtain the share of built-up land in 2014, which we denote by $\tilde{L}_{iy}/\mathcal{A}_i$, where $\mathcal{A}_i$ is the total area size of a municipality. Then, using the land market equilibrium condition (27), we have

$$\begin{aligned} \log \frac{\tilde{L}_{iy}}{\mathcal{A}_i} &= \log \left[\zeta \left(\frac{L_{icy}}{\mathcal{A}_i} + \frac{1}{\sigma - 1} \frac{w_{iy} M_{i2y}}{R_{iy}\mathcal{A}_i} + \gamma \frac{w_{iy} M_{i1y}}{\beta R_{iy} \mathcal{A}_i} \right) \right] + \epsilon_{iy} \\ &= \log \left(\zeta_0 \frac{L_{icy}}{\mathcal{A}_i} + \zeta_1 \frac{w_{iy} M_{i2y}}{R_{iy}\mathcal{A}_i} + \zeta_2 \frac{w_{izy} M_{i1y}}{\beta R_{iy}\mathcal{A}_i} \right) + \epsilon_{iy}, \end{aligned} \tag{34}$$

where $\zeta = \zeta_0$, $\sigma = (\zeta_0 + \zeta_1)/\zeta_1$ and $\gamma = \zeta_2/\zeta_0$ and $y = 2014$. Note that ζ is a normalization constant that is 1 in the theoretical model.

In Appendix C.4 we show the results of estimating (34) in order to obtain σ and γ. We also estimate regressions where we include controls Z_{iy}. First, we control for the composition

[15] The direction of the bias is not entirely clear. If workplace amenities and wages are positively correlated, not controlling properly for workplace amenities will lead to an overestimate of ε. However, if amenities and wages are negatively correlated, ε will be underestimated.

of employment in different industries because, say, manufacturing firms consume much more land than retail firms, although they are both supplying (final) goods. Further, we control for geographical factors that may affect construction costs and jointly impact land use, employment, and land rents. Finally, we include region fixed effects to control for any regional differences in construction costs. Hence, we estimate:

$$\log \frac{\tilde{L}_{iy}}{\mathcal{A}_i} = \log \left(\zeta_0 \frac{L_{icy}}{\mathcal{A}_i} + \zeta_1 \frac{w_{iy} M_{i2y}}{R_{iy} \mathcal{A}_i} + \zeta_2 \frac{w_{iy} M_{i1y}}{\beta R_{iy} \mathcal{A}_i} \right) + \chi Z_{iy} + \mu_{j \in z,y} + \epsilon_{iy}.$$

4.1.5 Step 5: Recovering productivities

In the final step we recover a set of location fundamentals. These are productivities of final and intermediate firms (E_{j1y} and E_{j2y}, respectively). Note that the values of E_{j1y} and E_{j2y} that are consistent with the equilibrium may be obtained as follows. First, E_{j1y} is obtained from (18).

Furthermore, solving (29) with respect to E_{j2} leads to

$$\begin{aligned} E_{j2y} &= \left[\frac{\left(E_{j1y}\left(1-\beta-\hat{\gamma}\right)\right)^{\frac{1}{\beta+\hat{\gamma}}}}{\hat{\sigma}} \left(\frac{\hat{\sigma} w_{jy}}{\hat{\sigma}-1} \right)^{1-\hat{\sigma}} \sum_{i=1}^{I} \sum_{m=1}^{2} \tau_{jimy}^{1-\hat{\sigma}} M_{i1y}^{\frac{\beta}{\beta+\hat{\gamma}}} L_{i1y}^{\frac{\hat{\gamma}}{\beta+\hat{\gamma}}} \left(\frac{1}{P_{iy}} \right)^{\frac{1-\beta\hat{\sigma}-\hat{\gamma}\hat{\sigma}}{\beta+\hat{\gamma}}} \frac{1}{R_{jy} \ell_{jy}} \right]^{\frac{1}{1-\hat{\sigma}}} \quad (35) \\ &= \frac{\hat{\sigma} w_{jy}}{\hat{\sigma}-1} \left[\frac{\left(1-\beta-\hat{\gamma}\right)^{\hat{\sigma}}}{\hat{\sigma} R_{jy} \ell_{jy} E_{j1y}^{\frac{1}{(\beta+\hat{\gamma})(\hat{\sigma}-1)}}} \sum_{i=1}^{I} \sum_{m=1}^{2} \hat{\tau}_{jimy}^{1-\hat{\sigma}} M_{i1y}^{\frac{\beta}{\beta+\hat{\gamma}}} L_{i1y}^{\frac{\hat{\gamma}}{\beta+\hat{\gamma}}} \left(\frac{1}{P_{iy}} \right)^{\frac{1-\beta\hat{\sigma}-\hat{\gamma}\hat{\sigma}}{\beta+\hat{\gamma}}} \right]^{\frac{1}{1-\hat{\sigma}}}, \end{aligned}$$

where

$$L_{i1y} = \frac{\hat{\gamma} M_{i1y} w_{iy}}{\beta R_{iy}}.$$

Thus, using Newton-Raphson and the estimated parameters $\{\hat{\kappa}_{\mathcal{R}}, \hat{\kappa}_{\mathcal{H}}, \hat{\theta}_{\mathcal{R}}, \hat{\theta}_{\mathcal{R}}, \hat{\varepsilon}, \hat{\sigma}, \hat{\gamma}\}$ we get $2I$ equations that yield the desired values for E_{j2y}.

In Appendix E we consider the extension of the model where we allow for agglomeration economies in the intermediate sector using estimated productivities E_{j2y}. We show that our results are not materially influenced by this extension.

4.2 Results

We report the results of the preferred estimation of the structural parameters in Table 5. For each parameter, we provide a more elaborate discussion in Appendix C, including several robustness analyses. First of all, we estimate the railway and highway commuting travel time elasticities and find that the elasticity with respect to travel time by train is $\hat{\varkappa}_{\mathcal{R}} = 0.0678$, while the elasticity is almost as high as when traveling by car ($\hat{\varkappa}_{\mathcal{H}} = 0.0531$). This implies that a ten-minute travel time increase by train changes the number of commuters by $(\exp(-0.0678 \times 10) - 1) \times 100\% = -49\%$,

Table 5 — Structural parameters

	Baseline results (1)
Commuting railway travel time elasticity, $\hat{\varkappa}_{\mathcal{R}} = \hat{\kappa}_R \hat{\varepsilon}$	0.0678***
	(0.0004)
Commuting road travel time elasticity, $\hat{\varkappa}_{\mathcal{H}} = \hat{\kappa}_H \hat{\varepsilon}$	0.0531***
	(0.0010)
Trade railway travel time elasticity, $\hat{\vartheta}_{\mathcal{R}} = \hat{\theta}_R(1-\sigma)$	-0.0020***
	(0.0002)
Trade road travel time elasticity, $\hat{\vartheta}_{\mathcal{H}} = \hat{\theta}_H(1-\sigma)$	-0.0242***
	(0.0008)
Worker heterogeneity, $\hat{\varepsilon}$	2.8525***
	(0.5519)
Elasticity of substitution, $\hat{\sigma}$	1.9961***
	(0.6244)
Share of expenditure on land by final firms, $\hat{\gamma}$	0.0720***
	(0.0048)
Fixed parameters:	
Share of households spent on land, α	0.2300
Share of firms spent on labor, β	0.5000
Number of location pairs	2,748,964
Number of locations	1,658

Notes: Standard errors are bootstrapped (250 replications) by work locations and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

while a ten-minute increase in travel time by car decreases the number of commuters by 41%. The observation that travel time elasticities are similar is in line with that the modal split between train and car is similar. More specifically, according to the *Nationwide Person Trip Survey* by the *Ministry of Land, Infrastructure, Transport, and Tourism,* in Japan, $20-50\%$ of the commutes are by car, while about 20% of the commutes are by train.[16] In Appendix C.1, we show that our finding of similar commuting time elasticities for traveling by train and car holds in various alternative specifications.

Moving on to the trade travel time elasticities, we find that the rail travel time elasticity, $\hat{\vartheta}_{\mathcal{R}}$, is equal to -0.0020 and highly statistically significant, implying that the expenditure share on intermediates decreases by 11.3% for a 1-hour increase in railway travel time. This travel time elasticity is about 3% of the commuting time elasticity, which is in line with Monte *et al.* (2018). Since less than 1% of the transportation of goods is by rail, we may expect that this elasticity captures face-to-face business relationships among intermediate and final firms. Travel time by road also matters: for a 1-hour increase in travel time by road the expenditure share decreases

[16] Note that if we include either railway travel time *or* travel time by road, we find elasticities of respectively 0.089 and 0.064, which are similar to the literature.

by 77%.[17] Hence, these results suggest that travel time costs by train are considerably lower. We believe this makes sense as the Shinkansen offers high levels of comfort so that work-related activities can easily be undertaken while traveling. When traveling by car this is considerably harder. In 2017, the total distance of domestic passenger transport in Japan amounted to approximately 605 billion passenger kilometers, with railway transport accounting for 72.3% of the transport distance, while motor vehicles only account for 11.3% (and air travel for 16.4%). Hence, for long-distance travel, the train is by far the most preferred transport mode, which is in line with $\left|\hat{\vartheta}_{\mathcal{R}}\right| << \left|\hat{\vartheta}_{\mathcal{H}}\right|$. Appendix C.2 shows that this finding is robust, also if we control for the possibility that people travel by airplane.

Regarding workers' heterogeneity parameter, we find that $\hat{\varepsilon} = 2.85$, which we obtain by a regression of the transformed wages $\tilde{w}_{it}$ on wages given by the data (see Appendix A.1). Recall that we include municipality and region-by-year fixed effects and instrument observed wages by a Bartik-style predicted employment measure based on employment shares in 1981. Our estimate of ε is on the low side as compared to the existing literature. Eaton and Kortum (2002) find $\hat{\varepsilon} = 8.28$ for international trade flows, Donaldson and Hornbeck (2016) find $\hat{\varepsilon} = 8.22$ for the US in 1890, while Ahlfeldt *et al.* (2015) find $\hat{\varepsilon} = 6.2$ for Berlin. However, it is hard to compare these results directly because these papers deal with completely different issues. By contrast, for the U.S., Monte *et al.* (2018) find $\hat{\varepsilon} = 3.3$, which is very close to our estimate.[18] In Appendix C.3 we provide a range of robustness checks to find that our estimate of $\hat{\varepsilon}$ is largely robust.

We further estimate the elasticity of substitution, σ, and the share of final firms spent on land, γ. We obtain $\hat{\sigma}$ and $\hat{\gamma}$ by a regression of the log of the share developed land on measures of land use by households, intermediate firms, and final firms. The literature does not provide clear priors on what σ should be. For example, Hsieh *et al.* (2019) work with the elasticity of substitution between occupations but have no information on this parameter. In their baseline model, they use $\sigma = 3$ but allow for values that vary from 1.05 to 10. We find that $\hat{\sigma} = 2.00$. Hence, intermediates are reasonably differentiated. We think this estimate is a useful benchmark for other studies that need to assume a value of the elasticity of substitution between intermediate services. We find that the share of expenditures of final firms on land is about 7%, which is very similar to that found by Valentinyi and Herrendorf (2008) for U.S. firms. In Appendix C.4, we show that the estimates of $\hat{\sigma}$ and $\hat{\gamma}$ are robust, albeit not always precise, across specifications when we include detailed fixed effects, industry employment shares, and geographical attributes.

[17] If we include only travel time by road as in Monte *et al.* (2018), we find an elasticity of about −0.011. For the travel time by rail it is −0.021. These are both comparable to the elasticities assumed in Monte *et al.*

[18] Moreover, our approach better addresses endogeneity concerns, such as unobserved workplace amenities, which leads to a somewhat lower estimate. Ahlfeldt *et al.* (2015) recover ε by comparing the variances of log-transformed wages to the variance of log observed wages. This approach would here lead to $\varepsilon = 77.79$, which is unrealistically large. This overestimate is likely to be a result of the correlation between unobserved workplace amenities and wages (Ahlfeldt *et al.*, 2020).

5 Counterfactual analyses

5.1 Model performance

We aim to assess whether our model can explain past changes in employment across locations by undertaking counterfactual experiments given $\{\alpha, \beta\}$ and the estimated parameters $\{\hat{\kappa}_{\mathcal{R}}, \hat{\kappa}_{\mathcal{H}}, \hat{\theta}_{\mathcal{R}}, \hat{\theta}_{\mathcal{H}}, \hat{\varepsilon}, \hat{\sigma}, \hat{\gamma}\}$, estimated amenities ($A_{i2014}$ and B_{i2014}), TFPs ($E_{i1,2014}$ and $E_{i2,2014}$), land use by intermediate firms (ℓ_{i2014}), and total land use (L_{i2014}).

Our model is static and therefore we only should evaluate long-term changes in employment. Hence, we go back to 1957 and aim to back-cast employment levels and compare the predicted employment levels given the 1957 transport network (so before the Shinkansen was built) to the values observed in the data. We describe the procedure to develop the counterfactual values in Appendix D.1. Since we do not have good data on the secondary road network in 1957, we use data on the road network from the early 1900s. We then regress the observed change in employment density in the data on the counterfactual change in employment:

$$\log\left(\frac{M_{i2014}}{M_{i1957}}\right) = \phi_0 + \phi_1 \log\left(\frac{M_{i2014}}{\hat{M}_i^{\mathcal{C}}}\right) + \epsilon_{iy},$$

where $\hat{M}_i^{\mathcal{C}}$ is the estimated counterfactual value given the transport network in 1957, M_{iy} are the observed employment in either 1957 or 2014; ϕ_0 and ϕ_1 are parameters to be estimated, and ϵ_{iy} is an error term. Because $\hat{M}_i^{\mathcal{C}}$ is estimated, we should incorporate the variance in the underlying model's parameters when calculating standard errors. In so doing, we obtain standard errors by bootstrapping the whole structural estimation approach and predict $\hat{M}_i^{\mathcal{C}}$ 250 times. The regressions are weighted by the employment in each municipality in 2014, both to estimate the average effect for the counterfactual analysis as well as to minimize the influence of outliers. Table 6 reports the results.

We find a positive elasticity, ϕ_1 of the predicted change on the observed change in employment. More specifically, a 1% increase in the predicted change in employment increases the observed change in employment by 0.17%. We are not surprised that this elasticity is not closer to 1 as there have been many changes in amenity and productivity levels that are likely to bring down ϕ_1.

One may worry that this result is mostly driven by the location fundamentals and by the relative position of municipalities within transport networks, *e.g.* because Tokyo is more centrally located. In column (2), we therefore include prefecture fixed effects that should mitigate this issue to a large extent. The elasticity almost doubles (*i.e.*, $\hat{\phi}_1 = 0.297$). Column (3) shows that the unweighted estimate delivers a somewhat lower coefficient.

TABLE 6 — MODEL PERFORMANCE: BACK-CASTING CHANGES IN EMPLOYMENT

(Dependent variable: the change in employment between 1957 and 2014)

	Baseline	+ Prefecture FE	Unweighted	Euclidean	+ Euclidean	Add 2014 empl
	(1)	(2)	(3)	(4)	(5)	(6)
	WLS	WLS	OLS	WLS	WLS	WLS
Predicted change in employment	0.1704**	0.2972***	0.1543***		0.2744***	0.3024***
1957-2014 *(log)*	(0.0825)	(0.0772)	(0.0391)		(0.0792)	(0.0774)
Predicted change in employment				0.1173**	0.0365	0.0067
using Euclidean distance *(log)*				(0.0481)	(0.0528)	(0.0484)
Employment density in 2014 *(log)*						0.0521
						(0.0368)
Prefecture fixed effects	No	Yes	Yes	Yes	Yes	Yes
Number of observations	1,658	1,658	1,658	1,658	1,658	1,658
R^2	0.008	0.269	0.256	0.256	0.269	0.275

Notes: In columns (1), (2)-(6) we weight the estimates by the employment in 2014. Standard errors are bootstrapped (250 replications) by the municipality and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

A main concern is that $\hat{M}_i^{\mathcal{C}}$, based on changes in transportation infrastructure, is correlated to unobserved changes in the relative positions of municipalities within Japan but independently of any transport network. For example, Tokyo has become more attractive but at the same time has seen improvements in transportation infrastructure. To address this concern, we calculate the counterfactual change in employment for the same location fundamentals *when the travel time between municipalities is given by the Euclidean distance.* The model-predicted employment density is a rather weak predictor of the change in employment density between 1957 and 2014, with an elasticity equal to 0.117 (see column (4)), which is considerably lower than that associated with the counterfactual employment density based on the transport network. In column (5), we include the counterfactual employment change based on Euclidean distance as a control variable and find that the impact of the counterfactual employment based on Euclidean distances is small and statistically insignificant. More importantly, the elasticity of the model-predicted employment change is again close to 0.3. Hence, *the predictions based on the actual transport network distance outperform the measure based on Euclidean distance.* This increases our belief that our model is able to capture relevant changes in employment due to changes in the transport network.

In column (6) we further control for the employment density in 2014, as denser areas may be on different trends. This does not appear to be the case as the coefficient is statistically insignificant. The elasticity of the predicted change in employment is similar to previous estimates (*i.e.*, $\hat{\phi}_1 = 0.302$).

In sum, we believe that this exercise shows that our model is very capable of reproducing the observed changes in the spatial distribution of employment.

Table 7 — Counterfactual experiments

	Experiment 1: No Shinkansen	Experiment 2: Extended Shinkansen	Experiment 3: No Highways
	(1)	(2)	(3)
Change in average travel time to employment by train *(in %)*	85.78	-10.86	0.02
	[83.24, 86.36]	[-10.99, -10.82]	[0.00, 0.04]
Change in average travel time to employment by road *(in %)*	0.10	0.03	58.62
	[-1.28, 0.45]	[-0.01, 0.11]	[58.62, 58.64]
Change in gross welfare *(in %)*	-21.71	1.60	-0.43
	[-46.67, -2.96]	[-13.93, 2.58]	[-11.05, -0.05]
Change in total production by final firms *(in %)*	-35.22	1.13	-0.85
	[-64.30, -4.67]	[-33.26, 3.19]	[-19.97, -0.02]
Change in total land rents *(in %)*	10.03	6.97	0.08
	[-9.54, 48.90]	[0.52, 144.34]	[-0.34, 48.46]

Notes: 95% confidence intervals are bootstrapped (250 replications) by municipality and in brackets.

5.2 Counterfactual experiments – aggregate effects

We consider three counterfactual experiments. First, we consider the effects of removing the Shinkansen. Second, we discuss the effects of the hypothetical realization of the planned extensions of the Shinkansen network. Third, as a comparison, we investigate the effects of removing Japan's highways.

