# RIDING ON THE NEW SILK ROAD: QUANTIFYING THE WELFARE GAINS FROM HIGH-SPEED RAILWAYS*

Mingzhi Xu
University of California, Davis

Job Market Paper

October 2017

**Abstract**

This paper studies the aggregate and distributional impacts of high-speed railways (HSR) in an economy with producer-supplier linkages. HSR connection generates productivity gains by improving firm-to-firm matching efficiency and leading firms to search more and better suppliers. We first provide reduced-form evidence that access to HSR significantly promotes exports at the prefecture level. This effect is stronger in regions closer to the HSR hubs. Then, we construct and calibrate a quantitative spatial equilibrium model to perform counterfactuals, taking into account trade, migration, and outsourcing. The quantitative exercise reveals that the construction of HSR between 2007 and 2015 increased China's overall welfare by 0.46%, but was also associated with an increase in national inequality. However, the rising inequality could be alleviated by reforms with the goal of reducing internal migration costs. In addition, we find gains from HSR are larger when labor migration costs are higher, implying HSR project is an ideal policy for countries like China which feature high internal migration barriers.

**Key Words:** Transport infrastructure, outsourcing, high-speed railways, welfare, inequality, China

**JEL Classification:** F10, F16, O18, R10, R12, R23

*We are extremely grateful to my advisers, Robert C. Feenstra, Giovanni Peri, Deborah L. Swenson and Ina Simonovska, for their guidance. We thank Samuel Kortum, Meixin Guo and Shu Shen for help offered on this project. We also thank Agostina Brinatti, Nicolas Caramp, Luca Macedoni, Christopher Meissner, Brendan Price, Sarah Quincy, Katheryn Russ, Sanjay Singh, Alan Taylor, Vladimir Tyazhelnikov and Qin Yu for helpful comments, as well as seminar participants at UC Davis. We especially thank Robert C. Feenstra for generously providing the China Customs Dataset and research support. Financial support from the China Government Scholarship is gratefully acknowledged. Please direct correspondence to: Mingzhi Xu, University of California, Davis, email: mzhxu@ucdavis.edu. All errors are mine.

# 1 Introduction

*"A major new high-speed rail line will generate many thousands of construction jobs over several years, as well as permanent jobs for rail employees and increased economic activity in the destinations these trains serve. High-speed rail is long overdue, and this plan lets American travelers know that they are not doomed to a future..."*

—Former U.S. President Barack Obama, 2009

Adequate infrastructure has been considered as the essential ingredient for a country's economic growth. While there has been rich research on the economic consequences of transportation infrastructure that reduces the cost of delivery of goods[1], considerably less attention has been paid to the infrastructure project that aims to decrease the costs of passenger travel, such as high-speed rail (HSR). Existing studies' enthusiasm on this increasingly popular transport mode is based on the partial evidence of HSR impacts on employment and firm productivity[2], but there is even little structural evidence for the economy-wide importance of HSR and its role in inequality. In this paper, we aims to fill this gap by studying the aggregate and distributional consequences of HSR from counterfactual simulation of the calibrated multi-region general equilibrium model with producer-supplier linkage, costly trade, and frictional mobility.

Our theoretical framework highlights the productivity gains at the supply-side, and we attribute it to the enhanced outsourcing motivation of firms led by the HSR-induced improvement in matching efficiency between producer and suppliers[3]. While our main contribution is to quantify the impact of HSR connection on welfare and study its distributional effects from skill and geographic dimensions, we also shed light on the channel through which internal migration frictions interact with regional outsourcing ability in affecting the welfare of the whole nation. We find gains from HSR are larger when labor migration costs are higher, implying HSR projects are an ideal policy for countries like China which feature high internal migration barriers.

To motivate the model, taking advantage of the rapid expansion of high-speed rails in China as a plausibly exogenous shock of improving firm-to-firm matching efficiency across Chinese cities over time, we find the connection to HSR contributes to the improvement of regional exporting performance, which is consistent with the recent studies. To address endogenous placement issue, we adopt an instrumental variable strategy based upon the construction of least cost path spanning tree

[1]Transportation infrastructure of this type mainly includes highway system and traditional railroads. In our study, we consider the Transportation infrastructure involved in long distance traveling, and we do not focus on the commuting transportation infrastructure, such as public bus, Bart, and subway, which are primarily used within a city.

[2]Most of the literature are empirical studies, from which one can only learn the net effect of HSR on some particular outcomes, without understanding the underlying channels nor the less obvious consequences induced by HSR. One exception is Bernard, Moxnes, and Saito (2015) which structurally studies the impact of HSR on firm's buy-supplier network formation while remains silent on the consumers. More discussions on existing literature is provided in the next section.

[3]HSR provides faster means of transportation that make the circulation of managers easier, which facilitates the transmission of information and increases the efficiency of monitoring. Consequently, potential gain sources from the improvement in firm productivity that is attributed to the rising incentive to outsourcing associated with the lower search and outsourcing costs.

networks. Complementing with IV regression, we carry out the event study to rule out the violation of parallel trend assumption that HSR placement could be based on the past and expected future growth, and the main finding remains robust. Besides the direct impact, we find HSR connection has positive spillovers to nearby regions, and such effect is stronger for the area closer to HSR hubs. As far as we know, we are the first to document the causal relationship between HSR network and exports in case of China.

The theoretical model which forms the basis for our analysis extends Eaton, Kortum, and Kramarz (2016) framework to an economy with multiple regions connected to each other through costly trade, frictional migration, and task outsourcing. On the production side, regions trade with each other and differ in both technology and their exposure to outsourcing. To capture the important mechanism through which HSR connection affect inequality, we consider unskilled labor to be substitutable to intermediate tasks. Production and regional outsourcing ability determine the demand of labors of different types, and the supply of goods that are used as final consumption and intermediate task input, in all regions. On the worker side, each worker decides where to work and live, based upon the expected utilities obtained from all potential destinations, which depends on the region-specific amenities, prices of consumption goods, wages as well as the idiosyncratic taste shocks. Worker's migration and consumption determine the supply of labor and the demand for goods, in all regions. In equilibrium, labor distribution and factor prices are jointly determined such that all markets clear.

We calibrate our model to the equilibrium of Chinese economy in 2007. In the calibration, a region in the model is either a province or a municipality in China. The key parameters of the model are migration costs which are constructed using the migration flow from China census and internal trade costs that are sourced from the China's 2007 extended IO table. The remaining parameters governing regional productivity and production technology are estimated using various macro data sources and using the auxiliary equations from the model.

Our quantitative exercise reveals that the construction of HSR between 2007 and 2015 increased China's overall welfare by 0.46%, but was also associated with an increase in national inequality. However, the rising inequality could be alleviated by the reform to reduce internal migration cost. Intriguingly, we find labor mobility and outsourcing ability are substitutable to each other in promoting welfare. The HSR project is found to generate greater welfare gains in an economy with high frictions in labor mobility, making it an ideal policy for China to improve the overall well-being. On top of this, we detect the differentiated pattern of mobilities for different types of workers in influencing welfare gains and inequality. When the unskilled labors are more mobile, the incentive of outsourcing becomes smaller. In the absence of outsourcing motivation, the shrink in the relative demand for unskilled workers would be quite limited in regions linked to HSR network, so is the benefits of being connecting to HSR. In contrast, when only reducing the migration cost of the skilled labors, there remains the substantial difference in the labor cost of using unskilled labor, and firms still have the incentive to outsourcing manufacturing tasks. As a result, the relative demand for unskilled (skilled) labor declines (rises) after being connected to HSR. In fact, as skilled labors become mobile, they become more willing to move to HSR regions where their relative return is high, and this increases their supply and decreases their cost in HSR regions, which in turn leads firms to enhance the incentive to outsourcing more manufacturing tasks further. The complementary relationship between skills and HSR raises the overall welfare, which is primarily through the increased relative return to skills attributed to outsourcing.