We calculate the average travel time to employment in 2014 and show in Appendix D.2 what areas are the most affected. We report the aggregate results in Table 7.

Experiment 1: No Shinkansen. Our first experiment shows that *removing the Shinkansen as a whole would have substantial negative welfare effects: the gross welfare would decrease by* 21%. We do not think this welfare effect comes as a surprise as the increase in average travel time will increase by 86%. One may be surprised to see that average travel times by road change a little. This is due to the reshuffling of the population and employment as a response to changes in railway travel times.

Do the positive welfare effects of the Shinkansen exceed construction and operating costs? The total annual discounted costs are about ¥1.8 trillion, while annual benefits are an order of magnitudes larger, *i.e.*, ¥117 trillion. Hence we find strong positive effects on the net welfare of the construction of the Shinkansen.[19] The confidence bands imply that the lower bound of annual

[19] From the *Ministry of Land, Infrastructure, Transport, and Tourism*, the construction of a km of Shinkansen costs about ¥14.5 billion in 2022 prices. If we extrapolate these costs to the full network, this amounts to ¥37 trillion. Given a discount rate of 3.5%, this amounts to about ¥1.3 trillion per year. This is probably an upper-bound estimate, as these are the costs of the planned Chuo Shinkansen if it would be using 'ordinary' Shinkansen technology rather than Maglev technology. More specifically, the planned costs for the Tokaido Shinkansen were

benefits are ¥17 trillion, which still vastly exceed the costs.

In line with the effects on welfare, the total output of the final sector decreases by 35% if the Shinkansen were to be removed.[20] Further, a priori the sign of the effect on land values is hard to guess. On the one hand, higher transport costs imply that firms can spend less on land and labor. On the other hand, because transport costs are higher without the Shinkansen, firms and people will concentrate more, which in turn raises land values. Indeed, in the presence of high transport costs, firms and people want to be close to each other (Pflüger and Tabuchi, 2010). Yet, it appears that total land revenues increase by 10% if the Shinkansen were to be removed. Recall that landlords are absentee, which explains most of the gap between the reductions between total output and welfare.

Experiment 2: Shinkansen extensions. In column (2) of Table 7, we consider the scenario in which all planned Shinkansen lines are built (see the dashed lines in Figure 1). This includes *(i)* an extension from Hokuto to Sapporo, *(ii)* an extension to Nagasaki, *(iii)* a link between Kanazawa and Kyoto, and more importantly, *(iv)* a project to connect Tokyo and Osaka by a 'Maglev' with a maximum speed of 505 km/h. This HSR is expected to connect Tokyo and Nagoya in 40 minutes, and Tokyo and Osaka in 67 minutes. Compared with the existing Tokaido Shinkansen, the Maglev line will cut traveling time by half. The commercial service is scheduled to start between Tokyo and Nagoya in 2027 and between Nagoya and Osaka in 2045.

Column (2) in Table 6 shows that the average travel time by train across all location pairs is reduced by 11%. We find positive welfare effects of the new Shinkansen lines: the gross welfare increases by 1.6%. However, this effect is not statistically significantly different from zero because the 95% confidence bands range from -9.8% to 2.3%. Hence, this implies that *the Japanese economy may not benefit from the planned extensions*. Since the current Shinkansen network already links almost all large Japanese cities, the planned extensions of the Shinkansen network are possibly an example of over-provision of transport infrastructure.

Experiment 3: No highways. Our final experiment is to compare the effects of high-speed rail to the effects of highways. We therefore consider removing all highways, which will imply an

¥200 billion, which was severely understated. The actual costs of construction were about ¥400 billion, which is ¥1,750 billion in 2022 prices (Smith, 2003). The segment between Tokyo and Shin-Osaka is 515km, implying a cost per km of ¥3.4 billion. The total Shinkansen network in 2014 was 2667km, hence the approximated total costs are about ¥9 trillion. We emphasize that the quantitative and qualitative conclusions are hardly affected by choosing this lower bound of construction costs. Operating costs are about ¥0.5 billion per km per year, which are also from the *Ministry of Land, Infrastructure, Transport, and Tourism*. There is some uncertainty about the actual operating costs. For the Hokkaido Shinkansen, statistics suggest lower operating costs of ¥0.1 billion per km per year. We once again emphasize that this does not affect our conclusions. Japanese GDP was ¥539 trillion in 2021.

[20] This result is reinforced once we allow for agglomeration economies (see Appendix E.3).

increase in average travel time by road of almost 60% (see column (3) in Table 7). We find that the welfare effects are considerably smaller than the effects of removing the Shinkansen: welfare decreases by only 0.43%, while total land revenues are not materially influenced. The reason why welfare costs of removing highways are considerably smaller than the welfare costs of removing the Shinkansen is that costs of traveling by road have been estimated to be considerably higher than by train (*i.e.*, $\left|\hat{\vartheta}_{\mathcal{H}}\right| >> \left|\hat{\vartheta}_{\mathcal{R}}\right|$) so that it is unattractive to ship intermediate services by car. Hence, removing highways will have limited effects because most intermediates are shipped by train. This is in line with descriptive data showing that railway passenger transport accounts for 72.3% of passenger-kilometers, while for roads account for only 11.3%.

5.3 Counterfactual experiments – local effects

We now consider the local effects of the counterfactual experiments. Because maps can be somewhat hard to read, we will put them in Appendix D.2 and instead repeat our econometric approach from Section 2.4 to study what happens to locations close to Shinkansen stations. More specifically, we estimate:

$$\log\left(\frac{M_j^{\mathcal{C}}}{M_{j2014}}\right) = \beta_0 + \beta_1 \mathcal{S}_{j2014} + \beta_2 \mathcal{X}_{j2014} + \epsilon_j, \qquad (36)$$

where β_1 captures the effect on the predicted change in local employment if locale j had a Shinkansen station within 10km in 2014, *i.e.*, $\mathcal{S}_{j2014}$, while β_2 captures the effect of having an 'ordinary' railway station or highway connection within 10km, *i.e.*, $\mathcal{X}_{j2014}$. Table 8 reports the results of (36).

Experiment 1: No Shinkansen. In columns (1) and (2) in Table 8 we show that locations that have a Shinkansen station in 2014 will lose employment once the Shinkansen network would be removed. The change is $(\exp(-0.2466) - 1) \times 100\% = -21.8\%$, which is similar to the reduced-form results (see column (3), Table 4). This improves the belief that our model makes meaningful predictions on changes in the spatial distribution of employment across Japan. It is clear that the Shinkansen network is interlinked with the 'ordinary' railway network, as locations that have a railway station within 10km would also witness a reduction in employment of 12.6% if the Shinkansen were to be removed. Unsurprisingly, municipalities near highways do not see much change when the Shinkansen were to be removed.

In Figure D.1 in Appendix D.2, we plot the spatial effects. First, we observe the largest travel time increases (up to 150%) along the corridor where the Shinkansen lines are and in municipalities that had a station. The effects are particularly large on Kyushu. Employment and residential

TABLE 8 — COUNTERFACTUAL EXPERIMENTS: REGRESSIONS
(Dependent variable: the change in counterfactual employment)

	Scenario 1: No Shinkansen		Scenario 2: Extended Shinkansen		Scenario 3: No Highways	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Shinkansen station in 2014 <10km	-0.2466*	-0.1901*	-0.0009	-0.0019	-0.0008	-0.0014
	(0.1293)	(0.1106)	(0.0010)	(0.0012)	(0.0008)	(0.0011)
Planned Shinkansen station <10km			0.0270	0.0139		
			(0.0209)	(0.0104)		
Railway station <10km	-0.1342*	-0.1194*	0.0080	0.0065	0.0007*	0.0000
	(0.0726)	(0.0706)	(0.0063)	(0.0045)	(0.0004)	(0.0002)
Highway <10km	-0.0046	0.0030*	0.0093	0.0061	-0.0042**	-0.0044**
	(0.0064)	(0.0016)	(0.0063)	(0.0039)	(0.0020)	(0.0021)
Geographic controls	No	Yes	No	Yes	No	Yes
Region fixed effects	No	Yes	No	Yes	No	Yes
Number of observations	1,658	1,658	1,658	1,658	1,658	1,658
R^2	0.2022	0.6680	0.0179	0.1149	0.1774	0.3460

Notes: Geographic controls include the log of area size, the log of population in 1872, the share of developed land, the distance to the coast, as well as the mean elevation. Standard errors are bootstrapped (250 replications) by municipality and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

population are particularly decreasing in areas that did not benefit from the Shinkansen. For example, Hokkaido in the North witnesses an increase in employment and population. Wage and land rent decreases are large (up to 50%) and strongly correlated to the reductions in travel time by rail, with the reductions being the greatest in Kyushu and in the Northeastern part of Honshu.

The effects on the Tokyo and Kyoto-Osaka metropolitan areas are positive, as Tokyo and Nagoya would be 1.2% and 2.2% smaller in terms of employment without the Shinkansen, respectively, which shows that the Shinkansen has somewhat amplified the 'hub effects' of these metropolitan areas. By contrast, Osaka would be 0.7% larger. In general, contrary to general belief, *the Shinkansen has generally enhanced growth of Tokyo even further*. This is in line with the results obtained by Qin (2017) for China.

Experiment 2: Shinkansen extensions. Let us turn to the local effects of the second experiment, which considers the planned extensions of the Shinkansen network. In Table 8 we show that employment within 10km of a planned Shinkansen station seems to imply employment growth of about 2%, although the result is not statistically significantly different from zero. Like in the previous scenario, the effect on ordinary railway stations is about 50% of this effect. Also Figure D.2 in Appendix D.2 highlights some positive effects on employment, wages and land rents in connected areas, such as Nagasaki and Sapporo. On the other hand, Tokyo, Nagoya and Osaka do not seem to benefit much from the new Maglev line.

Experiment 3: No highways. Columns (5) and (6) in Table 8 show that municipalities with a highway connection lose about 0.4% of employment if highways were to be removed. Although travel time by road is reduced substantially in parts of Kyushu and Honshu (see Figure D.3 in Appendix D.3), *changes in employment, residential population, land rents, and wages are considerably smaller than the effects triggered by removing the Shinkansen*. An important reason for this is that the travel-time costs of services by road are considerably higher (recall that $|\vartheta_{\mathcal{H}}| > |\vartheta_{\mathcal{R}}|$) so that highways are considerably less important in facilitating long-distance business-to-business travel.

6 Conclusions

This paper estimates the effects of large infrastructure investments on the geographic distribution of economic activities. As high-speed rail is on the rise in many countries because it is seen as a sustainable alternative to the airplane on medium-distance travels, we have chosen to focus on Japan which has the oldest HSR networks in the world, *i.e.*, the Shinkansen. This enables us to evaluate welfare and the long-run spatial effects of infrastructure investments in high-speed rail. To achieve our goal, we develop a new spatial quantitative model in which *(i)* input-output linkages are formed between intermediate firms providing services and final firms producing goods. Services are then shipped by train *and* by road; *(ii)* workers choose where to live and where to work; while they can commute by train *and* by car, and *(iii)* intermediate firms, final firms, and workers compete for land.

We apply the model using data from Japan, enabling us to estimate relevant model parameters, including the business travel time elasticity by railway and road, as well as the elasticity of substitution among intermediate services. We also show that our model is able to back-cast employment changes reasonably well, which underscores the importance of transport networks in explaining the spatial population and employment distributions.

We conduct counterfactual experiments in which *(i)* the entire Shinkansen network is removed, *(ii)* all planned Shinkansen lines are realized, and *(iii)* all highways are removed. The Shinkansen network has generated a sizable welfare gain of about 20% , with the lower bound being about 3%. By comparing the benefits to construction and operation costs, the benefits far outweigh the costs. Hence, the Shinkansen yields positive benefits to the Japanese economy. However, the Shinkansen has not been successful in promoting economic growth and development outside of Tokyo. Finally, the welfare effects of removing highways are considerably smaller than those generated by removing the Shinkansen, which underlines the importance of business-to-business travel.

References

[1] Abrantes, P. A., Wardman, M. R. (2011). Meta-analysis of UK values of travel time: An update. *Transportation Research Part A: Policy and Practice*, 45(1), 1-17.

[2] Ahlfeldt, G.M., Bald, F., Roth, D., Seidel, T. (2020). The Stationary Spatial Equilibrium with Migration Costs. *Mimeo*, London School of Economics.

[3] Ahlfeldt, G.M., Feddersen, A. (2018). From Periphery to Core: Measuring Agglomeration Effects Using High-speed Rail. *Journal of Economic Geography* 18: 355-90.

[4] Ahlfeldt, G.M., Redding, S.J., Sturm, D.M., Wolf, N. (2015). The Economics of Density: Evidence from the Berlin Wall. *Econometrica*, 83: 2127-89.

[5] Allen, T., Arkolakis, C., Li, X. (2015). On the Existence and Uniqueness of Trade Equilibria. *Mimeo*, Yale University.

[6] Banerjee, A., Duflo, E., Qian, N. (2020). On the Road: Access to Transportation Infrastructure and Economic Growth in China. *Journal of Development Economics*, 102442.

[7] Baum-Snow, N. (2007). Did Highways Cause Suburbanization? *Quarterly Journal of Economics* 122: 775-805.

[8] Baum-Snow, N., Brandt, L., Henderson, J.V., Turner, M.A., Zhang, Q. (2017). Roads, Railroads, and Decentralization of Chinese Cities. *Review of Economics and Statistics* 99: 435-48.

[9] Baum-Snow, N., Henderson, J.V., Turner, M.A., Zhang, Q., Brandt, L. (2020). Does Investment in National Highways Help or Hurt Hinterland City Growth? *Journal of Urban Economics* 115: 103124.

[10] Bayer, P., Timmins, C. (2007). Estimating Equilibrium Models of Sorting Across Locations. *Economic Journal* 117: 353-74.

[11] Behrens, K., Lamorgese, A.R., Ottaviano, G.I.P., Tabuchi, T. (2007). Changes in Transport and Non-transport Costs: Local vs Global Impacts in a Spatial Network. *Regional Science and Urban Economics* 37: 625-48.

[12] Behrens, C., Pels, E. (2012). Intermodal Competition in the London–Paris Passenger Market: High-Speed Rail and Air Transport. *Journal of Urban Economics*, 71: 278-88.

[13] Berger, T., Enflo. K. (2017). Locomotives of Local Growth: The Short- and Long-term Impact of Railroads in Sweden. *Journal of Urban Economics* 98: 124-38.

[14] Bernard, A.B., Moxnes, A., Saito, Y.U. (2019). Production Networks, Geography, and Firm Performance. *Journal of Political Economy* 127: 639-88.

[15] Blundell, R., Powell, J. L. (2003). Endogeneity in Nonparametric and Semiparametric Regression Models. *Econometric Society Monographs* 36: 312-357.

[16] Brinkman, J., Coen-Pirani, D., Sieg, H. (2015). Firm Dynamics in an Urban Economy. *International Economic Review* 56: 1135-64.

[17] Charnoz, P., Lelarge, C., Trevien, C. (2018). Communication Costs and the Internal Organization of Multi-plant Businesses: Evidence from the Impact of the French High-speed Rail. *Economic Journal* 128: 949-94.

[18] Ciccone, A., Hall, R. E. (1996). Productivity and the Density of Economic Activity. *American Economic Review* 86: 54-70.

[19] Combes, P.-P. and Gobillon, L. (2015). The Empirics of Agglomeration Economies. In: *Handbook of Regional and Urban Economics. Volume 5*, edited by G. Duranton, J.V. Henderson and W.C. Strange, 247-348. Amsterdam: Elsevier.

[20] Combes, P.P., Duranton, G., Gobillon, L. (2019). The Costs of Agglomeration: House and Land Prices in French cities. *Review of Economic Studies* 86: 1556-89.

[21] Coto-Millán, P., Inglada, V., Rey, B. (2007). Effects of Network Economies in High-speed Rail: the Spanish Case. *Annals of Regional Science* 41: 911-925.

[22] Davis, D.R., Weinstein, D.E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review 92*: 1269-89.

[23] De Rus, G., Inglada, V. (1997). Cost-benefit Analysis of the High-speed Train in Spain. *Annals of Regional Science* 31: 175-188.

[24] Degryse, H., Ongena, S. (2006). Distance, Lending Relationships, and Competition. *Journal of Finance*, LX(1), 231-266.

[25] Dericks, G., Koster, H.R.A. (2021). The Billion Pound Drop: The Blitz and Agglomeration Economies in London. *Journal of Economic Geography*, forthcoming.

[26] Dijkman, H., Koopmans, C. C., Vromans, M. (2000). Cost-Benefit Analysis of High Speed Rail. *CPB Report*.

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

[28] Donaldson, D. (2018). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review* 108: 899-934.

[29] Eaton, J., Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica* 70: 1741-79.

[30] Ellison, G., Glaeser, E.L. Kerr, W.R. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review* 100: 1195-213.

[31] Ethier, W. (1982). National and International Returns to Scale in the Modern Theory of International Trade. *American Economic Review* 72: 389-405.

[32] Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System, *Review of Economic Studies* 81: 1046-70.

[33] Garcia-López, M. À., Holl, A., Viladecans-Marsal, E. (2015). Suburbanization and Highways in Spain when the Romans and the Bourbons still Shape its Cities. *Journal of Urban Economics* 85: 52-67.

[34] Head, K., Mayer, T. (2000). Non-Europe: The Magnitude and Causes of Market Fragmentation in Europe. *Review of World Economics* 136: 285-314.

[35] Hilber, C. A., Vermeulen, W. (2016). The Impact of Supply Constraints on House Prices in England. *Economic Journal* 126: 358-405.

[36] Hollander, S., Verriest, A. (2016). Bridging the Gap: The Design of Bank Loan Contracts and Distance. *Journal of Financial Economics* 119: 399-419.

[37] Hsieh, C.-T., Hurst, E., Jones, C.I., Klenow, P.J. (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica* 87: 1439-74.

[38] JR Central (2020). Central Japan Railway Company Annual Report. Nagoya.

[39] Karadi, P., Koren, M. (2017). Cattle, Steaks and Restaurants: Development Accounting When Space Matters. *Mimeo*, Central European University.

[40] Kito, H. (1996). The Regional Population of Japan Before the Meiji Period *(in Japanese*). *Jochi Keizai Ronsyu* 41: 65-79.

[41] Kline, P., Moretti, E. (2014). Local Economic Development, Agglomeration Economics, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics* 129: 275-331.

[42] Kanemoto, Y. and Tokuoka, K. (2002). Proposal for the standards of metropolitan areas of Japan (in Japanese). *Oyo Chiikigaku Kenkyu* 7: 1-5.

[43] Koster, H.R.A. (2023). The Welfare Effects of Greenbelt Policy: Evidence from England. *Economic Journal* (conditionally accepted).