## 2 Related Literature

This paper primarily relates and contributes to the growing literature on the impacts of transportation infrastructure projects on economic development. One strand of literature concentrates on the role of infrastructure in reducing transport cost for shipping goods between regions and in reducing the commuting time for individuals within a city[4], and most of them focus on the aggregate and distributional effects, as well as the geographic distribution of economic activity. For instance, Donaldson (2016) examines the welfare gains resulted from the reduction of domestic trade cost using Indian's railroad's placement. The similar question has also been explored by Banerjee, Duflo, and Qian (2012) with China's highway system. Michaels (2008) studies the rising of skill premium in the local labor market induced by highway network. Faber (2014) and Duranton, Morrow, and Turner (2014) consider the effects of highways on the geographic pattern of economic activity[5]. Relative to these existing literature, our paper draws attention to a different means of transportation infrastructure, i.e., the inter-city passenger transportation that is featured by high-speed rails transportation.

The closest literature to our research that studies the impact of high-speed rails on the economic development includeBernard, Moxnes, and Saito (2015) on Japan, Charnoz, Lelarge, and Trevien (2016) on France, and Lin (2017) on China. Lin (2017) attributes the expansion in urban employment to the HSR induced market access, while the other two emphasize the tightened firm-to-firm linkage via HSR. Specifically, HSR provides faster means of transportation that make the circulation of managers easier, which facilitates the transmission of information and increases the efficiency of monitoring. As a consequence, firms become more willing to outsourcing the production tasks either to other firms (Bernard, Moxnes, and Saito (2015)) or to their remote affiliates (Charnoz, Lelarge, and Trevien (2016)). However, both of the studies focus on the performance of the individual firms and we regard them as the micro foundation to our studies. Advantageous to previous studies, the analysis of this paper takes a structural method, which allows us to speak to both the aggregate and distributional effects in a unified framework. It will be attractive to making policy that usually requires the tradeoff between the overall welfare gains from HSR and the corresponding effects on inequality. Besides the theoretical contribution, we extend this HSR literature by providing empirical evidence that access to HSR network substantially improves regional exporting performance. As far

[4] The detailed information regarding on this topic refer to the survey byRedding and Turner (2014).

[5] Faber (2014) studies the asymmetric impacts of highway system on the core and peripheral markets in China, while Duranton, Morrow, and Turner (2014) examine the highway's impact on the composition of trade for US cities.

as we know, we are the first to explore the causal relationship between HSR connection and exports.

The paper is also related to the large literature on the firm-to-firm connection in affecting production, which cast the micro foundation to our theoretical analysis. The common feature of these research is to emphasize the importance of face-to-face meeting that matters for monitoring and relationships in production. An improvement in this firm-to-firm linkage is associated with an enhance in firm performance, either through the direct face-to-face meeting among managers as documented by Cai and Szeidl (2016), the frequent air traveling as studied by Cristea (2011) and Giroud (2013), or through the land traveling by high-speed rails as studied in Bernard, Moxnes, and Saito (2015) and Charnoz, Lelarge, and Trevien (2016). In contrast and superior to these studies, we emphasize the creation of linkage in response to infrastructure shocks from the macro perspective and can capture the general equilibrium effect of the infrastructure placement that is missing in the previous literature.

Welfare gains from HSR in our paper originate from the rising capability to outsourcing the manufacturing task, which makes this article related to the studies on the determinants of domestic and foreign sourcing and the economic consequences, pioneered by the work of Feenstra and Hanson (1996). At the macro level, Wright (2014) finds the international outsourcing raises home country's productivity. From micro perspective, Amiti and Konings (2007), Goldberg, Khandelwal, Pavcnik, and Topalova (2010), Bøler, Moxnes, and Ulltveit-Moe (2012), Antras, Fort, and Tintelnot (2014) and Halpern, Koren, and Szeidl (2015) attributes the improvement in firm performance to the imported inputs that are imperfectly substitutable to the domestic inputs. In line with the literature that focuses on the productivity gains from outsourcing, we study how HSR influence a region's capability of outsourcing and develop a quantitative spatial equilibrium model to estimate regional outsourcing capability structurally[6]. This extension makes it possible to analyze the distributional effects of an improvement in a region's outsourcing ability along both skill and geographic dimensions, connecting this literature to a rich literature on trade and the inequality.

Finally, the model of this paper relates to the recent quantitative trade literature[7], pioneered by Eaton and Kortum (2002). The common policy instruments among these models are either to deal with trade cost for shipping products or the migration cost for moving labors, which makes them infeasible to directly capture the improvement in firm-to-firm matching efficiency due to HSR connection. To directly model the improvement in matching efficiency and to fit the economic environment of China, we adjust and extend Eaton, Kortum, and Kramarz (2016)'s framework to incorporate frictional labor mobility across multiple regions within a country. The quantitative results also highlight

[6]Previous work studying international outsourcing consider the reduction in trade costs as the primary driver of the increased offshoring activities across countries, while our focus is the improved buyer-seller matching efficiency. Though these two factors could both explain the rising demand for outsourcing, they have completely different implications on the consequential distributional effects.

[7]Recent studies on this topic include Alvarez and Lucas (2007), Arkolakis, Costinot, and Rodríguez-Clare (2012), Hsieh, Hurst, Jones, and Klenow (2013), Di Giovanni, Levchenko, and Zhang (2014), Caliendo, Parro, Rossi-Hansberg, and Sarte (2014); Caliendo and Parro (2015), Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016), Edmond, Midrigan, and Xu (2015) and others. More literature is surveyed in Costinot and Rodriguez-Clare (2013). Two applications of these quantitative method to China includes the contemporaneous papers by Fan (2015) and Tombe, Zhu, et al. (2015).

that labor mobility plays a crucial role in determining the distributional effects of HSR connection. To sum up, the paper provides a new framework to take care of trade, migration, and outsourcing in a tractable framework, which can be used to examine the welfare and inequality implications under the various different scenario.

The rest of the paper is organized as follows. We introduce the background of HSR of China and the provide data information in Section 3 and explores the causal relationship between HSR connection on exporting performance in Section 4. In Section 5, we present our multi-regional model of domestic sourcing. In Section 6, we calibrate the model parameters to match a set of moments at the initial equilibrium level in 2007, and the results are discussed in Section 7. We perform some counterfactual exercise and consider the welfare implication in Section 8. Finally, we provide an extended model with reservation demands for certain types of labor in production in Section 9, and make our conclusions in Section 10.

# 3 Background and Data Description

## 3.1 Background

High-speed rails usually refer to the railroads that could carry trains with a top speed at least greater than 250km/h according to the government document of Ministry of Railways of China (MRC) in 2013. Though the planning for high-speed railway could date back to early 1990, this idea was not actually considered until 2002 when the newly appointed governor of MRC proposed the "Great Leap Forward" strategy. The plan aims to expand the current railway capacity as well as the introduction of high-speed rails in China. Before the high-speed railway construction boom[8] in 2007, there was only one passenger-dedicated HSR between Qinhuangdao and Shenyang opened in 2003, but with an operating speed of $200km/h$ only[9]. In 2008, the State Council set the goal of increasing traditional railroad length to 120,000 kilometers and 16,000 kilometers for the HSR, aiming to form four north-south corridors and four east-west corridors in their Mid-to-Long Term Railway Development Plan[10], and HSR era began. Up to the end of 2015, HSR network has covered 190 cities (almost all the capital cities and the ones with a population greater than 500 thousand), with total HSR length as 19,000 kilometers. The expansion of HSR network in China from 2008 to 2015 is displayed in Figure A.3, where cities are marked blue if they were linked to the HSR network at the end of each year. Figure A.4 shows the planned HSR network to be constructed by 2020. Regarding pure length, China's

[8] There have been five waves of speed acceleration before the massive HSR being put into use, namely, in 1997, 1998, 2000, 2001 and 2004. However, these improvement aims at traditional trains, and the improved speed remains low, i.e., the average speed increased to $65.7km/h$ after the fifth speed acceleration in 2004.

[9] Technically speaking, the Qinhuangdao-Shenyang line was not used as the high-speed rails as the carrying speed was still below the minimum speed criteria, and it functioned as the pilot railway to test the feasibility of HSR operation as also noted by Ou, Richard, Jin, and Zhou (2014).

[10] The planning was later revised in 2015 to expanding traditional railroad length to 175,000 kilometers and 45,000 kilometers for the HSR by 2030. Other detailed background knowledge of HSR of China could refer Ou, Richard, Jin, and Zhou (2014), Lin (2017) and Qin (2017).