[44] Koster, H., Tabuchi, T., Thisse, J.- F. (2021). To be Connected or Not to Be Connected? The Role of Long-haul Economies. *Journal of Economic Geography* 22: 711-53.

[45] Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review* 70: 950-59.

[46] Krugman, P., Venables, A.J. (1995). Globalization and the Inequality of Nations. *Quarterly Journal of Economics* 110: 857-80.

[47] Larch, M., Wanner, J., Yotov, Y. V., Zylkin, T. (2019). Currency Unions and Trade: A PPML Re-assessment with High-dimensional Fixed Effects. *Oxford Bulletin of Economics and Statistics* 81: 487-510.

[48] Melo, P. C., Graham, D. J., Noland, R. B. (2009). A Meta-analysis of Estimates of Urban Agglomeration Economies. *Regional Science and Urban Economics* 39: 332-42.

[49] Ministry of Land, Infrastructure, Transport and Tourism (2019). *White Paper on Transport Policy.* Tokyo: Shobi Printing Co., Ltd.

[50] Monte, F., Redding, S.J., Rossi-Hansberg, E. (2018). Commuting, Migration, and Local Employment Elasticities. *American Economic Review* 108: 3855-90.

[51] NHK (2015). Survey on Commuting Times. *NHK Culture Research Institute.*

[52] Okamoto, C., Sato, Y. (2021). Impacts of High-speed Rail Construction on Land Prices in Urban Agglomerations: Evidence from Kyushu in Japan. *Journal of Asian Economics*, 76, 101364.

[53] Pflüger, M., Tabuchi, T. (2010). The Size of Regions with Land Use for Production *Regional Science and Urban Economics* 40: 481-89.

[54] Qin, Y. (2017). No County Left behind? The Distributional Impact of High-speed Rail Upgrades in China. *Journal of Economic Geography* 17: 489-520.

[55] Redding, S.J., Turner, M.A. (2015). Transportation Costs and the Spatial Organization of Economic Activity. In: *Handbook of Regional and Urban Economics. Volume 5*, edited by G. Duranton, J.V. Henderson, and W.C. Strange 247-348. Amsterdam: Elsevier.

[56] Rosenthal, S.S., and Strange, W.C. (2004). Evidence on the Nature and Sources of Agglomeration Economies. In: *Handbook of Regional and Urban economics. Volume 4*, edited by J.V. Henderson and J.-F. Thisse 2119-71. Amsterdam: Elsevier.

[57] Rose, L. A. (1992). Land Values and Housing Rents in Urban Japan. *Journal of Urban Economics* 31(2): 230-251.

[58] Saiz, A. (2010). The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics* 125: 1253-96.

[59] Sato, N. (2015). *History of Shinkansen* (In Japanese), Tokyo: Chuokoron-Shinsha.

[60] Smith, K. (2021). Chuo maglev costs soar. *International Railway Journal (online).*

[61] Smith, R. A. (2003). The Japanese Shinkansen: Catalyst for the Renaissance of Rail. *Journal of Transport History* 24: 222-237.

[62] Sugiyama, M. (2015). Automobile and Road Transport Policies in Japan. *Technical report.*

[63] Takayama, Y., Ikeda, K., Thisse, J- F. (2020). Stability and Sustainability of Urban Systems under Commuting and Transportation Costs. *Regional Science and Urban Economics* 84: 103553.

[64] Thomas, I. (2002). *Transportation Networks and the Optimal Location of Human Activities: A Numerical Geography Approach*. Cheltenham, U.K.: Edward Elgar.

[65] Valentinyi, A., Herrendorf, B. (2008). Measuring Factor Income Shares at the Sectoral Level. *Review of Economic Dynamics* 11(4), 820-835.

[66] Wrona, J. (2018). Border Effects without Borders: What Divides Japan's Internal Trade? *International Economic Review* 59: 1209-62.

[67] Zheng, S., Kahn, M.E. (2013). China's Bullet Trains Facilitate Market Integration and Mitigate the Cost of Mega City Growth. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 110: 1248-53.

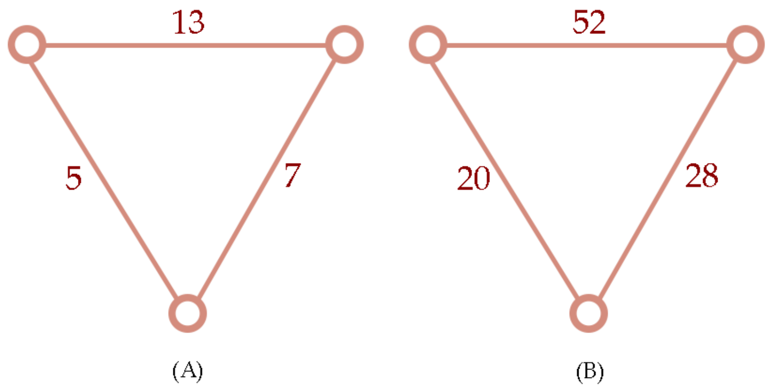


Figure 1: FIGURE A.1 — FROM TRADE LINKS TO TRADE VALUE: AN EXAMPLE

Appendix A. Data appendix

A.1 Trade in intermediates

Using input-output data at the sectoral level we normalize the link data from TSR to obtain trade values. Consider the following example in Figure A.1 where vertices represent municipalities. We focus on one industry-pair, say, finance and manufacturing. We then count the links between all municipalities for each industry-pair, for example. In Figure A.1.A, we have in total 25 links. Assume now that the observed value of trade from Input-output tables between these two industries in a certain year was 100. In this case, by multiplying the number of links by $100/25 = 4$, which implies that trade values are, respectively, 52, 20 and 28, as shown in Figure A.1.B. We repeat this normalization for each industry-pair for each year. To calculate the total value of trade between intermediate services and final firms for each municipality-pair, we then sum the normalized trade value across sector-pairs.

A.2 Wages

Data on wages at the municipality level are only available for the manufacturing sector, going back to 1979. However, we compiled a panel dataset at the municipality level of wages including all industrial sectors.[21] We first digitize hardcopy wage data on 47 *prefectures* for 8 industrial sectors, including manufacturing. Furthermore, we digitize hardcopy data on employment at the municipality level since 1979 for these 8 industrial sectors. Using data at the prefecture level, we estimate the following regression:

[21] The industrial sectors included are Construction, Electricity Production, Real Estate, Finance and Insurance, Manufacturing, Mining, Retail, Consumer Services, and Transportation.

TABLE A.1 — CALCULATING WAGES

(Dependent variable: the average wage at the prefecture level)

	1978	1981	1986	1991	1996	2001	2006	2009	2014
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Manufacturing wage *(log)*	0.2380*	0.3267***	0.3973***	0.4348***	0.4500***	0.3740***	0.4083***	0.2989***	0.2146***
	(0.1230)	(0.1086)	(0.1196)	(0.1250)	(0.1133)	(0.0796)	(0.0806)	(0.0717)	(0.0647)
Share construction workers	-2.1358*	-2.7058**	-1.0596	-1.4324	0.4353	-0.2360	-0.3830	-0.8728	-0.8281
	(1.2297)	(1.2723)	(1.2956)	(1.1525)	(1.2938)	(1.0700)	(0.8919)	(0.5804)	(0.5057)
Share workers in electricity	11.3916	13.0345*	7.9297	6.5816	-4.4572	-5.1545	-1.7843	-2.3662	-0.5150
and water supply	(7.0429)	(7.1504)	(7.4964)	(6.3646)	(7.3838)	(6.6586)	(6.2655)	(5.1301)	(5.0609)
Share workers in finance	3.5136	5.8406**	1.7358	7.9878***	6.2047***	6.2058***	7.1250***	6.7369***	6.6541***
and real estate	(2.6889)	(2.4597)	(2.9309)	(2.2331)	(2.0944)	(1.9118)	(2.4899)	(2.1552)	(1.7403)
Share workers in mining	39.4181	18.3458	18.8571	10.7510	-0.2813	-10.3284	-0.2088	0.6528	-1.1953
	(29.2538)	(21.1227)	(23.3166)	(15.5767)	(16.1820)	(10.9652)	(5.1147)	(3.1958)	(2.6583)
Share workers in retail	-0.2501	0.4745	0.2561	0.4115	1.2294*	0.0251	0.4391	0.2213	0.3302
	(0.7356)	(0.5698)	(0.6476)	(0.6593)	(0.6206)	(0.4932)	(0.5259)	(0.3853)	(0.3661)
Share workers in consumer services	-0.8389**	-0.9293***	-1.2054***	-0.5693	-1.3374***	-1.0721**	-1.4677***	-1.0226**	-1.1941***
	(0.3681)	(0.3288)	(0.4413)	(0.4249)	(0.4696)	(0.4777)	(0.4622)	(0.3957)	(0.3960)
Share workers in transport	3.1862***	2.3755***	2.9780***	-0.4233	-1.2968	1.1954	-0.7263	0.6543	1.1197
and logistics	(0.7486)	(0.6469)	(0.7581)	(1.1206)	(1.4353)	(1.2218)	(1.3213)	(0.9416)	(0.7223)
Observations	46	46	46	46	46	46	46	46	46
R^2	0.8016	0.8665	0.8307	0.7966	0.7875	0.8457	0.8491	0.8639	0.8766

Notes: The omitted category is the share of workers in manufacturing. Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

$$w_{zy} = \psi_y + \psi_y^{\mathcal{M}} w_{zy}^{\mathcal{M}} + \sum_{g=1}^{G} \psi_{yg} e_{zyg} + \epsilon_{zy},$$

where w_{zy} is the monthly wage observed in prefecture z in year y, $w_{zt}^{\mathcal{M}}$ is the (annual) manufacturing wage, e_{zyg} are employment shares in sectors $g = 1, ..., G$, ψ_y, $\psi_y^{\mathcal{M}}$, ψ_{tg} are coefficients to be estimated, and ϵ_{zy} denotes an error term. Hence, the above specification yields year-specific regressions of wages on manufacturing wages and employment shares.

In Table A.1 we show the results of these regressions. There appears to be a strong correlation (almost 0.893) between manufacturing wages and overall wages at the prefecture level. We find that the elasticity of manufacturing wages with respect to overall wages is about 0.3. Employment in financial services leads to markedly higher wages: a 10 percentage point increase in the share of workers in financial services is associated with a wage increase of about 80%. For consumer services, the effect is consistently negative with wages that are about 10% lower for a 10 percentage point increase in the share of workers in consumer services.

To calculate wages at the municipality level, we use the estimated parameters $\hat{\psi}_y$, $\hat{\psi}_y^{\mathcal{M}}$, $\hat{\psi}_{tg}$, together with the manufacturing wages and employment shares at the municipality level available in the data. Figure A.2 shows the average annual wages calculated at the municipality level. As expected, wages are higher in denser areas like Tokyo and Osaka. We find a correlation of log wages with log land values of 0.46, which we think is reasonable.

A.3 Undevelopable land

We construct a new measure of developable land for Japan based on land use maps and elevation data. We obtain information on lakes and water bodies from *OpenStreetMap*. Using these data, we only keep land in each municipality. Furthermore, we obtain information on elevation from the *AlosWorld3D* project, which provides elevation at a 30m by 30m resolution. We calculate the slopes of each grid cell and remove all grid cells that have slopes above 50% as Saiz (2010) shows that these are essentially undevelopable. Furthermore, we remove all land that is above 2000m above sea level, for which it is unlikely that there are permanent settlements.

Figure A.3 shows a map of undevelopable land for Japan. One may observe that the large cities (Tokyo, Osaka, Nagoya) are located in relatively flat areas with high shares of developable land. However, most of inland Japan is severely geographically constrained: we find that 20% of Japan is undevelopable. Still, in some municipalities, the share of undevelopable land is much higher and can be as high as 80% (*e.g.*, in mountainous areas in Central Japan). The share of 80% developable land is considerably higher than other estimates of 'inhabitable land', which would be 33% according to the *Social Indicators* by Prefecture from the *Statistics Bureau*. Inhabitable land excludes forests and does not use information on slopes.

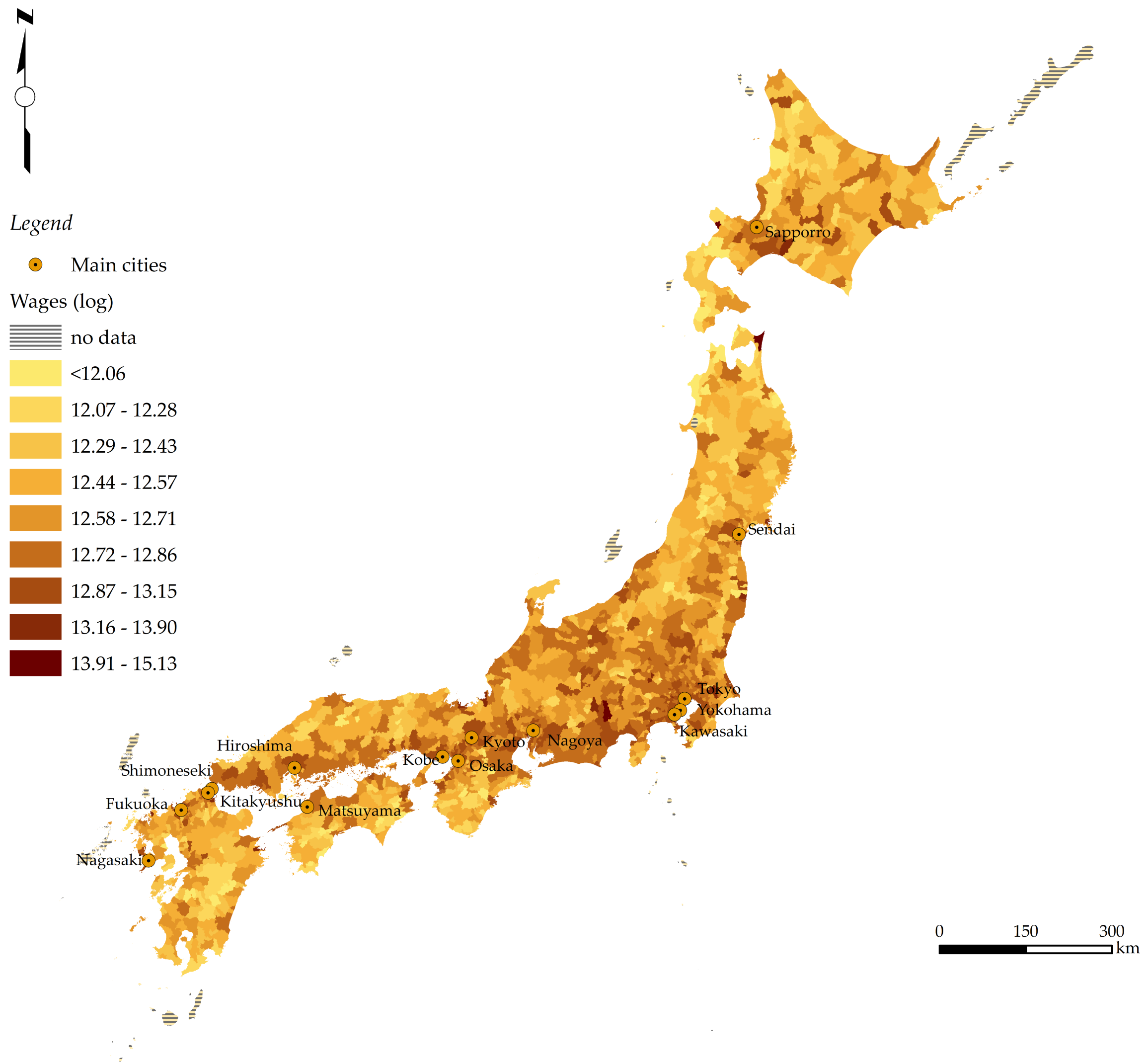


Figure A.2 — Annual wages 2001-2011 at the municipality level

However, forests are technically developable, although they are often protected. Furthermore, although slopes exceeding 20% could be developed, it is very costly to do so. Hence, we think our estimate of undevelopable land is best interpreted as a lower bound estimate of the amount of undevelopable land in Japan.[22] We use historical road networks and historical infrastructure plans. First, we have a hardcopy map of the *Seven-Circuit Road Network* in the 7$^{\text{th}}$ century.[23]

[22] Note that we use the amount of developable land area to calculate employment and population densities. If we were to use inhabitable land, this means that we would find high densities in remote areas with large patches of forests. We think it is unlikely that these areas benefit strongly from agglomeration economies. This provides another argument to use our measure of developable land.

[23] As for the route, see the map available at `https://www.mlit.go.jp/road/michi-re/1-1.htm`.

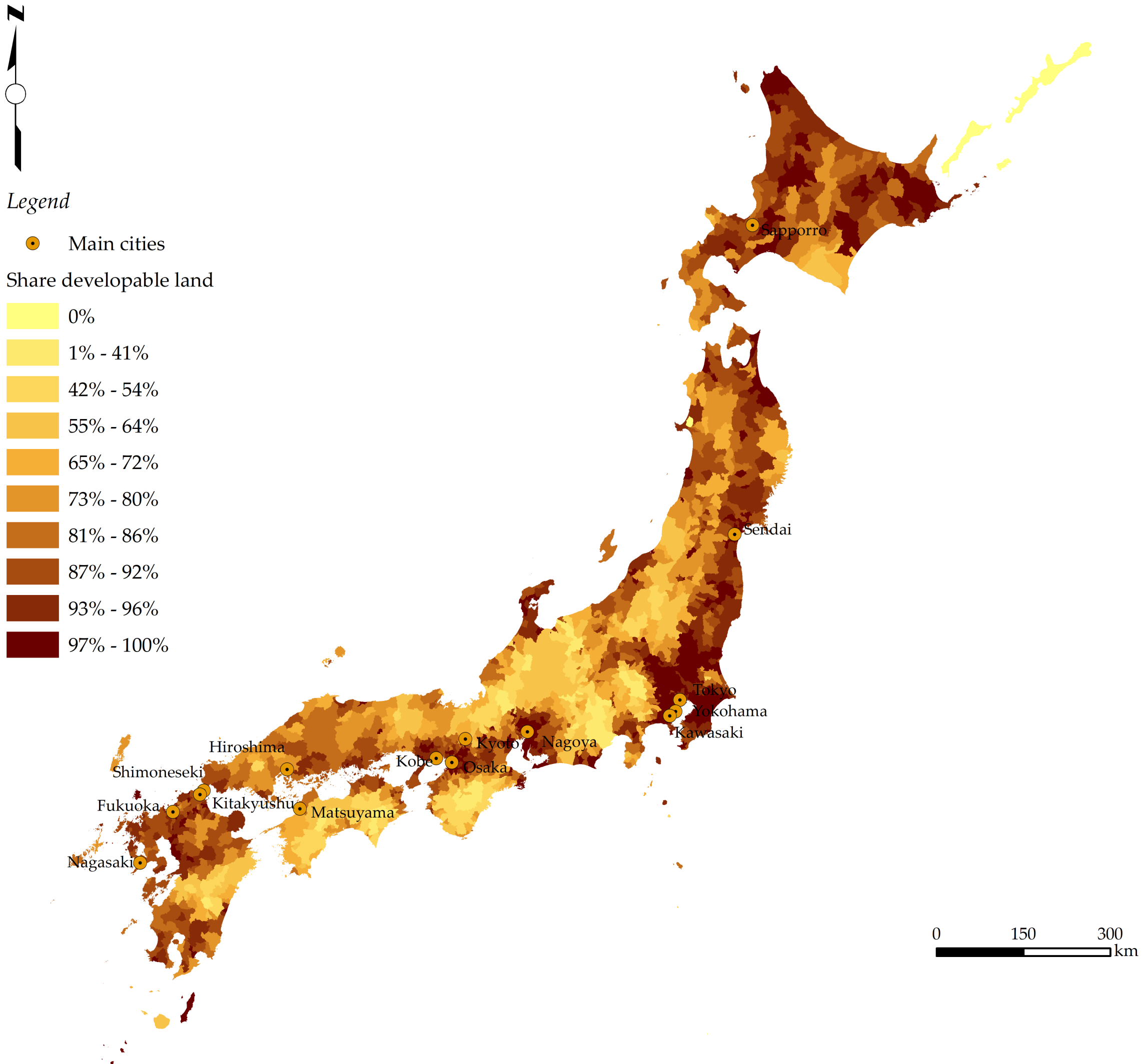


FIGURE A.3 — DEVELOPABLE LAND IN JAPAN

This road network was developed to connect the initial capital cities (*i.e.*, Nara and Kyoto) with many other cities in Japan. Second, we use hardcopy maps of the *National Road Plan* developed by the *Home Ministry* in 1943. The total length of highways was planned to be 5,490 km.[24]

A.4 Historical data

The planned network was motivated by the transport of military supplies. Third, we refer to the actual routes of roads and railways in 1900 obtained from the *National Land Numerical Information*. We manually georeference these historic transport maps to be able to link the data to current municipal data.

[24] As for the route, see the map available at `https://www.mlit.go.jp/road/michi-re/4-2.htm`.

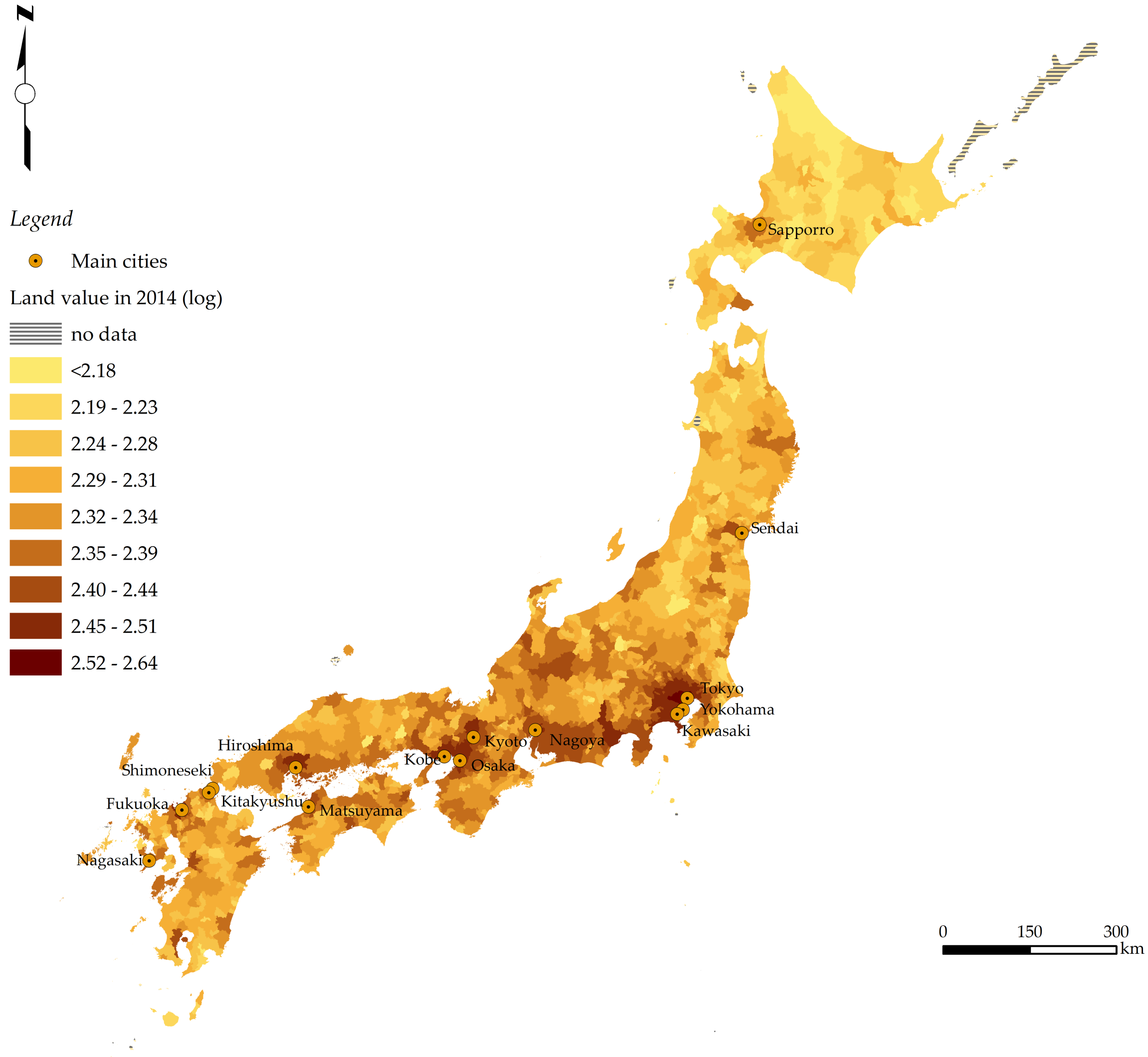


Figure A.4 — Land values in 2014 at the municipality level

The data on the local population, except for Hokkaido in 900, are taken from Kito (1996). Kito estimated the number of provincial population by using the information on the area of rice fields, which is available in *Wamyo Ruijusho*, a Japanese dictionary completed in 938. Although the estimates of population are available for 68 provinces in 900, the provincial boundaries are obviously different from current municipal boundaries. We address this issue by distributing the population in each province according to the share of land of each municipality in the corresponding province. By using the information on the number of archaeological remains, Takada (2017) estimated the population size in Hokkaido around the 9^{th} century as $37,000$. This number is distributed to each municipality in Hokkaido based on its land share. The local population in 1872, including that of Hokkaido, is obtained from Kito (1996), which is based on the *National Table on Family Registration* compiled by the *Home Ministry*.

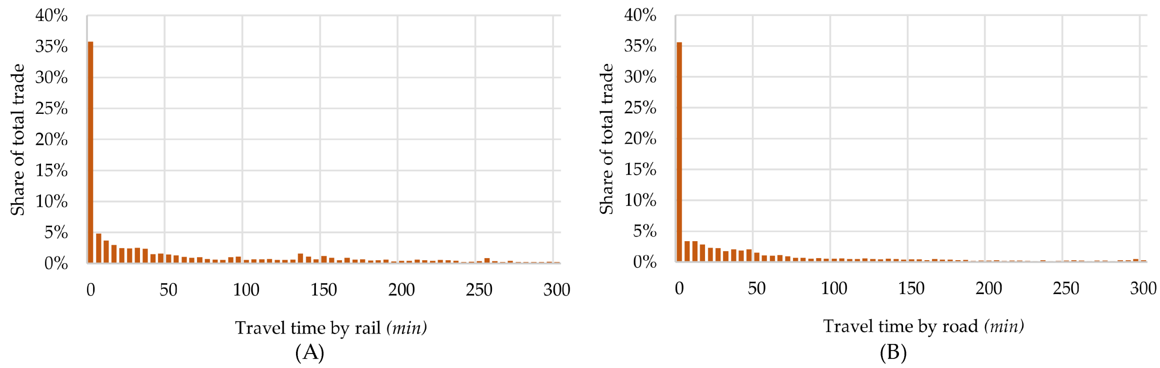


Figure 2: Figure A.6 — Histograms of travel time by trade value

A.5 Detailed descriptives

In Figure A.4 we show that land values are considerably higher in metropolitan areas like Tokyo, Osaka, and Nagoya.

In Figure A.5 we display a histogram of the commuting time distribution. Essentially, we count the share of people with a certain commuting time. In the left panel we show that about 5% of the commuting flows are within the same municipality. Then, we see a right-skewed distribution with essentially all commuting being within 120 minutes of traveling. The pattern for road travel time looks similar (see the right panel), although the travel time by road is generally somewhat longer. Hence, essentially all commutes are now within 240 minutes.

In Figure A.6 histograms related to travel time between firms are shown, implying that we also take into account the Shinkansen and highways. In the left panel we observe that 35% of the trade value occurs within the same municipality. Essentially all trade occurs within 5 hours of traveling by train or road. The histogram for travel time by road in the right panel looks similar, although it is somewhat smoother.

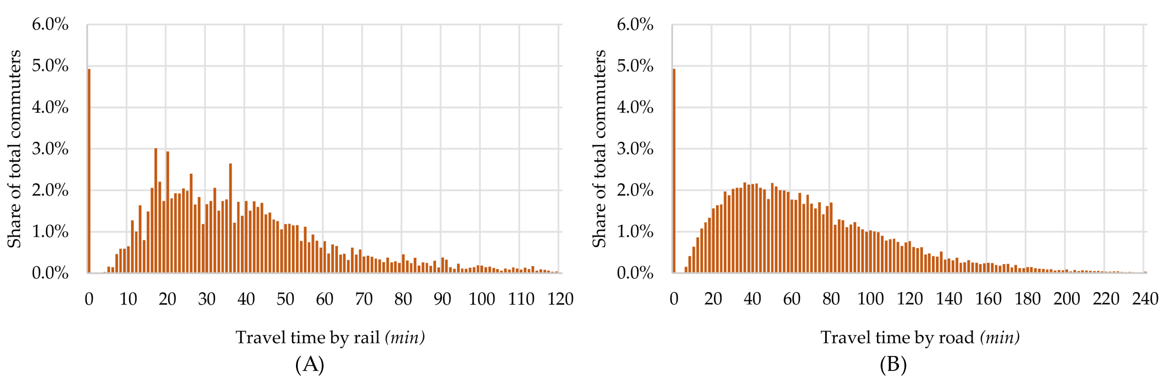


Figure A.5 — Histograms of commuting time by number of commuters

Table A.2 — Descriptive statistics for travel flow data

	(1) mean	(2) sd	(3) min	(4) max
Travel flow by rail	11,008	191,659	0	8,223,689
Travel flow by road	2,864.5	35,036	0	1,116,832
Travel flow by airplane	35.070	274.76	0	6,360.4
Travel flow by rail, selected	528.54	6,605.2	0	205,934
Travel flow by road, selected	58.417	229.43	0	4,988.2
Travel flow by airplane, selected	35.579	279.30	0	6,360.4

Notes: The data are from travel flows between prefectures in 2010 and 2015. The number of observations is $4,232$. For the selected flows, we exclude prefectures for which the centroid is within 75km of each other as well as within-prefecture flows to exclude most commuting flows. We then have $4,068$ observations.

A.6 Travel flows and trade in intermediate services

At the heart of our model contains a gravity model of intermediate services trade. Equation (26) shows that the expenditure share of the final sector on intermediate services from location i depends on the travel time by rail and road, which suggests that intermediate services trade may require business trips.

Business trips are journeys that employees take for work-related purposes. They can include visiting clients, attending meetings, conducting market research, or participating in conferences. Business trips often involve face-to-face interactions, as they require employees to engage with other people in order to achieve their objectives. We hypothesize that trade of intermediate service often requires face-to-face interactions and therefore a higher flow of services will imply a higher flow of business trips.

Here, we aim to provide more evidence that trade of intermediate services is indeed related to the number of business trips between two locations. For this we obtain data on travel flows by travel mode between *prefectures* for 2010 and 2015.[25] There are 46 prefectures in Japan so these are considerably larger than municipalities. Table A.2 reports descriptive statistics for the travel flows.. Travel flows include commuting flows, shopping trips and business-to-business journeys. It may seem that travel flows by train are overrepresented but that is because the travel flow by road is only taking into account cars and trucks owned by firms. We may try to isolate business trips by excluding within-prefecture flows as well as flows that have a destination within 75km of the prefecture's origin. Indeed, this reduces the travel flows by rail and road by more than 95%.

To further investigate the relationship between travel flows and trade of intermediate services

[25] These are the years for which we have data on trade of intermediate services *and* commuting flows. Please note that the measurement of travel flows by car is somewhat different before 2010. From 2010 onwards it includes only the travel by cars and trucks owned by firms.

we estimate simple regressions of the following form:

$$\log F_{ijy} = \beta_0 + \beta_1 \log S_{ijy} + \beta_2 \log n_{ijy} + \lambda_y + \epsilon_{jy}.$$

where F_{ijy} is the travel flow between i and j in year y, S_{ijy} is the value of intermediate services trade, n_{ijy} are the number of commuters, and λ_y denotes year fixed effects. Because the flow is sometimes zero (in about 10% of the cases for the total flow), we estimate the above equation by PPML. Further, S_{ijy} and/or n_{ijy} can also be zero, we use the inverse-hyperbolic sine transformation, which is approximately the same as a log transformation. Hence:

$$F_{ijy} = \exp(\beta_0 + \beta_1 \operatorname{arsinh} S_{ijy} + \beta_2 \operatorname{arsinh} n_{ijy} + \lambda_y).$$

We report the main results in Table A.3. In the first four columns we control for the number of commuters. Column (1) shows that the elasticity of intermediate services trade is 0.763. The elasticity of travel flows with respect to the number of commuters is considerably lower (0.483). Moreover, we can explain almost 95% of the variation in travel flows by just these two variables confirming that intermediate services trade and commuting flows are important determinants of travel flows. In the remaining columns we distinguish between three different travel modes: rail, road and air. Column (2) shows that the elasticity with respect to becomes stronger (0.855) once we focus on trips by rail, while the elasticity is low and statistically insignificant for travel by road (column (3)). This suggest that business trips are particularly made by train rather than by road, which is in line with descriptive statistics reported earlier (i.e., that the share of train travel in 2010 is 43.7% for trips between 300 and 500km, while it reaches almost 70% for trips between 500 and 700km). As expected, for air travel in column (4) we also find a strong trade elasticity of about 0.8 but we note that the flows in absolute numbers are much lower for air travel as compared to travel by rail.

One may be worried that simply controlling for commuting flows may not be sufficient as the relationship between travel flows and commuting flows may be nonlinear, for example because people living further away from work may commute fewer times a week. In columns (5)-(8) in Table A.3 we therefore take an alternate approach by excluding within-prefecture travel as well as flows that have a destination within 75km of the prefecture's origin. As shown in Table A.2 this removes 95% of the local flows, likely related to commuting and shops. We then show that the estimated elasticities are somewhat lower. Still, we find that the elasticity of intermediate services trade is about twice as large for travel by rail as compared to travel by road, which confirms the earlier finding that business travel is likely done mostly by rail.

To the extent one may be considered that the treatment of zero trade or commuters impact

Table A.3 — Travel flows, trade of intermediate services and commuting

(Dependent variable: the travel flow between prefecture i and j)

	Control for commuting flows				Exclude destination locations <100km of origin			
	All	Rail	Road	Air	All	Rail	Road	Air
	(1) PPML	(2) PPML	(3) PPML	(4) PPML	(5) PPML	(6) PPML	(7) PPML	(8) PPML
Trade of intermediate services *(ihs)*	0.7634***	0.8548***	0.2172	0.7915***	0.6289***	0.6883***	0.3418***	0.6259***
	(0.1039)	(0.1068)	(0.1384)	(0.1209)	(0.0903)	(0.1090)	(0.0138)	(0.0348)
Number of commuters *(ihs)*	0.4826***	0.4247***	1.0471***	-0.3142***				
	(0.1106)	(0.1135)	(0.2073)	(0.0781)				
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,232	4,232	4,232	4,232	4,068	4,068	4,068	4,068
Log-likelihood	-16,126,244	-17,525,255	-2,057,991	-271,693	-3,981,932	-3,959,871	-344,546	-251,354
Pseudo-R^2	0.940	0.921	0.961	0.455	0.510	0.492	0.343	0.483

Notes: Our explanatory variables are transformed using the inverse-hyperbolic sine transformation (arsinh). Standard errors are clustered by prefecture origin and destination in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

the results, we replicate the results but now take the log of S_{ijy}, while excluding pairs for which $S_{ijy} = 0$. Furthermore, because n_{ijy} is often zero for prefectures that are far apart, we control for a 5$^{\text{th}}$-order polynomial instead of taking the logarithm of n_{ijy} The results reported in Table A.4 are very much in the same ballpark and support the above conclusions.