HSR network is the longest in the world so far[11]. This rapid construction of HSR network, between 2008 and 2015, also provides researchers with the plausibly exogenous variation in each city's access to HSR, to study the causal impact of HSR on social and economic outcomes.

The main original objective of high-speed railways aims to provide the capital cities with faster ways of transportation, as documented in the Mid-to-Long plan. HSR lines are also expected to complement the existing transportation infrastructure as much as possible. Most of the lines are designed to carry bullet trains with an average speed of 350 $km/h$, which are much faster than most of the intercity lines (with average speed of 200 to 250 $km/h$, such as Guangzhou-Zhuhai Intercity Line), and other traditional top-speed lines (with average speed of 120 to 160 $km/h$, such as trains titled with K, T and Z). The traveling time between Beijing and Shanghai has considerably been shortened from 13 hours to about 5 hours after Jinghu High-Speed Railway opened to the public for commercial service in 2011. In the mean while, the HSR tickets are competitive compared to other transportation methods, and it varies depending on the speed of the train. For instance, the average cost of a second-class HSR fare with speed of 300 to 350 $km/h$ is roughly 0.077 dollars per kilometer, which is cheaper than the discounted air tickets and three times higher than tickets of the traditional express railways[12]. The average ticket cost is also about a quarter of the other country's high-speed rail fares. To a certain extent, this reflects high passenger density and occupancy rates, as well as low construction and operating costs for high-speed railways in China.

As an economical means of transportation, the ridership of HSR has grown rapidly. According to Ou, Richard, Jin, and Zhou (2014), the total HSR ridership is 672 million in 2013, roughly twice the amount of the ridership carried by air[13]. While the aircraft is still an attractive option for traveling, the reliability of HSR services, departure frequency, and comfort make it very competitive for most medium-distance trips. For short trips, some cities have begun to use the HSR as a regular commuting method, such as Baoding to Beijing and Wuxi to Shanghai. Among the passengers of HSR, business traveling account for a majority proportion, which is much higher than that of the passengers taking traditional trains[14]. Therefore, the development of HSR network has undoubtedly been believed to enhance the social-economic linkage between regions and to bring significant economic development benefit, which accrues to business and individuals even when themselves do not travel.

High-speed railways are also expected to insert tremendous influence worldwide in the future, and it has been made the key component for China government to revive the "New Silk Road" since China first announced her "One Belt, One Road" foreign policy in 2013. Up to 2015, thirty-eight countries and regions have planned to build high-speed rail lines and work mostly with China on

[11]The second longest country is Spain (2515 $km$ in 2013), followed by Japan ( 2388 $km$ in 2013) and France ( 2036 $km$ in 2013).

[12]The tickets of the traditional express train are usually hard to buy, and the service quality is significantly lower.

[13]The average annual growth rate of ridership is 39% for HSR, compared to 13% for air.

[14]For instance, according to on-the-ground surveys in Ou, Richard, Jin, and Zhou (2014), the business traveling account for 62% of the total ridership in Tianjin-Jinan HSR line, compared to 51% for traditional trains. In contrast, the proportion of tourist traveling by regular trains is greater than that of HSR.

constructing these HSR[15]. A global high-speed rail network (covering Asia, Europe, middle east and Africa) with a total length of 93,000 kilometers is planned, and 60% percent of these planned HSR was expected to be finished by the year 2030. The Chinese government has negotiated agreements to construct 34,700 kilometers for other countries, including 26,300 kilometers allocated for the "New Silk Road", which would allow passengers to travel 6,000 miles (from Shanghai to London) at speeds of up to 200 mph. The world is bound to be more integrated with the introduction of the fast means of transportation in the future, which will undoubtedly affect the living style, production distribution, as well as the well-being of each along the network. Understanding to what extent such an infrastructure project can improve production efficiency and living standard will be an important question from perspectives of both academia and politics.

### 3.2 Data Description

The prefecture level social economic variables source from China City Statistical Year Books from 2006 to 2015, such as population, average income, average ridership, usage of internet, and e.t.c[16]. The initial connect time of HSR for each city is collected through various methods. Our primary source is the major events section in China Railway Statistical Yearbook, where we search with each city's name for the documented connecting time. The second source of exact opening date is online news, where we search keywords such as city name, "Dongchezu", "Hexiehao" for the earliest documented time (using Google advanced search engine).We only include the HSR lines that run at an average speed of 250 $km/h$, which leaves us 173 cities connected with HSR by 2014. The regional export information sources from China Customs that is aggregated at prefecture level.

To calibrate the model parameters on migration costs, we use the bilateral migration flows calculated from two sources of data. The first is the 2000 Population Census which provides detailed population counts at the county (sub-prefecture) level and by skilled type (individuals with a college degree or above). The second is the aggregated 2005 1% Population Survey, which provides the aggregated migration flow at bilateral province level. We source the share of skilled workers for each migration flow from 2000 Census and calculated the migration flow by labor type for 2005, which is used as the initial condition of the economy. To calibrate the structural parameters on trade cost, we use the 2007 national input-output table that reports the bilateral trade flows for all provinces and a variety of sectors, and we aggregate it at the bilateral province level. Finally, high-resolution data of land usage, land cover, and elevation, which are used to construct the least cost path HSR routes, are obtained from Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences.

[15]See reference: https://www.nextbigfuture.com/2015/02/china-spending-to-build-40000-miles-of.html

[16]In the event study, we use earlier years to also control for the pre-treatment trend.

# 4 Motivational Evidence

Our empirical investigation is guided by Bernard, Moxnes, and Saito (2015) suggesting that connection to HSR improves firm's productivity, by decreasing firms' search cost and increasing the incentive to outsourcing associated with lower travel time. In line with the literature, we provide new evidence that sheds light on the enhanced buyer-supplier relationships induced by HSR in China, by exploring the causal impact of HSR connection on the exporting performance[17] at the prefecture level. Conceptually, better means of intercity transportation increases the matching efficiency by lowering the cost of face-to-face interactions across spaces[18], which leads to more efficient buyer-seller relationships. Consequently, we expect the average productivity to rise for regions after being linked to HSR network, which is reflected by the expansion of exports after HSR connection.

## 4.1 Difference in Differences Specification

The data described in the previous section are used to estimate the effects of HSR connection on the changes of export. The baseline estimation strategy uses a difference in difference specification in the form of:

$$\ln(y_{ct}) - \ln(y_{c0}) \quad = \quad \beta Connect_{ct} + \eta X_{ct} + \gamma_c + \epsilon_{ct} \tag{1}$$

where $y_{ct}$ denotes the outcome variable of prefecture $c$ in year $t$, $\gamma_c$ is a prefecture fixed effect, $Connect_{ct}$ is the dummy variable that indicates whether prefecture $c$ was connected to the HSR network between 2006 - 2014, and $X_{ct}$ is a vector of prefecture control variables as explained in data description. The error term $\epsilon_{ct}$ is clustered at the province level. The identification assumption of difference-in-difference estimations is the parallel trends of export between HSR-connected cities and the other cities without the connection of HSR.

To check the parallel trend assumption, we run a variation of (1), controlling for the leads and the lags of the initial connection dummies

$$\ln(y_{ct}) - \ln(y_{c0}) \quad = \quad \sum_{m=1}^{3} \beta_m FirstConnect_{c,t-m} + \sum_{n=1}^{3} \beta_n FirstConnect_{c,t+n} + \eta X_{ct} + \gamma_c + \epsilon_{ct} \tag{2}$$

where $FirstConnect_{c,t}$ is the dummy variable that indicates whether prefecture $c$ is first connected to the HSR network in year $t$. It switches to unity only if the HSR line connecting city is opened in

[17] We believe exporting performance is less affected by the other confounding factors (such as any demand changes) other than the changes to the supply side. This may also be the reason that the impact of HSR on overall income could be both positive and adverse as founded in Qin (2017) and Lin (2017). As far as we know, we are the first to explore the causal relationship between HSR connection and exporting performance. This section motivates our theoretical analysis that emphasizes the productivity gains (changes of supply side) from HSR connection.