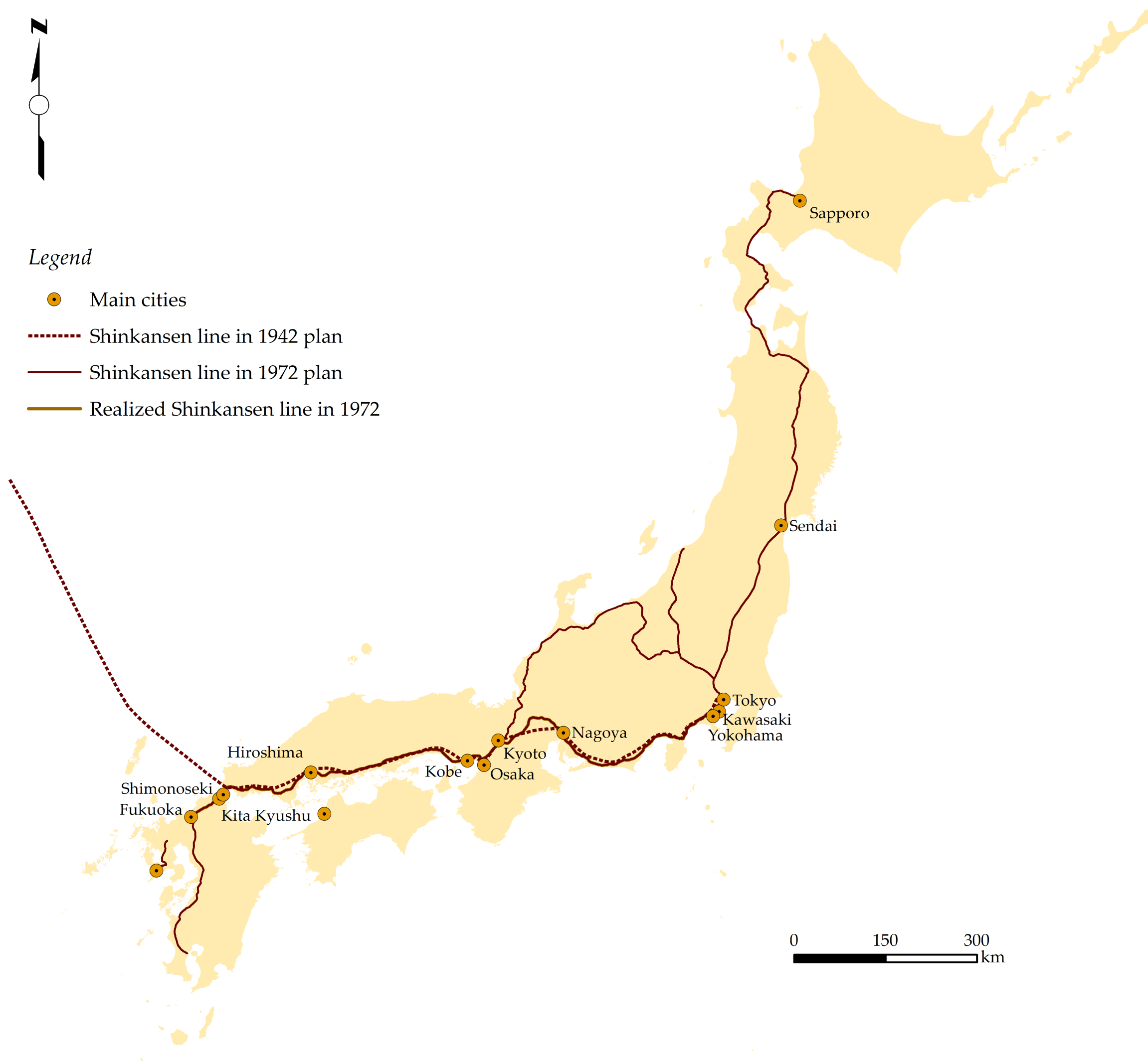


Figure A.7 — Overview of planned Shinkansen routes

A.7 Historic Shinkansen plans

We report maps of the planned Shinkansen network in Figure A.7. The idea behind the 1942 plan was to link Tokyo to Shimonoseki and even further to Beijing. The 1972 plan included most Shinkansen lines, except the planned Chuo Shinkansen between Tokyo, Nagoya, and Osaka, as well as the 'Mini'-Shinkansen lines.

Table A.4 — Travel flows and trade of intermediate services (log-transformation)
(Dependent variable: the travel flow between prefecture i and j)

	Control for commuting flows				Exclude destination locations <100km of origin			
	All	Rail	Road	Air	All	Rail	Road	Air
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	PPML	PPML	PPML	PPML	PPML	PPML	PPML	PPML
Trade of intermediate services *(log)*	0.6405***	0.7227***	0.1040	0.8320***	0.7514***	0.7758***	0.4974***	0.7930***
	(0.2297)	(0.2432)	(0.1298)	(0.0713)	(0.1183)	(0.1279)	(0.0703)	(0.0142)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of commuters, 5^{th}-order polynomial	Yes	Yes	Yes	Yes	No	No	No	No
Number of observations	1,471	1,471	1,471	1,471	1,313	1,313	1,313	1,313
Log-likelihood	-15,415,077	-16,917,544	-1,307,782	-208,855	-3,594,634	-3,729,968	-219,507	-202,951
Pseudo-R^2	0.926	0.903	0.967	0.403	0.365	0.337	0.236	0.389

Notes: Standard errors are clustered by prefecture origin and destination in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

Appendix B. Reduced-form results

B.1 Event studies

To provide some additional support for a causal interpretation of the positive effect of the opening of a Shinkansen station on average travel times and employment densities we undertake an event-study analysis:

$$\left\{\log \overline{T}_{\mathcal{R},jy}, \log \frac{M_{jy}}{\mathcal{A}_i}\right\} = \beta_0 + \sum_{\tau=-35}^{45} \beta_{\tau 1} \mathcal{S}_{j,y-\tau} + \beta_2 \mathcal{X}_{jy} + \lambda_j + \lambda_y + \epsilon_{jy}.$$

The coefficients $\beta_{\tau 1}$ are then dependent on 10-year windows before and after the opening of a Shinkansen, denoted by τ. We report the results in Figure B.1.

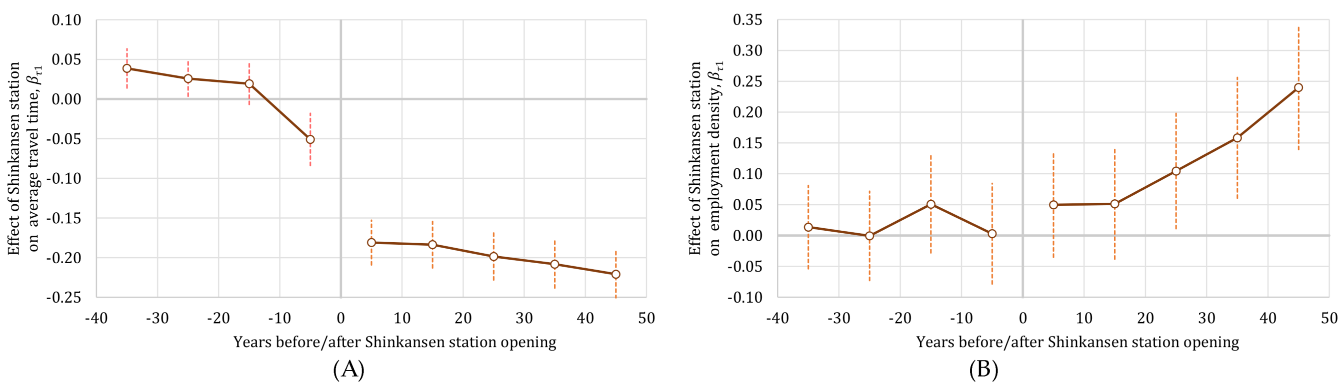


FIGURE B.1 — REDUCED-FORM RESULTS: EVENT STUDIES

In Panel A, we take average travel times as dependent variable. We show that the effect of a Shinkansen station on average travel times is already adjusting a few years before a Shinkansen station is opened. We think this is possible because other railway connections were already improved before the opening of the Shinkansen station.

In Panel B, we do not find evidence for pre-trends in employment density. After the opening of a station, the employment density increases by about 5% but this appears to be a short-run effect. After 45 years, the effect of a Shinkansen station on employment density has increased to about 25%.

B.2 First-stage results

Here we report first-stage results. In the first column of Table B.1 we show that a planned station in 1972 increases the probability of getting a Shinkansen station. For each 10-year increase, the probability increases by 10.2 percentage points. The effect of a planned station in the 1942 plan

Table B.1 — Reduced-form estimates: first-stage results
(Dependent variable: Shinkansen station <10km)

	Baseline specification		Intermediate places			
	Instrument: 1972 plan×year	Instrument: 1942 plan×year	Instrument: 1972 plan×year		Instrument: 1942 plan×year	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Planned Shinkansen station <10km×year of observation	0.0102***		0.0077***	0.0093***		
	(0.0004)		(0.0005)	(0.0005)		
Planned Shinkansen station in 1942 <10km×year of observation		0.0086***			0.0058***	0.0063***
		(0.0009)			(0.0010)	(0.0010)
Railway station <10km	0.0182***	0.0197***	0.2725***	0.1304**	0.2659***	0.1306**
	(0.0045)	(0.0053)	(0.0455)	(0.0523)	(0.0512)	(0.0542)
Highway <10km	0.0352***	0.0547***	0.0626***	0.0085	0.0823***	0.0188
	(0.0039)	(0.0044)	(0.0134)	(0.0142)	(0.0139)	(0.0160)
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Preference×year fixed effects	No	No	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	18,678	18,678	4,554	4,543	4,554	4,543
R^2	0.8187	0.7840	0.8012	0.8731	0.7833	0.8483

Notes: In columns (3)-(6) we exclude municipalities that are centres of metropolitan or so-called micropolitan areas. We further exclude areas that are further than 10km of a current or future Shinkansen line. Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

increases the probability by 8.6% for each 10-year increase. Both instruments are strong, although the 1972 plan is inevitably somewhat stronger.

In the analyses where we make selections on what municipalities we include (such as those close to current and future Shinkansen lines) the first-stage results are not materially influenced. We discuss the second-stage results of this approach in Appendix B.4.

B.3 Heterogeneous effects

The impact of a Shinkansen station on accessibility and employment density likely depends on where the Shinkansen station is opened. If the station is opened in an area that is already very well connected, effects may be smaller than, say, in a peripheral area that was poorly connected. We therefore consider estimating semiparametric regressions of the form:

$$\left\{\log \overline{T}_{\mathcal{R},jy}, \log \frac{M_{jy}}{\mathcal{A}_j}\right\} = f\left(\mathcal{S}_{jy}, \mathcal{X}_{jy}\right) + \lambda_j + \lambda_y + \epsilon_{jy}, \tag{37}$$

where $f(\cdot)$ is a nonparametric function of $\mathcal{S}_{jy}$ and $\mathcal{X}_{jy}$, which is dependent on the employment density in 1957. We assume $f\left(\mathcal{S}_{jy}, \mathcal{X}_{jy}\right) = \beta_{j1}\mathcal{S}_{jy} + \beta_{j2}\mathcal{X}_{jy}$. We estimate $f(\cdot)$ by local linear estimation techniques, implying that for each unique value of employment density in 1957 we obtain values for β_{j1} and β_{j2}.[26] To estimate λ_j and λ_y we apply Robinson's (1983) estimation procedure.[27]

Figure B.2 shows the results. Panel A suggests that the heterogeneity in the effect of a Shinkansen station on average travel time is small, although we observe that travel time reductions are somewhat larger in dense areas in 1957. By contrast, Panel B, where we test the impact of a Shinkansen station on employment density, shows considerable heterogeneity in the effect. We find that the effect of a Shinkansen station is positive and statistically significant below a density of about 250 workers per ha, which is about the employment-weighted mean employment density in 1957. However, for higher densities the effect turns out to be negative. For example, for a density of $\exp(10) = 36,316$ workers per ha, employment density is reduced by 25% when a Shinkansen station is opened. What we learn here is that the effect of a Shinkansen connection is likely very heterogeneous.

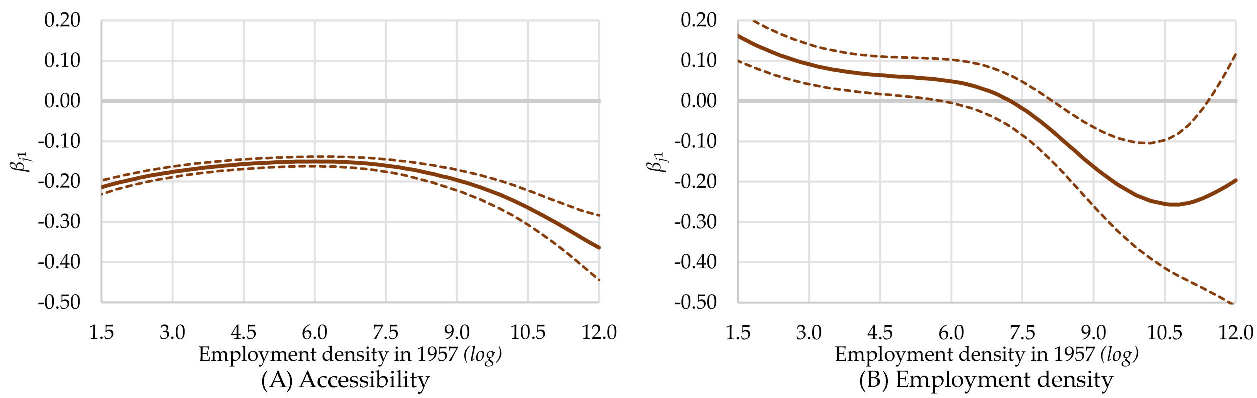


FIGURE B.2 — REDUCED-FORM RESULTS: HETEROGENEOUS RESULTS

Let us also consider to instrument for the Shinkansen station dummy. Because a non-linear IV estimation will yield invalid results, we adopt a control function approach. This implies that we estimate the following first stage:

$$\mathcal{S}_{jy} = \tilde{f}\left(\mathcal{I}_{jy}, \mathcal{X}_{jy}\right) + \tilde{\lambda}_j + \tilde{\lambda}_y + \tilde{\epsilon}_{jy},$$

where $\mathcal{I}_{jy}$ are the instruments as discussed in Section 2.4.3. We estimate the above specification using the same approach as in equation (37). We then insert the estimated first-stage residuals as a control function in (37). We report the results in Figure B.3. We show in Panel A that the

[26] We use a Gaussian kernel and set the bandwidth equal to 1. The results are pretty insensitive to these choices.

[27] This procedure regresses $\log \overline{tt}_{\mathcal{R},jy}$, $\log M_{jy}/\mathcal{A}_i$ and municipality and year dummies non-parametrically on $\mathcal{S}_{jy}$ and $\mathcal{X}_{jy}$ using local linear regressions. We then obtain residuals for the dependent variables ($\log \overline{tt}_{\mathcal{R},jy}$, $\log M_{jy}/\mathcal{A}_i$) and municipality and year dummies. The residuals of the dependent variables are then regressed on the dummy residuals using OLS, which identifies λ_j and λ_y. The final step is to replace the dependent variables by $\log \overline{tt}_{\mathcal{R},jy} - \hat{\lambda}_j - \hat{\lambda}_y$ and $\log M_{jy}/\mathcal{A}_i - \hat{\lambda}_j - \hat{\lambda}_y$ and regress these residuals non-parametrically on $\mathcal{S}_{jy}$ and $\mathcal{X}_{jy}$ to obtain the coefficients of interest: β_{j1} and β_{j2}.

effect on average travel times is comparable to the local linear estimates where we do not correct for endogeneity.

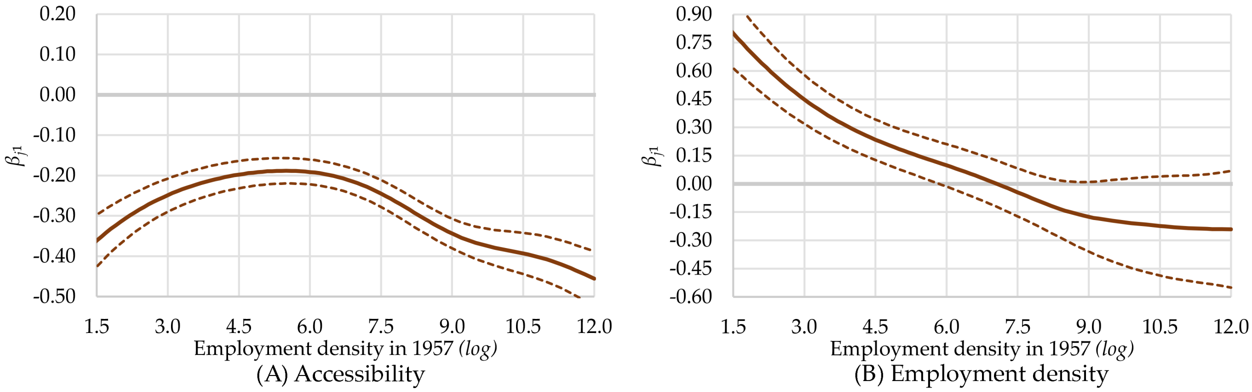


Figure B.3 — Reduced-form results: heterogeneous results, control function

The effects of a Shinkansen station on employment density are larger. For example, for low densities in 1957, the coefficients imply increases in employment density exceeding 50%. Interestingly, we still find negative effects of about 25% for municipalities that had a high employment density in 1957. The effects are, however, somewhat imprecise. Still, the overall pattern is the same as in Figure B.2: areas with a low employment density benefit considerably more from a Shinkansen station than areas with a higher employment density.

B.4 Intermediate places

Koster *et al.* (2022) are concerned with the analysis of the effects of Shinkansen stations openings on so-called intermediate places, which are localities that were not the main reason for constructing a Shinkansen line but happened to be in between two larger cities. For example, while Toyohashi with a population of 300 thousand was too small of a city to attract a dedicated Shinkansen line, the fact that it is located in between Tokyo and Nagoya meant that a Shinkansen station was opened in Toyohashi. They then find that a Shinkansen station leads to sizable reductions in employment density ranging from $10 - 40\%$.

The aim of this paper is different as we aim to study the overall effects of the Shinkansen network so we are interested in the effects beyond intermediate places. Still, we aim to show here that we can replicate their findings although we find that the average effect across all municipalities is positive (see Table 4). Following Koster *et al.* (2022), we use the definition from 2015 of Metropolitan Employment Areas (MEAs) and Micropolitan Employment Areas (McEAs) provided by Kanemoto and Tokuoka (2002). In our final sample, there are 100 MEAs and 117 McEAs. We exclude the so-called 'central municipalities' that belong to the MEAs and McEAs because they are likely the reason for a Shinkansen line to have been constructed. Further, we

TABLE B.2 — REDUCED-FORM ESTIMATES: INTERMEDIATE PLACES
(Dependent variable: the log of employment density)

	Baseline specifications		Instrument: 1972 plan×year		Instrument: 1942 plan×year	
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Shinkansen station <10km	-0.1129***	-0.1168***	**-0.3096*****	**-0.1841***	**-0.3739**	**-0.8298*****
	(0.0347)	(0.0422)	(**0.1199**)	(**0.0943**)	(**0.2513**)	(**0.2629**)
Railway station <10km	0.0177	-0.0708	0.0692	-0.0620	0.0861	0.0223
	(0.0531)	(0.0536)	(0.0626)	(0.0550)	(0.0854)	(0.0769)
Highway <10km	0.0981***	0.1253***	0.1129***	0.1265***	0.1177***	0.1380***
	(0.0199)	(0.0291)	(0.0221)	(0.0292)	(0.0261)	(0.0316)
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Preference×year fixed effects	No	Yes	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,554	4,543	4,554	4,543	4,554	4,543
R^2	0.9696	0.9760				
Kleibergen-Paap F-statistic			198.6	376	32.21	38.38

Notes: **Bold** indicates instrumented. In columns (3) and (4) we instrument Shinkansen <10km with an interaction term of whether the municipality is within 10km of a planned Shinkansen station with the year of observation. In columns (5) and (6) we use a dummy indicating that a municipality is within 10km of the planned line in 1942 interacted with the year of observation. We exclude municipalities that are centres of metropolitan or so-called micropolitan areas. We further exclude areas that are further than 10km of a current or future Shinkansen line. Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

only include municipalities that are within 10 kilometers of a current or future Shinkansen line. In this way, we compare places that are arguably similar in unobservables, except that one receives a Shinkansen station at some point. We report the results in Table B.2.

Columns (1) and (2) are OLS estimates. We show that a Shinkansen station now changes employment density by $(\exp(-0.1129) - 1) \times 100\% = -10.7\%$ so that connected intermediate places seem to lose employment. The results are essentially identical if we control for regional trends by including prefecture-by-year fixed effects (see column (2)).

In columns (3) and (4) we instrument for the Shinkansen station dummy with the 1972 plan interacted with year. Recall that the first-stage results indicated that the instrument is sufficiently strong. The effects are now somewhat larger: a Shinkansen station seems to reduce employment by 27%. The effect is somewhat smaller if we include prefecture-by-year fixed effects in column (4), albeit imprecise.

Columns (5) and (6) use the 1942 plan×year instrument. The instrument is then somewhat weaker and the second-stage results show large standard errors. Still, the point estimates are negative and sizable, although we think the estimate in column (6) is on the high side with a reduction in employment density exceeding 50%.

We reemphasize that these results are not contradicting the positive results shown in Section 2.4 because the intermediate places approach has a different control group: peri-urban areas that are close to Shinkansen lines but do not have a Shinkansen station. Moreover, the analyses where we allow for heterogeneous effects of Shinkansen stations showed that negative effects on employment density can occur for places with an above-mean employment density in 1957.

Appendix C. Structural estimation

C.1 A gravity model for commuting trips

We report results for different specifications of the commuting gravity equation in Table C.1. We begin by taking into account all location pairs that are within two hours traveling. We find that a ten-minute travel time increase by train changes the number of commuters by $(\exp(-0.0678 \times 10) - 1) \times 100\% = -49\%$. The impact of road infrastructure is similar but somewhat smaller: a ten-minute increase in travel time by car decreases the number of commuters by 41%. The observation that travel time elasticities are similar is in line with the observation that the modal split between train and car in Japan is similar. More specifically, according to the *Nationwide Person Trip Survey* by the *Ministry of Land, Infrastructure, Transport and Tourism,* $20 - 50\%$ of the commutes are by car, while about 20% of the commutes is by train. If we include either travel time by rail *or* travel time by road, we find elasticities of respectively 0.089 and 0.064, which are similar to the literature (see, *e.g.*, Ahlfeldt *et al.*, 2015).

The finding that the travel time elasticity for train travel time is stronger may be related to the lack of comfort in heavily congested commuter trains in large metropolitan areas like Tokyo and Osaka. In the remaining columns of Table C.1 we provide robustness checks for the baseline effects. First, in column (2) we focus on locations within 75km of each other, as commutes over longer distances may be considered unlikely. This reduces the observations by about 10%, but the results are not much affected. Hence, excluding location pairs that are far from each other does not affect the results. One may also be concerned that a high share of zero flows thwarts our estimates. This does not seem to be the case: when we exclude zero flows in column (3), the coefficients are very similar as compared to column (1).

In the remaining columns we address the issue of possible reverse causality. In column (4) we keep location pairs that were connected in a 1942 highway and the initial plan for a high-speed railway line to connect Tokyo to Shimonoseki and further to Beijing. This railway line was mainly to lower the costs of transporting passengers and cargo. Similarly, we obtained data on the 1942 National Highway Plan. We observe that this does not impact much our results, even though we only keep about 10% of the location pairs. In column (5) we use the road network in the 8^{th} century, which was centered around the capital Nara. We keep about 14% of the location pairs that were connected by a road in the 8^{th} century, but the results are once again very similar.

Column (6) studies whether the travel time elasticities depend on the relative travel times of both transport modes. One may argue that if train travel times are low compared to travel by road, this may attract more commuters, which in turn increases congestion and lowers comfort levels. By contrast, if road travel time is low relative to train travel times, more commuters may

TABLE C.1 — COMMUTING GRAVITY MODELS

(Dependent variable: the number of links)

	Baseline PPML	Locations <75km	Commuters >0	Connected in 1942 plan	Connected in 8^{th} century	Travel time ratio	Add controls
	(1) PPML	(2) PPML	(3) PPML	(4) PPML	(5) PPML	(6) PPML	(7) PPML
Travel time by train *(min)*, $\hat{\varkappa}_{\mathcal{R}}$	-0.0678***	-0.0806***	-0.0650***	-0.0683***	-0.0701***	-0.0884***	-0.0676***
	(0.0005)	(0.0007)	(0.0005)	(0.0011)	(0.0009)	(0.0008)	(0.0005)
Travel time by road *(min)*, $\hat{\varkappa}_{\mathcal{H}}$	-0.0531***	-0.0580***	-0.0529***	-0.0581***	-0.0601***	-0.0766***	-0.0528***
	(0.0011)	(0.0008)	(0.0009)	(0.0024)	(0.0020)	(0.0040)	(0.0011)
Travel time by train *(min)*/ Travel time by road *(min)*						0.0314***	
						(0.0009)	
East-west 'border'							-0.9758***
							(0.0625)
On same island							0.1128
							(0.1338)
Residential location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Workplace location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	476,625	422,921	326,342	46,558	357,317	476,625	476,625
Log-likelihood	-21,461,627	-6,409,245	-20,427,138	-2,412,660	-17,695,186	-21,461,627	-21,368,198
Pseudo-R^2	0.976	0.970	0.975	0.993	0.978	0.976	0.976

Notes: We use the number of commuters between municipalities as dependent variable. Standard errors are bootstrapped (250 replications) by sellers locations and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

be inclined to take the car leading to traffic congestion and a stronger travel time elasticity $\varkappa_{\mathcal{H}}$. We therefore include the ratio of travel time by rail and travel time by road as an additional control variable. We find that a standard deviation increase in the commuting time ratio implies an increase in the railway commuting time elasticity of $0.218 \times 0.0314 = 0.00685$, which is about 7.5% of the main effect. Hence, despite the effect of the commuting time ratio being statistically significant, the difference is economically negligible. Further, even though the commuting time elasticities may slightly differ, in practice the relative ratio of commuters' preferences for the train and car is not so important because we focus on the Shinkansen, which is hardly used by commuters. As a result, the relative ratio of commuting times by train and car is not affected in our counterfactual scenarios.

In the final column of Table C.1 we add two controls, one capturing whether a location pair has to cross the east-west 'border' as defined by Wrona (2018), and a dummy indicating whether the location pair is on the same island. We find that the travel time elasticities are virtually the same as the baseline specification. Interestingly, we find a strong negative effect of the east-west 'border' suggesting that fewer people commute between east and west Japan. However, this may also capture the fact that people do not commute much between prefectures. We find a positive coefficient of being on the same island, which makes sense, but note that the coefficient is rather imprecise. In any case, it is reassuring that the travel time elasticities are not materially influenced when including these controls.

C.2 A gravity model for business trips

In Table C.2 we report gravity models for final firms' input networks. The dependent variable is then the normalized trade value between the final and intermediate sectors for a given location pair, which follows directly from our model. Column (1) is the baseline specification, which shows a strongly significant railway travel time elasticity, suggesting that the trade between final and intermediate firms reduces by 2.0% for a 10-minute increase in railway travel time. This travel time elasticity is about 3% of the commuting time elasticity, which makes sense. Business trips are likely much less frequent so the travel time elasticity is expected to be much lower.

We further find that trade between final and intermediate firms reduces by 24% for a 10-minute increase in travel time by road. These results suggest that travel time costs by train are considerably lower than travel time by road. We think this makes sense as Japanese high-speed trains offer high levels of comfort so that work-related activities can easily be undertaken while traveling. When traveling by car this is considerably harder. In 2017, the total distance of domestic passenger transport in Japan amounted to approximately 605 billion passenger kilometers, with railway transport accounting for 72.3% of the transport distance, while motor vehicles only

TABLE C.2 — GRAVITY MODELS OF PRODUCTION NETWORKS
(Dependent variable: the value of intermediate trade linkages)

	Baseline PPML	All firms	Including air travel	Connected in 1942 plan	Connected in 8^{th} century	Connected by Shinkansen	Travel time ratio	Add controls
	(1) PPML	(2) PPML	(3) PPML	(4) PPML	(5) PPML	(6) PPML	(7) PPML	(8) PPML
Travel time by train *(min)*, $\hat{\varkappa}_{\mathcal{R}}$	-0.0020***	-0.0023***	-0.0014***	-0.0024***	-0.0053***	-0.0053***	-0.0009***	-0.0015***
	(0.0002)	(0.0001)	(0.0002)	(0.0006)	(0.0004)	(0.0004)	(0.0002)	(0.0001)
Travel time by road *(min)*, $\hat{\varkappa}_{\mathcal{H}}$	-0.0242***	-0.0164***	-0.0395***	-0.0279***	-0.0446***	-0.0446***	-0.0237***	-0.0247***
	(0.0008)	(0.0004)	(0.0015)	(0.0061)	(0.0051)	(0.0051)	(0.0008)	(0.0007)
Travel time by air *(min)*, $\hat{\varkappa}_{\mathcal{A}}$			-0.0419***					
			(0.0015)					
Travel time by train *(min)*/ Travel time by road *(min)*							-0.0015***	
							(0.0003)	
East-west 'border'								-0.2612***
								(0.0409)
On same island								0.4143***
								(0.0547)
Residential location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Workplace location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1,214,668	1,929,565	1,214,668	72,332	103,040	155,499	1,214,668	1,214,668
Log-likelihood	-21,705,246	-64,267,914	-13,640,402	-3,678,054	-18,340,792	-3,940,105	-6,297,798	-21,670,425
Pseudo-R^2	0.899	0.933	0.918	0.951	0.906	0.951	0.936	0.899

Notes: We use the number of commuters between municipalities as dependent variable. In column (6), municipalities that are within 25km of a Shinkansen station are considered to be connected to the Shinkansen network. Standard errors are bootstrapped (250 replications) by sellers locations and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

account for 11.3% (and air travel for 16.4%). Hence, for long-distance travel, the train is by far the most preferred transport mode, which is in line with the finding that the travel time elasticity by train is considerably lower than the travel time by car. If we include only travel time by road like in Monte *et al.* (2018), we find an elasticity of about -0.021, while for travel time by rail it is -0.011. Both elasticities are comparable to what Monte *et al.* (2018) obtained.

In the remainder of Table C.2 we provide robustness of the results. In column (2) we show that our results are not driven by the selection of single-plant firms for which we know the exact location. If we also include multi-plant firms (for which we only know the location of the headquarters), the results are very similar.

In column (3) we show that our elasticities are not so much affected by air travel – a possible omitted variable. We calculate the travel time by airplane using the network of passenger flights in 2014. To calculate total travel time from each origin to each destination, we calculate the travel time from each municipalities' centroid to the nearest airport by train, but excluding Shinkansen trains. Then, we assume that it will take 45 minutes to check in and clear security. We find that $\hat{\vartheta}_{\mathcal{A}} = -0.0419$. The decay parameter related to travel time by road is now also somewhat stronger ($\hat{\vartheta}_{\mathcal{R}} = -0.0395$). Hence, it seems that traveling by road and airplane imply similar travel time costs, which is in line with the idea that comfort levels when traveling by airplane are lower because one has to wait on the airport, change to other modes when arriving at the airport, etc. Because the estimated travel cost parameter of air travel is similar to that of travel by road, the share of intermediates shipped by airplane will be low. In line with the estimates, the share of travelers using the airplane is also relatively small. In what follows, we thus disregard air travel.

In columns (4)-(6) we investigate whether reverse causality is an issue by focusing on location pairs that were linked by highways or railways in the initial plans of 1942/1943 (column (4)), the road network in the 8^{th} century (column (5)), or municipalities that are within 25km of a Shinkansen station (column (6)). Although the number of observations is greatly reduced, we obtain similar findings. Thus, travel time costs for trains are lower than for cars.

In column (7) we check whether the travel time elasticities depend on the relative travel times of both transport modes. One may argue that if train travel times are low compared to travel by road, this may attract more business travelers, which in turn increases prices and congestion and lowers comfort levels. We therefore include the ratio of travel time by rail and travel time by road as an additional control variable. 2 standard deviation increase in the travel time ratio implies an increase in the railway travel time elasticity of $0.223 \times (-0.0015) = -0.0003$, which is about 15% of the mean of the baseline main effect. Like with the results on commuting time elasticities we observe a statistically significantly different travel time elasticity when train travel times are relatively high (although the sign is the opposite). However, the effect is again economically negligible.

TABLE C.3 — ESTIMATING THE HETEROGENEITY PARAMETER, ε
(Dependent variable: the log of transformed wages, $\hat{\tilde{w}}_{it}$)

	Baseline specification	Flexible instrument	Year f.e.	Prefecture-by-year f.e.	+ Industry shares	Include ≥1991
	(1)	(2)	(3)	(4)	(5)	(6)
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Wage *(log)*, $\hat{\varepsilon}$	2.8525***	2.7455***	6.0564***	1.8901***	3.6205***	1.9916***
	(0.5519)	(0.5582)	(0.8190)	(0.5685)	(0.7066)	(0.3878)
Industry employment shares (9)	No	No	No	No	Yes	No
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Region×year fixed effects	Yes	Yes	No	No	Yes	Yes
Prefecture×year fixed effects	No	No	No	Yes	No	No
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	6,632	6,632	6,632	6,632	6,632	13,264
Kleibergen-Paap F-statistic	11.34	17.40	13.54	15.98	12.06	4.37

Notes: We instrument wages with the predicted employment in each municipality in each year. In column (2) we add a squared term of predicted employment as an additional instrument. Bootstrapped standard errors (250 replications) are clustered at the municipality level and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

Column (8) in Table C.2 includes two additional controls: a dummy indicating whether a location pair has to cross the east-west 'border', following Wrona (2018), and whether they are on the same island. First of all, the railway travel time elasticity is once again very robust and hardly affected by the inclusion of these variables. The signs of the included variables have the expected signs. The coefficient on the east-west border indicates that trade is reduced by 23% when two municipalities are on different sides of the border. The order of magnitude is comparable to the estimates of Wrona (2018). The effect of being on the same island is much stronger: the coefficient indicates that being on the same island increases trade flows by about 50%. More importantly, the travel time elasticities are hardly affected by the inclusion of these control variables.

C.3 Identifying workers' heterogeneity

In Table C.3 we report the results of regressions to recover the heterogeneity parameter ε (recall that we rely on wage data between 2001 and 2014).

Column (1) displays the baseline specification. We include municipality and region-by-year fixed effects and instrument wages by a Bartik-style predicted employment measure, based on employment shares in 1981. The first stage in Table C.4 reveals that predicted employment has a positive effect on wages: a standard deviation increase in predicted employment is associated with a wage increase of 24%. The instrument is sufficiently strong as the first-stage F-statistic is 11.3. Going back to Table C.3, we find that a wage increase of 1% increases 'transformed' wages

Table C.4 — Estimating the heterogeneity parameter, first-stage results
(Dependent variable: the log of wage, w_{it})

	Baseline specification	Flexible instrument	Prefecture-by-year f.e.	Year f.e.	+ Industry shares	Include ≥1981
	(1)	(2)	(3)	(4)	(5)	(6)
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Predicted employment	0.2389***	0.4860***	0.2622***	0.2148***	0.1782***	0.2502***
	(0.0687)	(0.1575)	(0.0698)	(0.0686)	(0.0447)	(0.0714)
(Predicted employment)^2		-0.0091				
		(0.0125)				
Industry employment shares (9)	No	No	No	No	Yes	No
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Region×year fixed effects	Yes	Yes	No	No	Yes	Yes
Prefecture×year fixed effects	No	No	No	Yes	No	No
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	6,632	6,632	6,632	6,632	6,632	13,264
R^2	0.846	0.847	0.842	0.852	0.905	0.695

Notes: Bootstrapped standard errors (250 replications) are clustered at the municipality level and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

by 2.9%, implying that the heterogeneity parameter is $\varepsilon = 2.85$.

Our estimate of ε is on the low side as compared to the existing literature. Eaton and Kortum (2002) find an estimate of 8.28, but it is based on international trade flows rather than on intra-national commuting flows. Hence, their estimate is arguably not directly comparable to ours. Ahlfeldt *et al.* (2015), who rely on commuting flows within a city (*i.e.*, Berlin), find an estimate of ε of 6.2. However, they recover ε by comparing the variances of the log of transformed wages to the variance of the log observed wages. In this way, however, one may find a strong overestimate of ε because the variances may also relate to each other due to unobserved workplace amenities (Ahlfeldt *et al.*, 2020). If we were to recover ε by comparing variances, we would find that $\varepsilon = 77.79$, which is clearly unrealistically high. Our approach addresses endogeneity concerns in a better way, which leads to a lower, but more realistic, estimate.

In the remaining columns of Table C.3 we investigate the robustness of this estimate. Column (2) includes a squared term of predicted employment in order to investigate whether non-linearity in the instrument may improve the power of our estimate. This appears to have very limited effects, as the estimate is only slightly higher and the standard error only marginally lower. In column (2) in Table C.4 reporting first-stage results, we find that the coefficient of the squared term is statistically insignificant.

Going back to Table C.3, in column (3) we include year fixed effects instead of the more detailed region-by-year fixed effects. We find that the estimate more than doubles in size, suggesting that controlling for regional trends is important. When we include more detailed prefecture-by-year

Table C.5 — Estimating the elasticity of substitution and land expenditure, σ and γ
(Dependent variable: the log of the share of developed land, $\log(\tilde{\mathcal{L}}_{iy}/\mathcal{A}_i)$*)*

	Baseline specification	Exclude $>95^{th}$ perc.	Exclude $>99^{th}$ perc.	+ Geographic controls	+ Industry sector shares	+ Prefecture f.e.
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	OLS
Elasticity of substitution, $\hat{\sigma}$	1.9961***	1.8724***	2.1332***	1.8462***	2.4777	3.0986
	(0.6244)	(0.5803)	(0.7913)	(0.3475)	(3.9616)	(4.5641)
Share spent on land for final firms, $\hat{\gamma}$	0.0720***	0.0699***	0.0770***	0.0640***	0.1452***	0.1013***
	(0.0048)	(0.0012)	(0.0020)	(0.0002)	(0.0103)	(0.0028)
Geographic controls (3)	No	No	No	Yes	Yes	Yes
Industry employment shares (9)	No	No	No	No	Yes	Yes
Region fixed effects	No	No	No	No	Yes	Yes
Prefecture fixed effects	No	No	No	No	No	Yes
Number of observations	1,550	1,604	1,638	1,550	1,550	1,550
R^2	0.839	0.823	0.843	0.851	0.914	0.929
First-stage F-statistic						

Notes: Bootstrapped standard errors (250 replications) are clustered at the municipality level and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

fixed effects in column (4) the estimate is similar to our baseline estimate, albeit somewhat lower (*i.e.* $\varepsilon = 1.89$).