[18] The appropriateness of a supplier's product for the buyer's purpose is the key to the formation of buyer-supplier linkage. Without HSR, it would be less convenient for the buyers (e.g. firm managers of the eastern coast in China where labor wages are high) to frequently travel to the buyer's place (e.g. firm managers of the inner land in China where labor wages are low) to screen the provision of the outsourced input so that they are appropriate for this particular firm.

year $t$. $FirstConnect_{c,t-m}$ ($m = 1, 2, 3$) stands for the the $m$-th lead, and $FirstConnect_{c,t+n}$ ($n = 1, 2, 3$) stands for the the $n$-th lag. Including the leads allow us to control for the pre-HSR effects of future HSR connection as a placebo test, and helps to disentangle the anticipatory effects from the actual connection effects. Controlling for the lags enable us to trace the treatment effects in the future years after the initial placement of HSR. In reality, we would expect the anticipatory effects play no role and is irrelevant for exports, and expect the treatment effect to be significantly positive for the lags.

### 4.2 Least Cost Path Spanning Tree Networks

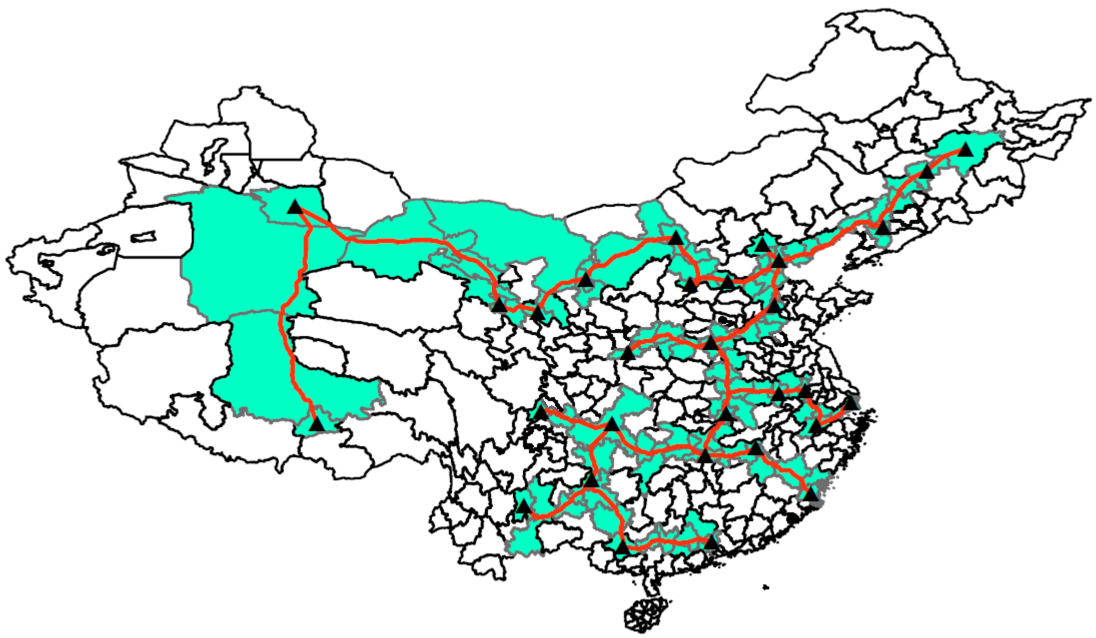


Figure 1: The Least Cost Path Spanning Tree Network

Estimation of (1) using OLS would imply the assumption that the selection of nodal cities in HSR network was randomly assigned, which would be strong. As stated in the official documents of The Ministry of Railway of China, the objective of this HSR grid is to connect all provincial capitals and other major cities with faster transportation. The placement of HSR lines, according to the Ministry of Railway of China, should be targeted on those politically important and economically prosperous cities and give them the priority to be linked in HSR network. It is likely that the policy makers would choose to build HSR near the cities that they expect to have a higher economic growth that is unobserved by the researchers. In this sense, the OLS estimates would impose an upward bias in the case of such correlation. To address this endogenous placement problem, we construct a least-cost path spanning tree network as an instrument for the actual HSR connection, in a similar way with Faber (2014), to evaluate China's HSR expansion. For robustness, we also construct the Euclidean spanning tree network as the alternative instrument. Both instruments corresponds the question that which routes the central government would like to construct if the only policy objective had been to link all the provincial capitals on a single network (as stated in the official report), subject to the total construction minimization[19].

The following paragraphs briefly explain the spanning tree instruments and a more detailed description is provided in the empirical appendix. The construction method follows Faber (2014). The first step is to compute the least cost paths between all capital city pairs on the basis of remote sensing data of land elevation, land cover and land use. To do so, we adopt a simple construction cost function that linearly increases with land slope gradients and lands that are covered by water, wetland, or built structure. Then we use Dijkstra's optimal route algorithm to construct the least cost HSR paths between all bilateral destinations. In the second steps, we use Kruskal's minimum spanning tree algorithm to generate a single continuous network subject to minimization of the global construction cost[20]. Figure 1 depicts the least cost path (LCP) network of HSR.

### 4.3 Estimation Results

Table 1: The Impact of HSR Network Connection on Export

| Dept. var | OLS | | IV Approach | | |
|---|---|---|---|---|---|
| $\ln y_{ct} - \ln y_{c0}$ | (1) | (2) | (3) Euclid IV | (4) Slope IV | (5) LCP IV |
| *Connection to HSR* | 0.529*** | 0.194*** | 0.994* | 1.705*** | 0.896* |
| | (0.0414) | (0.0615) | (0.507) | (0.489) | (0.486) |
| First stage F-Stat | - | - | 11.16 | 11.10 | 11.55 |
| Observations | 1,681 | 1,433 | 1,433 | 1,433 | 1,433 |
| R-squared | 0.422 | 0.574 | 0.506 | 0.332 | 0.522 |
| City FE | YES | YES | YES | YES | YES |
| Other Control | NO | YES | YES | YES | YES |

Notes: Each point estimate stems from a separate regression. All regressions include prefecture fixed effects. Euclid IV denotes the Euclidean distance spanning tree instrument. Slope IV stands for the cost path spanning tree instrument that uses average terrain slope gradient. LCP IV represents the least cost path spanning tree instrument. Other controls include per capita GDP, population, average rideship and internet coverage at prefectural level. Robust standard errors are clustered at the group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

In the conceptual framework, we hypothesize that connection to HSR promotes the manager-to-manager matching efficiency which allows firms to search better intermediate tasks suppliers and increases their motivation to outsourcing the manufacturing tasks. In consequence, firm productivity

[19] As our instrument variable, i.e., the least-cost-path HSR network, is time-invariant while we control for city fixed effect in the baseline regression, we instrument $HSR_{ct}$ using interaction term of least-cost-path HSR network and year dummies.

[20] The LCP based upon Euclidean distance and land slope gradients are constructed in the same method as explained above, but they differ in considering different land factors when calculating the construction cost. For instance, the LCP based on land slope gradients assigns higher construction costs to land parcels if it has a steeper slope.

rises and exporting performance improves. In this section, we directly test the implications of HSR connection on the improvement of exporting performance.

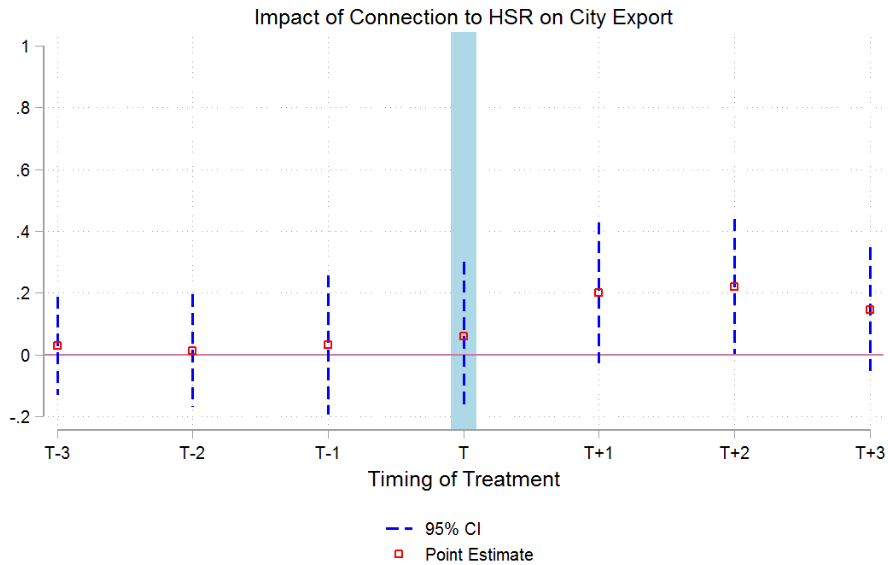
Table 1 presents the the difference-in-difference estimation results on the prefecture-level export. Controlling for city fixed effects, Column (1) reports the estimation of OLS, based on which connection to HSR significantly promotes a city's exporting performance by 52% annually. Column (2) to Column (4) reports the results for the least cost path instruments according to Euclidean distance,average gradient, and the constructing cost, respectively. The F-statistics of the first stage indicates that all these instruments are relevant to the actual HSR placements. As the point estimates remain significantly positive, we confirm that connection to HSR contributes to the city's export performance.

Figure 2 shows the event study results on city export specified by (2). Panel (a) displays the point estimates on the leads and lags of initial HSR connection using all samples (including the cities that are not planned to be connected by 2015). However, if the decision of HSR placement is based on past economic performance and the expected future growth, such assumption might be violated. To reduce the endogeneity issue of this type, we restrict the sample to the prefectures that are either connected or planned to be connected to HSR by 2015, the result of which case is displayed in Panel (b). It is clear that the coefficients of the leads are insignificant from zero, which supports the parallel trends assumption that the connection to HSR network does not correlate with previous trends in city's exporting performance. In the meanwhile, export grows gradually after being connected to HSR, and such effect is largest in the two-period lag. The treatment effect of HSR connection remains robust and is found to be stronger after we refine sample to the cities that are either planned or constructed by 2015, as shown in Panel (b). More robustness checks are presented in Table A.2 of the appendix[21], where we report the point estimates using the non-capital cities, as well as using the post financial crisis periods. The finding remains barely changed.

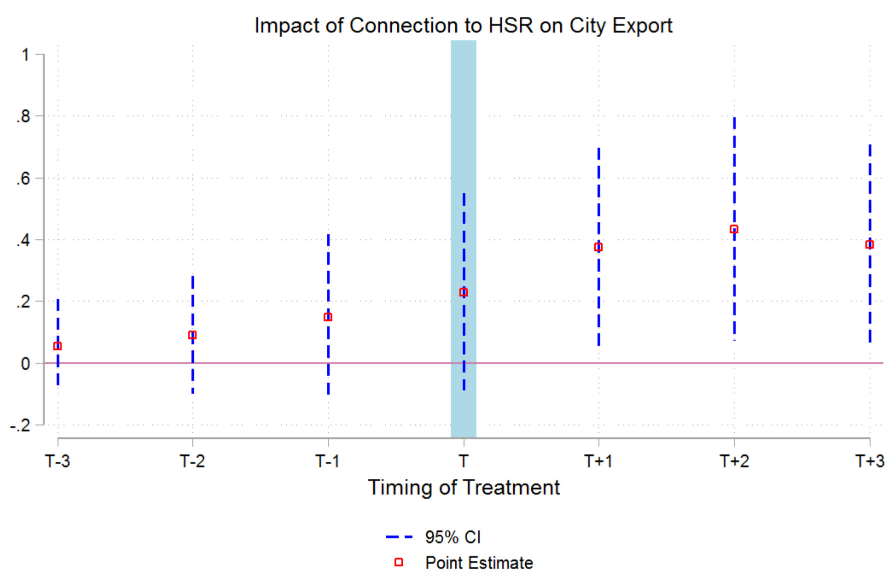
Besides the direct impact on the connected cities, access to HSR could also bring positive spillover effects to the cities which are not directly linked to HSR, since passengers from these unconnected cities could use HSR to travel to other destinations by transferring at nearby HSR hubs. To test the indirect impact of HSR connection, we consider prefecture-level city $c$ is initially "connected" to the HSR network in year $t$, if, among the cities that are located within a particular distance $d$ from $c$, the earliest connecting time is year $t$ . Given the broader definition of connection, we use the specification of (A.2) to estimate the HSR impact by distance $d$. Figure 3 displays the various point estimates of HSR impact by connecting time for different values of distance $d$. A coefficient greater than zero indicates the positive spillover effect of the connection of HSR hub city on the regions within the corresponding distance. Firstly, the event study of the pseudo-connection confirms the parallel trends assumption and the positive spillover effect of HSR connection, since the point estimates are only significantly positive during the post-connection period while remains insignificant before the actual connection. Secondly, the spillover effect is observed to decline with distance. As displayed in the

[21] Panel (a) and (b) Figure 2 corresponds the column (1) and (2) of Table A.2.

figure[22], the positive spillover effect is strongest for cities located within 60$km$ to 120$km$ to the HSR hub, which is sizably larger than the regions located farther away[23] (240$km$ to 300$km$). From the regression table, we also find that it takes a longer time to have a positive effect for the cities located farther from the HSR hubs, i.e., the point estimates become significantly positive after a three-year lag.




(a) The Overall Sample




(b) The Refined Sample

Figure 2: Event Study of Connection to HSR

[22]The detailed information refers to Table A.3 in the appendix.
[23]The positive effect of HSR almost vanishes for the long distance such as 300 $km$.

In our conceptual framework, HSR connection promotes export through the channel of productivity gains via the improved firm-to-firm matching efficiency. We distinguish the ordinary regime from the processing regime among export and uses the processing export as a placebo test. As processing trade regime firms are not subject to this benefit of HSR connection[24], we expect HSR connection would have no impact on the exporting performance of processing regime. We repeat the specification (A.2) and the point estimates are presented in Figure A.7[25]. Panel (a) reports the HSR impact on ordinary export, which is consistent with our earlier findings. Panel (b) reports the results of placebo test using processing export as the outcome variable. It is displayed that HSR connection does not have an impact on the processing activity, which verifies the causal effect of HSR connection on ordinary export increases.

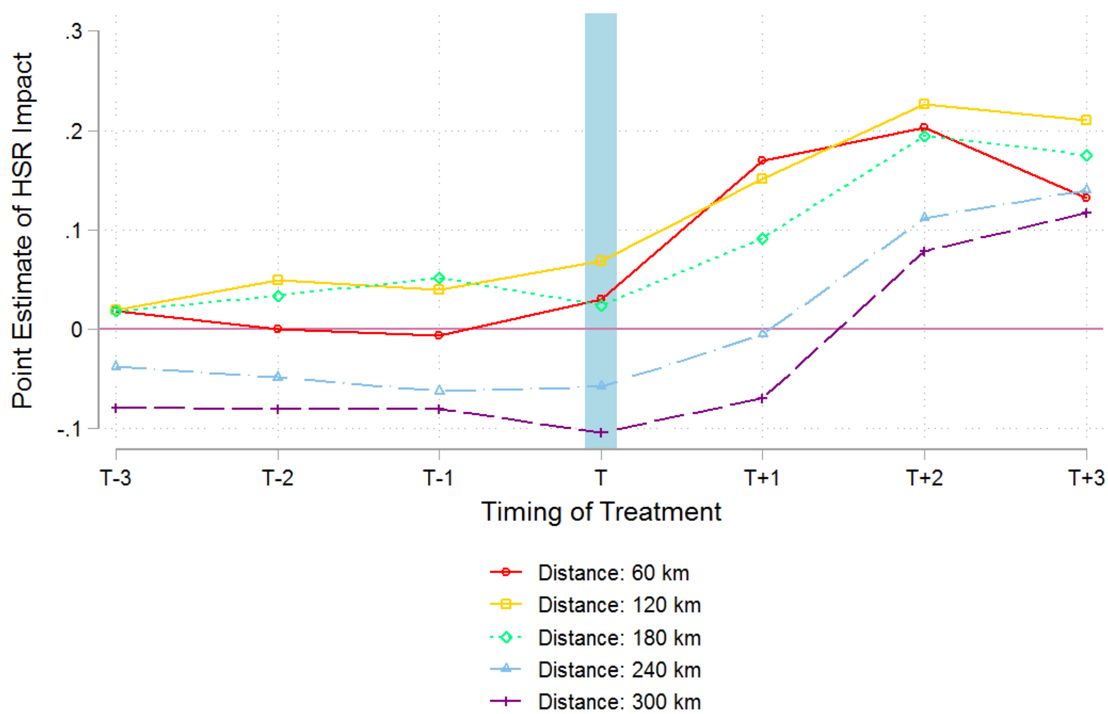



Figure 3: Spillover Impacts of HSR Connection by Distance

Summarizing, in this section, taking advantage of the rapid expansion of high-speed rails in China as a plausible source of exogenous shocks of improving firm-to-firm matching efficiency (by

[24]The matching process between foreign and domestic firms to contract a processing activity is far from being affected by HSR, since there are other major barriers to affect the formation of processing partnership such as incomplete contract environment, long distance between China and other countries and e.t.c.