Column (5) directly controls for employment shares in different sectors to focus on variations in calculated wages that are due to manufacturing wages (see Appendix A.2). This makes little difference to the estimate as $\varepsilon = 3.62$.

In the last column of Table C.3 we include data from 1991 onwards and find that the estimated heterogeneity parameter is somewhat lower than the baseline estimate, albeit not significantly so.

C.4 Recovering the elasticity of substitution and land expenses

In Table C.5 we estimate (34), by regressing the share of developed land on land use by residents, intermediate and final firms, to estimate the elasticity of substitution, σ, and the share of final firms spent on land, γ. To scale the data appropriately (for example, because land rents are normalized to have a geometric mean of 1), we estimate a normalization parameter ζ. Further, recall that $\zeta = \zeta_0$, $\sigma = (\zeta_0 + \zeta_1)/\zeta_1$ and $\gamma = \zeta_2/\zeta_0$.

In column (1) we show the most parsimonious specification. We find that $\hat{\sigma} = 2.00$, which suggests that intermediate services are reasonably differentiated. We find that the share of firms spent on land is 7.2%. Valentinyi and Herrendorf (2008) find 5% for U.S. manufacturing firms. Furthermore, according to a 2019 survey on Japanese firms that are owned by a single person, it appears that expenditures on land by final firms are 3%, which is only slightly lower.

In columns (2) and (3) we investigate the sensitivity of the results to our assumption to exclude the upper 2.5% values related to the explanatory variables (capturing land consumption by residents, intermediate firms, and final firms). We find similar values of $\hat{\sigma}$ and $\hat{\gamma}$ in column (2) where we exclude the upper 5% values. Column (3) excludes the upper 1% values, which leads to slightly less precise values for $\hat{\sigma}$.

Column (4) adds geographic controls, while column (5) adds further controls for the industrial composition of an area and region fixed effects. Column (6) includes more detailed prefecture fixed effects. It is shown that adding controls does not materially change the results, as the estimate of $\hat{\sigma}$ hovers around 2, while $\hat{\gamma}$ is about 0.07, but increases to maximally 14% if more detailed controls are added.

Appendix D. Counterfactuals

D.1 Counterfactuals – iterative procedure

Since we have obtained the parameters $\{\hat{\kappa}_{\mathcal{R}}, \hat{\kappa}_{\mathcal{H}}, \hat{\theta}_{\mathcal{R}}, \hat{\theta}_{\mathcal{H}}, \hat{\varepsilon}, \hat{\sigma}, \hat{\gamma}\}$, $\alpha = 0.23$ and $\beta = 0.5$, wages w_i, land rents R_i, residential amenities $\hat{A}_i$, workplace amenities $\hat{B}_j$, productivities $\left\{\hat{E}_{1j}, \hat{E}_{2j}\right\}$, land consumption per intermediate firm ℓ_j, and the amount of land available $\hat{L}_i$, we can undertake counterfactuals. We choose the following starting values: $M_{i1}^{\mathcal{C}} = M_{i1}$, $M_{i2}^{\mathcal{C}} = M_{i2}$, $N_i^{\mathcal{C}} = N_i$, $K_i^{\mathcal{C}} = K_i$, $R_i^{\mathcal{C}} = R_i$, $w_i^{\mathcal{C}} = w_i$, $R_i^{\mathcal{C}} = R_i$, where $\mathcal{C}$ refers to counterfactual values.

We adopt the following iterative procedure.

1. Given $\{\hat{\kappa}_{\mathcal{R}}, \hat{\kappa}_{\mathcal{H}}\}$, calculate counterfactual commuting costs $t_{ijm}^{\mathcal{C}}$.

2. Given $\{\hat{\theta}_{\mathcal{R}}, \hat{\theta}_{\mathcal{H}}\}$, calculate counterfactual trade costs of intermediates $\tau_{jim}^{\mathcal{C}}$.

3. We determine wages by solving (18):

$$\hat{E}_{i1} - \left(\frac{P_i^{\mathcal{C}}}{1-\beta-\hat{\gamma}}\right)^{1-\beta-\hat{\gamma}} \left(\frac{w_i^{\mathcal{C}}}{\beta}\right)^{\beta} \left(\frac{R_i^{\mathcal{C}}}{\hat{\gamma}}\right)^{\hat{\gamma}} = 0,$$

   where*

$$P_i^{\mathcal{C}} = \left[\sum_{j=1}^{I}\sum_{m=1}^{2} K_i^{\mathcal{C}} \left(\frac{\sigma}{\sigma-1}\frac{\tau_{jim}^{\mathcal{C}} w_j^{\mathcal{C}}}{\hat{E}_{j2}}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}.$$

   We use Newton-Raphson to find the unique vector of wages that satisfies the above condition.

4. Land rents are determined using (27):

$$\hat{L}_i - \sum_{j=1}^{I}\sum_{m=1}^{2} \frac{\alpha w_j^{\mathcal{C}}}{t_{ijm}^{\mathcal{C}} R_i^{\mathcal{C}}} \frac{(\hat{A}_i/(t_{ijm}^{\mathcal{C}}\left(R_i^{\mathcal{C}}\right)^{\alpha}))^{\hat{\varepsilon}}(M_{j1}^{\mathcal{C}} + M_{i2}^{\mathcal{C}})}{\sum_{r=1}^{I}\sum_{m=1}^{2}(\hat{A}_r/(t_{rjm}^{\mathcal{C}}\left(R_r^{\mathcal{C}}\right)^{\alpha}))^{\hat{\varepsilon}}} - \frac{\hat{\gamma} w_i^{\mathcal{C}} M_{i1}^{\mathcal{C}}}{\beta R_i^{\mathcal{C}}} - \frac{w_i^{\mathcal{C}} M_{i2}^{\mathcal{C}}}{(\hat{\sigma}-1) R_i^{\mathcal{C}}} = 0.$$

   Again, we apply Newton-Raphson to find the unique vector of land rents that satisfies the above condition.

5. Then, we determine employment in the final sector using (35):

$$\hat{E}_{j2} - \frac{\hat{\sigma} w_j^{\mathcal{C}}}{\hat{\sigma}-1}\left[\frac{(1-\beta-\hat{\gamma})^{\hat{\sigma}}}{\hat{\sigma} R_j^{\mathcal{C}} \ell_j \hat{E}_{j1}^{\frac{1}{(\beta+\hat{\gamma})(\hat{\sigma}-1)}}}\sum_{i=1}^{I}\sum_{m=1}^{2}(\tau_{jim}^{\mathcal{C}})^{1-\hat{\sigma}}\left(M_{i1}^{\mathcal{C}}\right)^{\frac{\beta}{\beta+\hat{\gamma}}}\left(L_{i1}^{\mathcal{C}}\right)^{\frac{\hat{\gamma}}{\beta+\hat{\gamma}}}\left(\frac{1}{P_i^{\mathcal{C}}}\right)^{\frac{1-\beta\hat{\sigma}-\hat{\gamma}\hat{\sigma}}{\beta+\hat{\gamma}}}\right]^{\frac{1}{1-\hat{\sigma}}} = 0,$$

where $L_{i1}^{\mathcal{C}} = \left(\hat{\gamma} w_i^{\mathcal{C}} M_{i1}^{\mathcal{C}}\right) / \left(\beta R_i^{\mathcal{C}}\right)$. Then, employment in the intermediate sector is given by $M_{i2}^{\mathcal{C}} = M_j^{\mathcal{C}} - M_{i1}^{\mathcal{C}}$.

6. Given $t_{ijm}^{\mathcal{C}}$, $\hat{A}_i$, $R_i^{\mathcal{C}}$, $M_j^{\mathcal{C}}$ and $\hat{\varepsilon}$, we solve for the counterfactual population $N_i^{\mathcal{C}}$ in each location i by using (8):

$$N_i^{\mathcal{C}} = \sum_{j=1}^{I} \sum_{m=1}^{2} \frac{(\hat{A}_i / \left(t_{ijm}^{\mathcal{C}} \left(R_i^{\mathcal{C}}\right)^{\alpha}\right))^{\hat{\varepsilon}}}{\sum_{r=1}^{I} \sum_{m=1}^{2} \left(\hat{A}_r / \left(t_{rjm}^{\mathcal{C}} \left(R_r^{\mathcal{C}}\right)^{\alpha}\right)\right)^{\hat{\varepsilon}}} M_j^{\mathcal{C}}.$$

7. Given $t_{ijm}^{\mathcal{C}}$, $\hat{B}_j$, $w_j^{\mathcal{C}}$, $N_i^{\mathcal{C}}$and $\hat{\varepsilon}$, we solve for the counterfactual employment $M_j^{\mathcal{C}}$ in each location j by using (6):

$$M_j^{\mathcal{C}} = \sum_{i=1}^{I} \sum_{m=1}^{2} \frac{(\hat{B}_j w_j^{\mathcal{C}} / t_{ijm}^{\mathcal{C}})^{\hat{\varepsilon}}}{\sum_{k=1}^{I} \sum_{m=1}^{2} \left(\hat{B}_k w_k^{\mathcal{C}} / t_{ikm}^{\mathcal{C}}\right)^{\hat{\varepsilon}}} N_i^{\mathcal{C}}.$$

8. Given $M_{i2}^{\mathcal{C}}$, $R_i^{\mathcal{C}}$, $w_i^{\mathcal{C}}$, ℓ_j, and $\hat{\sigma}$, we determine the number of intermediate firms:

$$K_i^{\mathcal{C}} = \frac{w_i^{\mathcal{C}} M_{i2}^{\mathcal{C}}}{\ell_j (\hat{\sigma} - 1) R_i^{\mathcal{C}}}.$$

9. For loops l and $l+1$, we have $M_{j,l+1}^{\mathcal{C}} = \varrho M_{j,l+1}^{\mathcal{C}} + (1-\varrho) M_{j,l}^{\mathcal{C}}$, $M_{j1,l+1}^{\mathcal{C}} = \varrho M_{j1,l+1}^{\mathcal{C}} + (1-\varrho) M_{j1,l}^{\mathcal{C}}$, $M_{j2,l+1}^{\mathcal{C}} = \varrho M_{j2,l+1}^{\mathcal{C}} + (1-\varrho) M_{j2,l}^{\mathcal{C}}$, $N_{j,l+1}^{\mathcal{C}} = \varrho N_{j,l+1}^{\mathcal{C}} + (1-\varrho) N_{j,l}^{\mathcal{C}}$, $K_{j,l+1}^{\mathcal{C}} = \varrho K_{j,l+1}^{\mathcal{C}} + (1-\varrho) K_{j,l}^{\mathcal{C}}$, $w_{j,l+1}^{\mathcal{C}} = \varrho w_{j,l+1}^{\mathcal{C}} + (1-\varrho) w_{j,l}^{\mathcal{C}}$, $R_{j,l+1}^{\mathcal{C}} = \varrho R_{j,l+1}^{\mathcal{C}} + (1-\varrho) R_{j,l}^{\mathcal{C}}$, where we set $\varrho = 0.1$. We then repeat steps (3)-(8) to reach new equilibrium values $M_j^{\mathcal{C}}$, $M_{1j}^{\mathcal{C}}$, $M_{2j}^{\mathcal{C}}$, $N_i^{\mathcal{C}}$, $w_j^{\mathcal{C}}$, and $R_j^{\mathcal{C}}$when the differences $M_{j,l+1}^{\mathcal{C}} - M_{j,l}^{\mathcal{C}}$ and $N_{i,l+1}^{\mathcal{C}} - N_{i,l}^{\mathcal{C}}$ are sufficiently small.

We check that this counterfactual procedure is able to replicate the population and employment values in 2014 – our base year.

D.2 Counterfactual experiments – maps

Here we plot the changes in travel times, employment, residential population, rents, and wages for the three counterfactual experiments. First, we consider the effects of removing the Shinkansen. Second, we discuss the effects of the planned extensions of the Shinkansen. Third, as a comparison, we investigate the effects of removing highways.

In Figure D.1 we focus on the first scenario where we remove the Shinkansen.

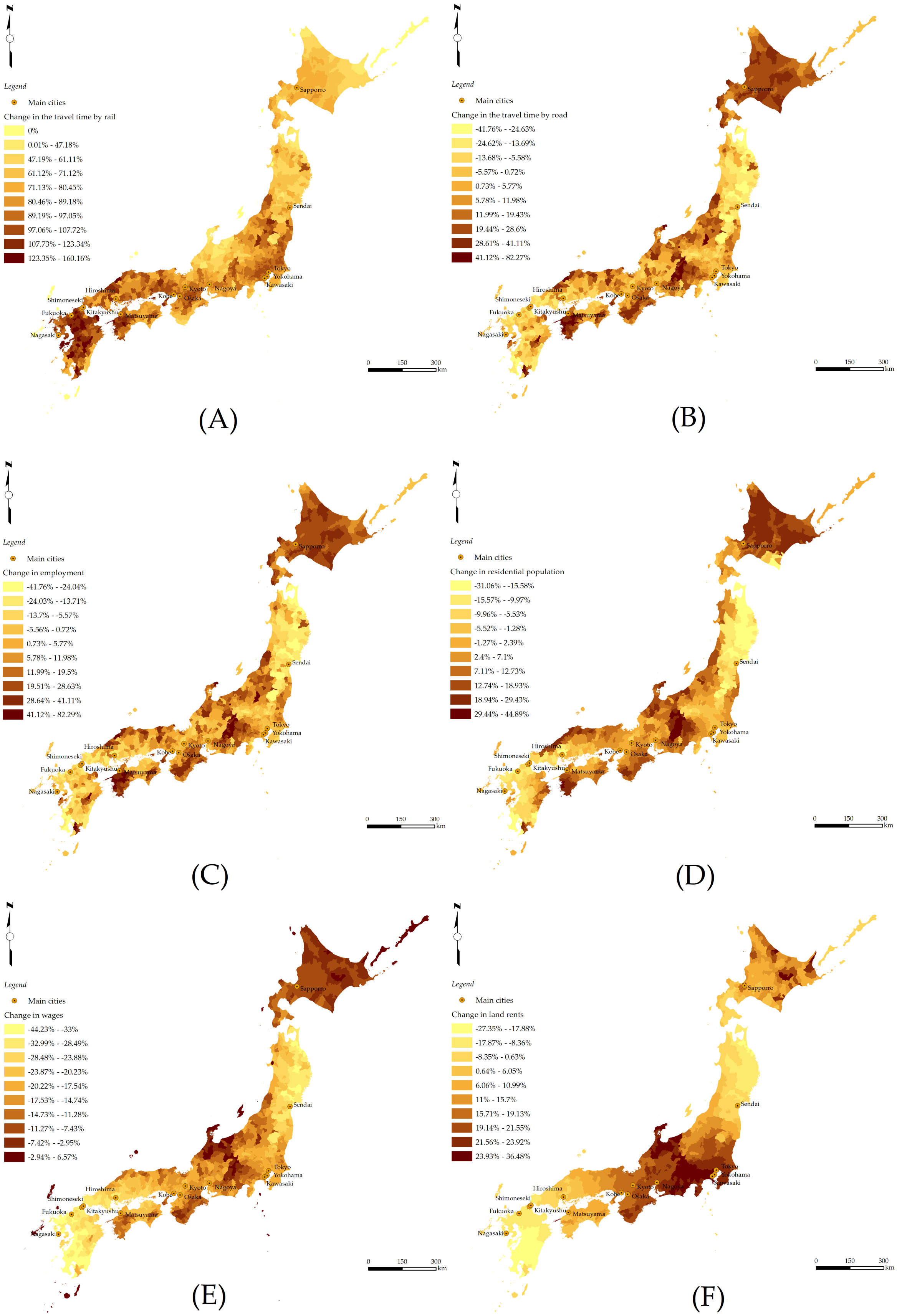


Figure D.1 — Counterfactual experiment 1: No Shinkansen

In Figure D.1A we consider travel time changes if the Shinkansen would not have been built. Kyushu would be much affected and would witness a travel time increase of up to 150%. Effects are sizable along other corridors of the Shinkansen network (*e.g.*, Osaka-Hiroshima and Tokyo-Sendai-Aomori) where travel time changes are substantial. Figure D.1B again shows smaller effects for the average travel time by road due to the reshuffling of jobs.

We find large population increases in the northern part of Kyushu, on Shikoku, and in the Wakayama and southern part of Nara prefectures (to the south of Osaka). Hence, removing the Shinkansen would particularly favor peripheral areas outside the large metropolitan areas, which is in line with the reduced-form results (Section 2.4).

In Figure D.2 we plot changes for the second scenario where we consider the effects of realizing planned extensions. There are particularly large travel time changes in Nagasaki, as well in near Sapporo (see Figure D.2A), which are now directly connected to the Shinkansen network. Moreover, we see reductions in travel time of up to 15% in Osaka, Nagoya, and Tokyo metropolitan areas. Although the road network is left unaffected, we observe also some small travel time changes by road (see Figure D.1B) due to a reshuffling of employment across locations.

For some cities, the effects are considerably larger. For example, Aomori's employment increase is about 11%. By contrast, the employment of large metropolitan areas of Tokyo, Osaka, and Nagoya change by 0.00%, 0.35%, and −0.022%, respectively, so they are hardly affected.

Figure D.3A shows the reductions in travel time by rail when the highway network is removed. These are now considerably smaller than the changes in travel time by road (see Figure D.3B). Particularly the northern part of Kyushu and the area around Tokyo are the most affected by the removal of highways. Again, we find reductions in employment particularly in peripheral areas that are not too far away from urban centers, but the local effects are considerably smaller compared to the local effects of removing the Shinkansen.

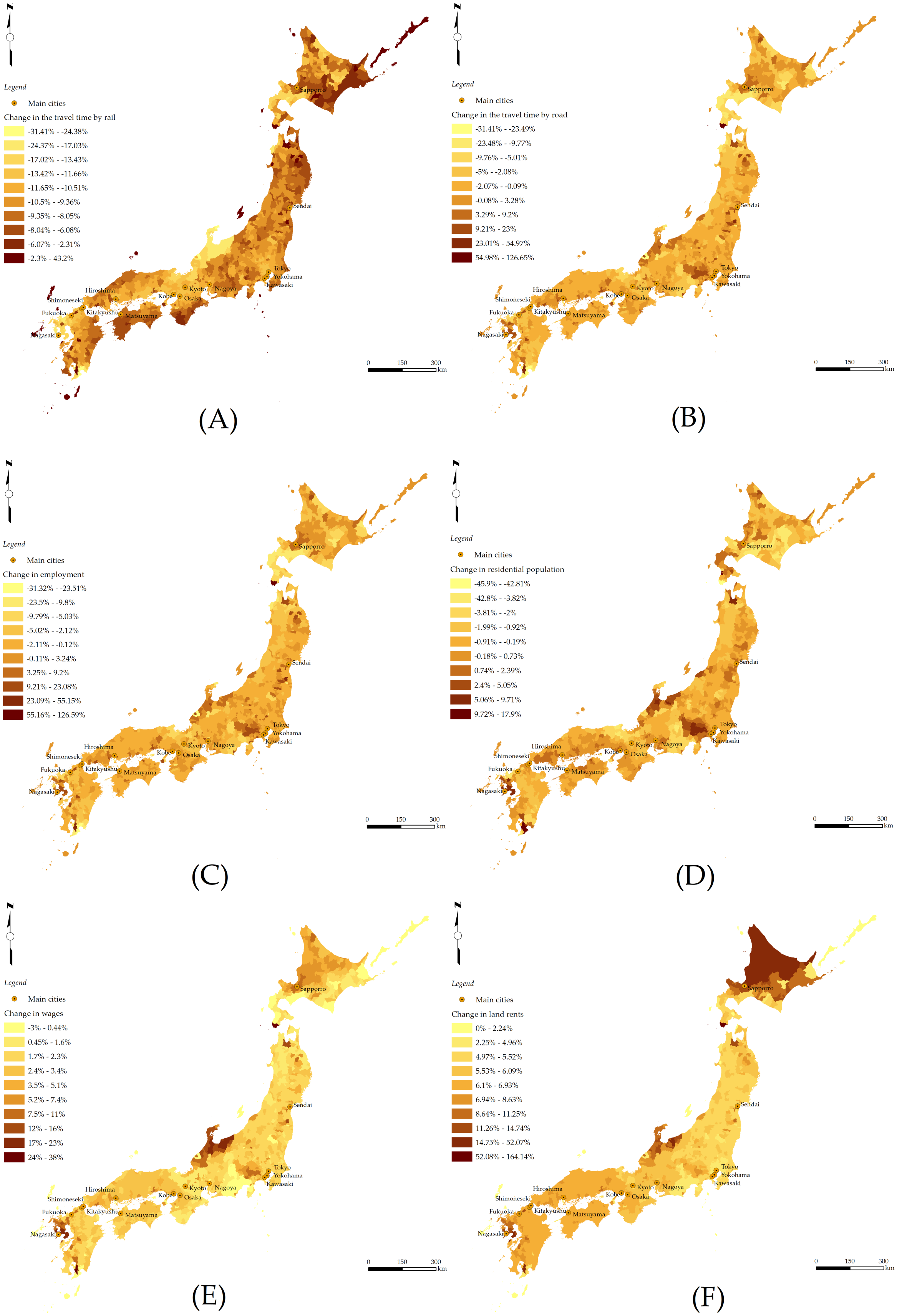
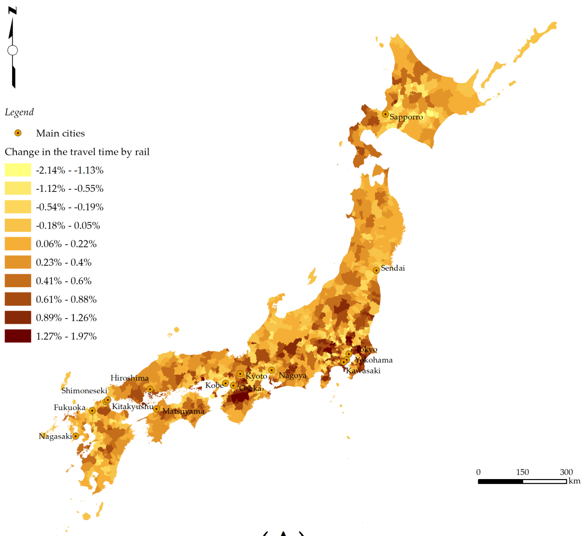
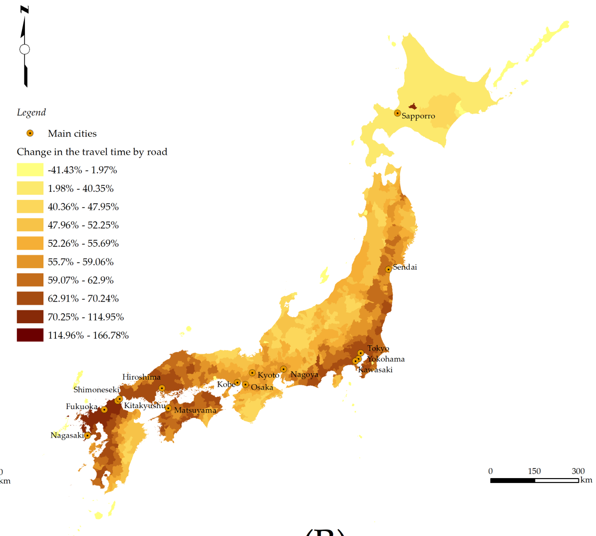


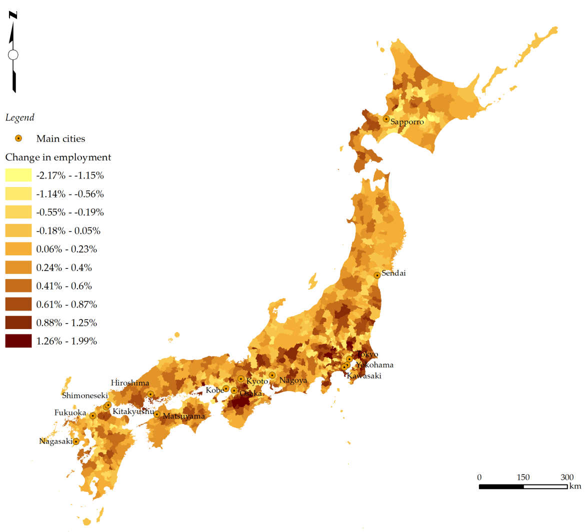
Figure D.2 — Counterfactual experiment 2: Extended Shinkansen



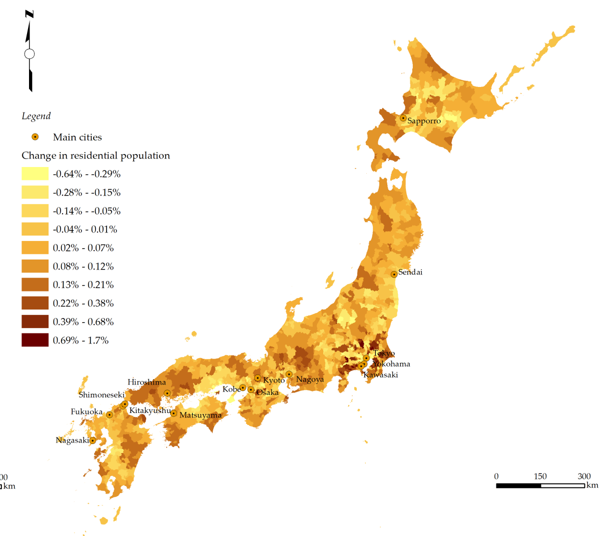
(A)



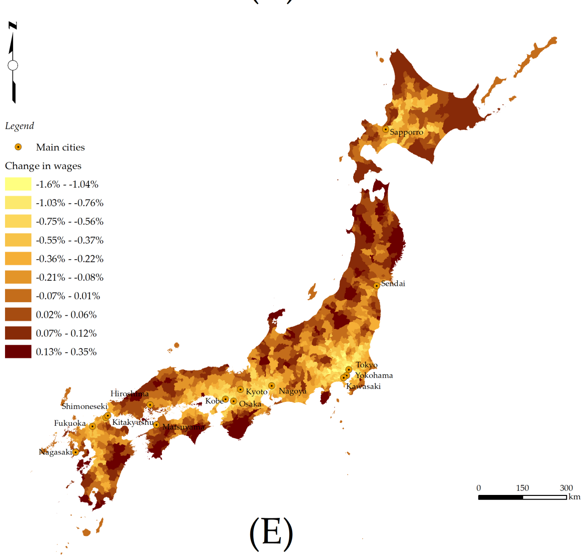
(B)



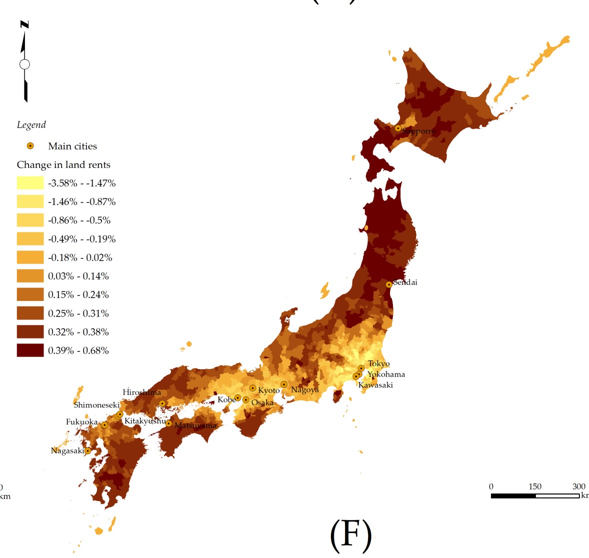
(C)



(D)



(E)



(F)

Figure D.3 — Counterfactual experiment 3: No highways

Appendix E. Agglomeration economies

E.1 Estimation of agglomeration economies

In this appendix we consider the extension of the quantitative model where we allow for agglomeration economies in the intermediate sector. We re-iterate that in the baseline version of our model, we allow for agglomeration economies through input-output linkages between final and intermediate firms, which makes it attractive for firms to locate close together (Krugman and Venables, 1995; Ellison *et al.*, 2010).

Following Ahlfeldt *et al.* (2015), we assume that the TFP of intermediate firms depends on the employment density of intermediate firms in j. More specifically, by log-linearising (30), we have

$$\log E_{j2y} = \tilde{e}_{2jy} + \lambda \log\left(\frac{M_{j2y}}{L_i}\right) + \epsilon_{jy},$$

where $\tilde{e}_{2jy} = \log e_{2jy}$ is a constant, λ is the agglomeration elasticity, and M_{j2y}/L_i is the employment density of intermediates in municipality j in year y.

It is well known that λ may be biased if more productive workers sort themselves into more productive areas because of amenities or because there are unobserved natural advantages correlated to employment density (Combes and Gobillon, 2015). To control for sorting, we will instrument for employment density using the population density in 725. The validity of such an approach rests on the assumption that past population densities are uncorrelated to current unobserved locational endowments and sorting patterns. Usually this assumption is debatable when using data from, say, one or two centuries ago. However, given that Japan looked very different $1,300$ years ago (with the capital being Nara instead of Tokyo or Kyoto), we believe that this assumption is reasonable. As a robustness check, we also use more recent population densities from 1872 and alternatively instrument for employment density using elevation and precipitation.

E.2 Results

We report the results in Table E.1. Column (1) shows the results of the preferred estimation where we regress TFPs on the employment density of the intermediate sector, while instrumenting the latter with population density in 725. The instrument is sufficiently strong: the first-stage F-statistic is 224. If we turn to the first-stage results in Table E.2, we find an elasticity of historic population density with the current intermediate employment density of 0.877. Hence, historic population density is strongly associated with current employment density, which concurs with Davis and Weinstein (2002). Going back to our preferred estimate of agglomeration economies in column (3) in Table E.1, we find $\hat{\lambda} = 0.062$. The literature on agglomeration economies suggests

Table E.1 — Agglomeration economies in the intermediate sector
(Dependent variable: the log TFP of intermediate firms, $\log \hat{E}_{iy}$)

	Instrument: Pop in 725	No instrument	Instrument: Pop in 1872	Instrument: Elevation	Exclude large metro areas
	(1) 2SLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) 2SLS
Employment density of intermediate employment, $\hat{\lambda}$	0.0616	0.1236	0.0895	0.0636	0.0528
	(0.0645)	(0.1175)	(0.0873)	(0.0795)	(0.0575)
Region×year fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	6,613	6,629	6,629	6,629	5,529
R^2		0.494			
Kleibergen-Paap F-statistic	223.957		541.3	874.1	93.4

Notes: In columns (1) and (5) we instrument employment density of intermediate firms by the log of population density in 725, while in column (3) we use the log of population density in 1872. The instruments in column (4) are precipitation and the mean elevation in a municipality. Bootstrapped standard errors (250 replications) are clustered at the municipality level and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

Table E.2 — Agglomeration economies, first-stage results
(Dependent variable: the log of employment density in intermediates

	Instrument: Pop in 725	Instrument: Pop in 1872	Instruments: Elevation	Exclude large metro areas
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Population density in 725 *(log)*	0.8765***			0.8702***
	(0.0513)			(0.1905)
Population density in 1872 *(log)*		0.9289***		
		(0.0899)		
Elevation *(sd)*			-1.1502***	
			(0.0413)	
Industry employment shares (9)	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	6,613	5,529	6,613	1,651
R^2	0.373	0.797	0.667	0.310

Notes: Bootstrapped standard errors (250 replications) are clustered at the municipality level and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

that λ ranges between $0.02 - 0.07$, and so this estimate falls within this range (Rosenthal and Strange, 2004; Combes and Gobillon, 2015). However, the estimate is not statistically significant because we take into account the variance in TPFs, $\hat{E}_{iy}$, caused by the other regression parameters. If we would ignore this, the estimate would be statistically significant at the 1% level (with a T-statistic exceeding 5).

In column (2) we do not instrument for current employment density of intermediate services employment. As expected, the estimate is somewhat inflated (0.124) because we do not control

Table E.3 — Counterfactual experiments with agglomeration economies

	Experiment 1: No Shinkansen	Experiment 2: Extended Shinkansen	Experiment 3: No Highways
	(1)	(2)	(3)
Change in average travel time to employment by train *(in %)*	85.12	-10.86	0.02
	[84.45, 86.18]	[-10.92, -10.69]	[-0.01, 0.04]
Change in average travel time to employment by road *(in %)*	-0.35	0.02	58.63
	[-0.55, 0.13]	[0.00, 0.34]	[58.60, 58.63]
Change in gross welfare *(in %)*	-34.26	0.52	-0.55
	[-53.98, -2.23]	[-28.78, 1.11]	[-15.19, 0.18]
Change in total production by final firms *(in %)*	-99.97	-13.64	-3.05
	[-99.97, -4.22]	[-99.27, -0.03]	[-37.23, 0.03]
Change in total land rents *(in %)*	189.58	20.23	1.72
	[-2.56, 264.20]	[0.36, 465.88]	[-3.41, 74.73]

Notes: 95% confidence intervals are bootstrapped (250 replications) by municipality and in brackets.

well for unobserved characteristics of locations and sorting.

Columns (3) and (4) investigate the robustness of the results when using alternative instruments. When using population density in 1872, the point estimate is only a little higher, albeit still statistically insignificant. The first stage is, unsurprisingly, stronger than when using population density in 725 (the Kleibergen-Paap F-statistic is now 541). Column (4) uses, in our view, a somewhat less convincing, instrument based on geography: elevation. We find that this instrument is very strong. In column (4) in Table E.2 we show that a standard deviation increase in elevation is associated with a $(\exp(-1.150)-1)\times 100\% = -68\%$ change in employment density of the intermediate sector. Going back to column (4) in Table E.1, we find a very similar, estimate of agglomeration economies ($\hat{\lambda} = 0.064$).

One may be concerned that this estimate is driven by the largest metropolitan areas: Tokyo, Osaka, and Nagoya. However, if we exclude municipalities in these metropolitan areas, we find a very similar estimate of $\hat{\lambda}$ in column (5).

All in all, using a conventional strategy to identify agglomeration economies we find estimates that fall well within the range suggested by the literature although the estimates are imprecise.

E.3 Counterfactual experiments – aggregate effects

Let us consider the results for the counterfactuals in the presence of agglomeration economies in the intermediate sector. To obtain the counterfactual values, we add an extra step to the iterative procedure in Appendix D.1. That is, after step 8, we update the TFPs related to intermediate firms as a function of the changes in employment density of intermediate firms. We replicate Table 7 in Table E.3.

Table E.4 — Counterfactual experiments with agglomeration economies: regressions
(Dependent variable: the change in counterfactual employment)

	Scenario 1: No Shinkansen		Scenario 2: Extended Shinkansen		Scenario 3: No Highways	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Shinkansen station in 2014 <10km	-0.1505	-0.1275	0.0007	0.0002	-0.0013	-0.0017
	(0.1097)	(0.0910)	(0.0050)	(0.0053)	(0.0032)	(0.0044)
Planned Shinkansen station <10km			0.0152	0.0066		
			(0.0192)	(0.0092)		
Railway station <10km	-0.0753	-0.0827	0.0043	0.0038	0.0004	-0.0003
	(0.0657)	(0.0652)	(0.0045)	(0.0052)	(0.0006)	(0.0017)
Highway <10km	-0.0061**	-0.0120***	0.0073	0.0054	-0.0046	-0.0047
	(0.0031)	(0.0036)	(0.0058)	(0.0046)	(0.0044)	(0.0037)
Geographic controls	No	Yes	No	Yes	No	Yes
Region fixed effects	No	Yes	No	Yes	No	Yes
Number of observations	1,658	1,658	1,658	1,658	1,658	1,658
R^2	0.0895	0.6844	0.0086	0.0746	0.1547	0.3664

Notes: Geographic controls include the log of area size, the log of population in 1872, the share of developed land, the distance to the coast, as well as the mean elevation. Standard errors are bootstrapped (250 replications) by municipality and in parentheses; *** $p < 0.01$, ** $p < 0.5$, * $p < 0.10$.

Looking at the aggregate results, we find that the welfare reductions are much larger than the baseline results if the Shinkansen were to be removed: welfare would decrease by 34%. However, the implied confidence bands are also larger so that the estimate is not statistically significantly larger. In any case, the presence of agglomeration economies strengthens our conclusions that the Shinkansen has generated strong positive welfare effects. The conclusion that the planned extension implies in oversupply of transport infrastructure is reinforced because gross welfare effects are again essentially zero.

E.4 Counterfactual experiments – local effects

One may be concerned that, despite the aggregate effects being similar, the changes in the spatial distribution of economic activities as a result of changes in travel time may be quite different. We therefore replicate Table 8, in which we regress the predicted employment density on dummies indicating whether the municipality has a Shinkansen station, railway station, or highway connection in 2014. It is straightforward to observe that the coefficients are similar to the results without agglomeration economies in the intermediate sector, albeit less precise. To sum up, we can safely conclude that *local density economies in the intermediate sector do not drive the effects of transport infrastructure.*