[25]The detailed information are provided in the Table A.4 of appendix. The point estimates of Figure A.7 corresponds to the column (1) and (7) in Table A.4.

reducing the passenger travel cost) across Chinese cities over time, we find the connection to HSR contributes to the improvement of exporting performance at the city level. To deal with the endogenous railway placement issue, we follow Faber (2014) and adopt an instrumental variable strategy based upon the construction of least cost path spanning tree networks. Complementing with IV regression, we carry out the event study to rule out the violation of parallel trend assumption that HSR placement could be based on the past and expected future growth, and this further confirms the positive causal impact of HSR connection on regional exports. Our point estimates are robust to various instruments, the alternative specifications, and different samples. Besides the direct impact, we also find HSR connection has positive spillovers across the nearby regions as well, and such impact decreases with distance to HSR hubs. In line with Bernard, Moxnes, and Saito (2015), this section extends the literature by providing empirical evidence that access to HSR network substantially improves regional exporting performance. As far as we know, we are the first to explore the causal relationship between HSR connection and exports. This section motivates the following theoretical analysis that attributes the export growth to the productivity gains resulted from outsourcing.

## 5 The Model of Regional Outsourcing

In the theoretical study, we extend the firm-to-firm trade model of Eaton, Kortum, and Kramarz (2016) in a multi-region context by introducing an individual migration decision and labor market dynamics. The basic structure is as follows. Firms require service (e.g. firm management) and manufacturing tasks as inputs into the production process. The service tasks can only be performed by themselves and the manufacturing tasks can be outsourced to firms elsewhere[26]. The final output can be used for consumption or as the intermediate input to perform the manufacturing tasks for the other firms, and the two purposes can not be substitutable with each other. Firms looking for inputs connect with sellers in different locations randomly, which depends on the regional specific matching efficiency. An improvement in the regional matching efficiency (due to HSR connection) leads firms to search more and better suppliers, which decreases the marginal production cost and increases the overall productivity. The basic structure of the model is depicted in Figure 4. Since there is substantial productivity across regions, and since Hukou system restricts labor mobility, we assume the heterogeneous innate productivity across regions and introduce an individual migration decision to estimate the distributional effect of HSR among regions and different skill groups.

We consider an economy consisting of a discrete number $N$ regions, indexed by $c = 1, 2, ..., N$. Each region $c$ is initially populated with two types of worker (skilled and unskilled; white-collar and blue-collar), and the mass of workers are denoted as $\bar{L}_c^W$ and $\bar{L}_c^B$, respectively. Labor mobility across regions is allowed but subject to friction, which is specified later.

[26]We will release this assumption by allowing service tasks also to be outsourced with an intensity smaller than that of manufacturing tasks.

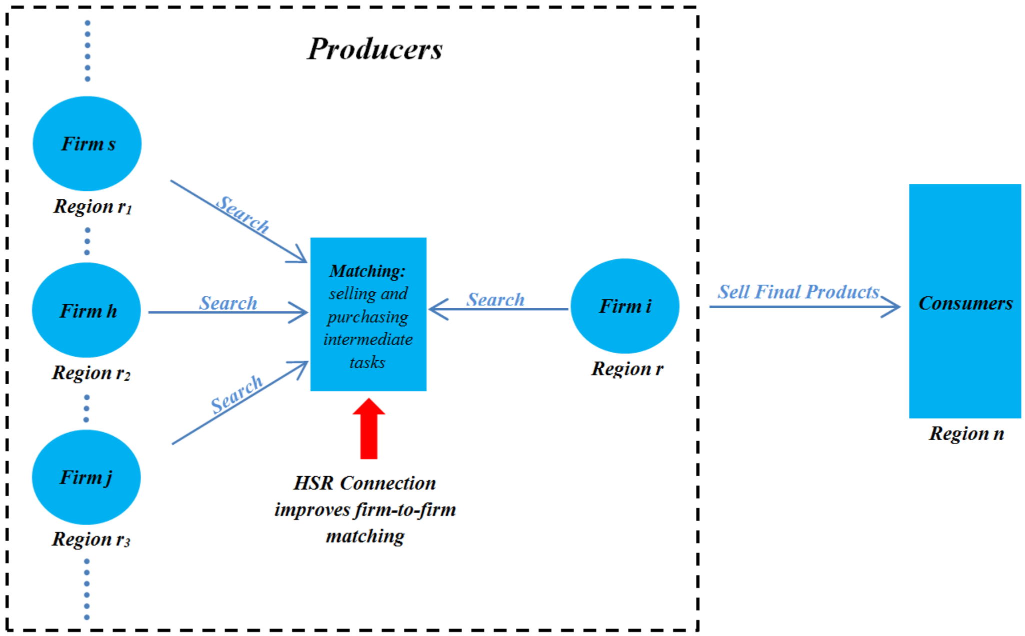



Figure 4: Model Diagram

## 5.1 Producer

### 5.1.1 Technology

A firm $j$ in location $i$ can produce a quantity of output $Q_i(j)$ by combining two intermediate tasks (service and manufacturing tasks, which are denoted as $S$ and $M$, respectively) according to the production function

$$Q_i(j) \quad = \quad z_i(j) \prod_{k \in \{S,M\}} b_k^{-1} \left( \frac{m_{k,i}(j)}{\beta_k} \right)^{\beta_k} \tag{3}$$

where $z_i(j)$ stands for producer $j$'s Hicks-neutral productivity, $m_{k,i}(j)$ denotes the input of task $k$, $b_k$ is a constant to be specified later, and $\beta_k$ is the Cobb-Douglas share of task $k$, which satisfy that $\beta_k > 0$ and $\beta_S + \beta_M = 1$.

The intermediate tasks can be performed either by the labor appropriate for that task or with an input produced by a firm. We assume the skilled labor is capable of producing the service task and the unskilled labor is capable of producing the manufacturing task. If firm uses labor to produce task $k$, it pays the wage $w_{k,i} \equiv w_i^{l(k)}$ for the labor of type $l(k) \in \{W, B\}$ where $k \in \{S, M\}$. The worker productivity of performing a task is assumed to be worker-firm specific and is denoted as $q_{k,i}(j)$, which is a random variable drawn from an extreme value distribution to be explained later. In the meanwhile, producer $j$ is also able to contact with a set of suppliers of the intermediate input that can be alternatively used to perform the tasks. In finding the intermediate inputs, buyers match with an integer number of suppliers due to search frictions. The variable profit is assumed to be split via Nash bargaining, and we assume the buyers have all the bargaining power[27] as Eaton, Kortum, and Kramarz (2016) for simplicity. We assume the labor and the intermediate inputs are perfect substitutes for performing the tasks, and firm chooses the optimal method of production to minimize the cost.

The unit cost firm $j$ pays to perform task $k$ depends on the labor wage and the lowest price available to it, which is denoted as $c_{k,i}(j)$ and expressed below

$$c_{k,i}(j) \quad = \quad min\left\{\frac{w_{k,i}}{q_{k,i}(j)},\, c_{k,i}^{min}(j)\right\} \tag{4}$$

where $c_{k,i}^{min}(j)$ is the lowest price of intermediate input available to firm $j$. The firm's marginal cost of delivering a unit of its output to destination $n$ is thus

$$c_{ni}(j) \quad = \quad \frac{d_{ni}}{z_i(j)}\prod_{k\in\{S,M\}}\frac{c_{k,i}(j)^{\beta_k}}{b_k} \tag{5}$$

where $d_{ni} \geq 1$ denotes the iceberg transportation cost of delivering one unit of final output sourcing from $i$ to $n$,with $d_{ii} = 1$ for $\forall i$. To make the result tractable, we follow the assumptions made by Eaton, Kortum, and Kramarz (2016) on the distributions for productivity, the labor productivity in performing a task as well as the distribution of prices of intermediate input.

Firstly, we assume the measure of potential producers in each region $i$ with productivity greater than $z$ ($z_i(j) \geq$z) is given by

$$\mu_i^Z(z) \quad = \quad T_i z^{-\theta} \tag{6}$$

where $T_i$ stands for the endowment of technology in region $i$ and $\theta \geq 0$ is the shape parameter governing the similarities, i.e, a larger value of $\theta$ implies the greater similarities.

Second, the worker productivity ($q_{k,i}(j)$) of performing task $k$ in firm $j$ is assumed to follow an extreme value distribution with $c.d.f.$ as

$$F(q) \quad = \quad exp(-q^{-\phi}) \tag{7}$$

where $\phi \geq 0$ reflects the similarity of worker's productivities across tasks and firms with restriction $\phi \leq \theta$.

Lastly, we will show that the measure of firms who sell in country $i$ with the marginal cost below $c$ is derived in form as given

$$\mu_i(c) \quad = \quad \Upsilon_i c^{\theta} \tag{8}$$

[27]It implies that the price of intermediate input is driven down to its unit cost.

where $Y_i$ is an endogenous variable reflecting the overall capacity of region $i$, including technology level, labor market condition, trade barriers as well as the efficiency of search for suppliers. In contrast, the specification of (6) and (7) are primitives of the model ($T_i$, $\theta$ and $\phi$ are exogenous parameters.).

### 5.1.2 Matching with Intermediate Input Suppliers

Matching between buyers and suppliers is random[28]. The encounter intensity with which a seller whose unit cost is $c$ encounters a buyer in region $n$ searching to fulfill the purpose $k$ (from a seller perspective) is

$$e_{k,n}(c) = \lambda_{k,n}\mu_n(c)^{-\gamma} \tag{9}$$

where $\mu_n(c)$ is the measure of firm whose marginal cost is below $c$ as defined in (8), $\gamma$ captures the degree of congestion by the lower cost sellers[29] with restriction $\gamma \in (0,1)$. The new parameter $\lambda_{k,n}$ governs the outsourcibility of task $k$[30], on which the access to high-speed Rail could insert an influence. Specifically, linking to HSR increases $\lambda_{k,n}$, as it becomes more convenient for a buyers to search sellers and supervise them so that the contracted intermediate inputs are proper for production usage. For tractability and realisticity, we assume that the service task cannot be outsourced such that $\lambda_{\text{Service},n} = 0$. The number of all the quotes that a buyer could receive for task $k$ with the price no greater than $c$ follows Poisson distribution. The Poisson parameter (the expected number of price quotes per buyers) is denoted as $\rho_{k,n}(c)$, which is derived by aggregating across all sellers in location $n$ given (8) and (9):

$$\begin{aligned}\rho_{k,n}(c) &= \int_0^c e_{k,n}(x)d\mu_n(x) \\ &= \frac{1}{1-\gamma}\lambda_{k,n}Y_n^{1-\gamma}c^{\theta(1-\gamma)}\end{aligned} \tag{10}$$

With restriction $\gamma \in (0,1)$, the expected number of price quotes below $c$ grows large with $c$, i.e., many less productive suppliers are able to serve the any given buyer which increases the expected number of price quotes.

Firm $j$ can perform task $k$ at a cost no greater than $c$ if both the labor cost and the lowest price quote of intermediate input is no less than $c$. The probability that firm $j$ in country $i$ is not able to encounter any suppliers with price lower than $c$ for task $k$ is $exp(-\rho_{k,i}(c))$ by Poisson distribution. On the other hand, the labor cost is no less than $c$ if $w_{k,i}/q_{k,i}(j) \geq c$ and the corresponding probability

[28]As fully explained in Eaton, Kortum, and Kramarz (2016), matching in this framework can be interpreted as coming into contact with intermediate input suppliers, and it could also relate to the appropriateness of the supplier's product for the buyer's purpose.

[29]When the number of firms offering lower cost increases, the probability of being able to match with a appropriate buyer for any supplier will decrease.

[30]In original paper, the authors interpret $\lambda_{k,n}$ as how easy it is for a seller to build the contact with a buyer for task $k$.

is $F(w_{k,i}/c)$ according to (7). As the two events are independent by assumption, the distribution of lowest cost of performing task $k$ in country $i$ is

$$
\begin{aligned}
G_{k,i}(c) &= 1 - exp(-\rho_{k,i}(c))F(w_{k,i}/c) \\
&= 1 - exp\left[-\left(\frac{1}{1-\gamma}\lambda_{k,i}Y_i^{1-\gamma}c^{\theta(1-\gamma)} + w_{k,i}^{-\phi}c^{\phi}\right)\right]
\end{aligned} \tag{11}
$$

We followEaton, Kortum, and Kramarz (2016) by restricting $\gamma$ as

$$
\gamma \equiv \frac{\theta - \phi}{\theta}
$$

under which, the parameters affecting the cost heterogeneity due to labor efficiency is the same with the parameters governing the heterogeneity due to intermediate inputs at a given task for a given buyer. The cost distribution for task $k$ can be further simplified as

$$
G_{k,i}(c) = 1 - exp\left(-\Xi_{k,i}c^{\phi}\right) \tag{12}
$$

where

$$
\Xi_{k,i} \equiv v_{k,i} + w_{k,i}^{-\phi} \tag{13}
$$

and

$$
v_{k,i} \equiv \frac{\theta}{\phi}\lambda_{k,i}Y_i^{\phi/\theta} \tag{14}
$$

$\Xi_{k,i}$ reflects the technology efficiency of performing task $k$ that is influenced by both labor cost ($w_{k,i}^{-\phi}$) and intermediate input efficiency ($v_{k,i}$). The lower labor cost ($w_{k,i}$) and the greater overall capacity ($Y_i$) of region $i$ implies the high level of efficiency ($\Xi_{k,i}$) of performing task $k$. Notice that a risen matching efficiency $\lambda_{k,i}$in region $i$, due to being connected to HSR, will raise the productivity of producing task $k$ as well.

**Proposition 1.** *The probability of fulfilling task k by hiring workers is $w_{k,i}^{-\phi}/\Xi_{k,i}$, while the firm outsources the task from the lowest-cost supplier with probability $v_{k,i}/\Xi_{k,i}$.*

The proposition gives us the comparative statics. We see directly that the probability of outsourcing the production of task $k$ increases when there is improvement of matching efficiency (a higher value of $\lambda_{k,i}$) or higher labor cost ($w_{k,i}$). An important caveat is that the comparative statics on matching efficiency take the wage as given, which doesn't take the general equilibrium into consideration. As the service tasks cannot be outsourced $\lambda_{S,i} = 0$, the share of labor force in service task is equal to unity as expected.

Note that $w_{k,i}^{-\phi}/\Xi_{k,i}$ is also the aggregate labor share in producing task $k$ in region $i$. The aggregate share of labor of type $l$ in the total production cost in region $i$ is

$$\beta_i^{l(k)} = \beta_k \times \frac{w_{k,i}^{-\phi}}{\Xi_{k,i}} \tag{15}$$

Consequently, the overall labor share of in the production in region $i$ is consequently written as:

$$\begin{aligned}\beta_i^L &= \sum_{k\in\{S,M\}} \beta_i^{l(k)} \\ &= \beta_S + \beta_M \frac{w_{M,i}^{-\phi}}{\Xi_{M,i}}\end{aligned} \tag{16}$$

Even though the basic assumption of production function is Cobb-Douglas, the overall labor share depends on the manufacturing wages and other factors.

#### 5.1.3 Cost Distribution

We now turn to the cost distribution of firm $j$ in region $i$, given that task specific cost follows the distribution shown in (12). The marginal cost of firm depends on the efficiency of each task given in (4). The expected measure of firms from location $i$ that can deliver their output to destination $n$ at a cost no greater than $c$ is derived as (derivation is provided in the Appendix):

$$\mu_{ni}(c) = T_i \Xi_i d_{ni}^{-\theta} c^{\theta} \tag{17}$$

where $\Xi_i$ is the technology parameter reflecting the efficiency gain due to outsourcing, and $\Xi_{k,i}$ is defined in (13):

$$\Xi_i \equiv \prod_{k\in\{S,M\}} (\Xi_{k,i})^{\theta\beta_k/\phi} \tag{18}$$

Aggregating across all the sourcing cities $i$, the measure of potential sellers in region $n$ that can deliver a unit of good at a cost below $c$ is

$$\mu_n(c) = \sum_i \mu_{ni}(c) = \Upsilon_n c^{\theta} \tag{19}$$

where $\Upsilon_n \equiv \sum_i T_i \Xi_i d_{ni}^{-\theta}$. This expression is consistent with the specification in (8). We also learn from the definition that the overall capacity of region $i$ depends on the exogenous technology efficiency $T_i$, trade costs $d_{ni}$, wages $w_{k,i}$ as well as the matching efficiency $\lambda_{k,i}$. Combining (13), (14) and the definition of $\Upsilon_n \equiv \sum_i T_i \Xi_i d_{ni}^{-\theta}$, we can solve the vector of $\Upsilon_n$ by solving the fixed point of the following system of equations

$$\Upsilon_n = \sum_i T_i d_{ni}^{-\theta} \prod_{k\in\{S,M\}} \left( \frac{\theta}{\phi} \lambda_{k,i} \Upsilon_i^{\phi/\theta} + w_{k,i}^{-\phi} \right)^{\theta\beta_k/\phi} \tag{20}$$

In the appendix, we show that $\Upsilon_n$ can be computed by iterative procedure under our assumptions. The measure of sellers sourcing from $i$ with a cost below $c$ in destination $n$ is $\mu_{ni}(c)$ defined in (17), and (19) is the total measure of active firms in $n$. Therefore, we are able to calculate the probability that the potential seller selling in $n$ at a unit cost below $c$ is from $i$ is:

$$\pi_{ni} = \frac{T_i \Xi_i d_{ni}^{-\theta}}{\sum_j T_j \Xi_j d_{nj}^{-\theta}} \tag{21}$$

With the assumption of continuum of producers, $\pi_{ni}$ is the share of the source $i$ in the purchases of destination $n$.

Note that when firm is not allowed to outsource the tasks $\lambda_{k,i} = 0$ for $\forall k$ and $\phi$ approaches to $\theta$, the total measure of active firm in region $n$ is

$$\mu_n = \sum_i T_i \left(\bar{w}_i d_{ni}\right)^{-\theta} c^\theta \tag{22}$$

where $\bar{w}_i$ is the average wage of region $i$ that is defined as $\bar{w}_i \equiv \prod_{k\in\{S,M\}} w_{k,i}^{\beta_k}$. The trade share of $i$ in market $n$ is

$$\pi_{ni} = \frac{T_i \left(\bar{w}_i d_{ni}\right)^{-\theta}}{\sum_j T_j \left(\bar{w}_j d_{nj}\right)^{-\theta}} \tag{23}$$

Both (22) and (23) are consistent with the results of Eaton and Kortum (2002) and Eaton, Kortum, and Kramarz (2011) .

**Proposition 2.** *The above assumptions over technology are consistent with an aggregate production function for region i, which satisfies the function form:*

$$Q_i = \left(L_{S,i}\right)^{\beta_S} \left[\left(\psi \left(L_{M,i}\right)^{\phi/(\phi+1)} + (1-\psi)\left(I_{M,i}\right)^{\phi/(\phi+1)}\right)^{(\phi+1)/\phi}\right]^{\beta_M} \tag{24}$$

*where $L_{k,i}$ denotes the total labor force employed in fulfilling task $k \in \{M, S\}$, $I_{M,i}$ are the intermediates used for manufacturing tasks, and $\psi \equiv 1/\left[1+\Gamma(1+1/\phi)^{\phi/(1+\phi)}\right]$.*

Equation (24) provides some intuitions on how the HSR affect the aggregate production schedule. As shown in the appendix, the average price of intermediate input, $\bar{p}_{M,i}$, is negatively related with $\frac{\theta}{\phi}\lambda_{k,i}Y_i^{\phi/\theta}$. An improvement in matching efficiency $\lambda_{k,i}$ will decrease the average costs of using intermediate tasks, which can be regarded as the composite input-biased technology shocks that reduce the demand for unskilled workers all else equal. As a result, it unambiguously raises skill premium and widen the welfare inequality between skilled and unskilled workers.

### 5.2 Consumer

Workers are risk neutral with a linear preference in a consumption index: $U_{odi} = C_{odi}$, where $C_{odi}$denotes the consumption index for worker $i$, born in region $o$ (origin) and working in location $d$ (destination). The consumption bundle depends on the consumption of the service goods ($C_{S,odi}$); consumption of the manufacturing goods ($C_{M,odi}$); and an idiosyncratic shock that is specific to the individual worker and varies with the worker's born and working places ($z_{odi}$). This idiosyncratic shock captures the idea that workers could have idiosyncratic reasons for migrating to different regions. Particularly, we assume the service goods and manufacturing goods are destination specific, i.e., $C_{k,odi} = C_{k,di}$ $\forall k \in \{S, M\}$; the aggregate consumption bundle is assumed to take the Cobb-Douglas form:

$$C_{odi} = \kappa z_{odi} C_{S,di}^{\alpha_S} C_{M,di}^{\alpha_M} \tag{25}$$

where $\kappa$ is a constant[31], $\alpha_k$ denotes the Cobb-Douglas share of need $k$, with $\alpha_k > 0$ and $\alpha_M + \alpha_S = 1$.

Different from Eaton, Kortum, and Kramarz (2016), we assume consumers are only able to fulfill the demand by purchasing from the local active firms. The meeting of consumers and firms are subject to matching frictions, and we posit that the intensity with which a firm who charge price $p$ in region $d$ encounters a consumer seeking to fulfill the need of $k$ is

$$\tilde{e}_{k,d}(p) = \tilde{\lambda}_{k,d} \mu_d(p)^{-\tilde{\gamma}} \tag{26}$$

where $\tilde{\lambda}_{k,d}$ governs the efficiency of matching between firms and consumers for the need $k \in \{S, M\}$ in region $d$, and $\tilde{\gamma}$ captures the extents to which lower price firms reduce the ability of a firm to match with a consumer, $\mu_d$ denotes the measure of active firms in region $d$, as defined in (19). We aggregate across the measure of potential firms of different prices, and the number of potential quotes that a consumer receives for need $k$ with price below $p$ follows Poisson distribution, which is similar with the assumption on firms. The Poisson parameter (the average number of price quotes) is given as:

$$\begin{aligned} \tilde{\rho}_{k,d}(p) &= \int_0^p \tilde{e}_{k,d}(x) d\mu_d(x) \\ &= \frac{1}{1-\tilde{\gamma}} \tilde{\lambda}_{k,d} Y_d^{1-\tilde{\gamma}} p^{\theta(1-\tilde{\gamma})} \end{aligned} \tag{27}$$

where we require $\tilde{\gamma} < 1$ such that the expected number of received price quotes increases with the unit cost. According to Poisson distribution, the probability that a consumers is not able to encounter a firm with price lower than $p$ is $exp\left(-\tilde{\rho}_{k,d}(p)\right)$. Therefore, the price distribution for need $k$ is expressed as

$$\tilde{G}(p) = 1 - exp\left[-\tilde{v}_{k,d} p^{\theta(1-\tilde{\gamma})}\right] \tag{28}$$

where $\tilde{v}_{k,d} \equiv \frac{1}{1-\tilde{\gamma}} \tilde{\lambda}_{k,d} Y_d^{1-\tilde{\gamma}}$ is related with how easy it is for a consumer to be able to purchase the product with low prices.

Perceiving the idiosyncratic tastes $z_{odi}$, the indirect utility of a consumer of type $l \in \{W, B\}$ facing the prices of each need $k$ ($p_S$ and $p_M$) is given as

$$V_{odi}(l) = \frac{z_{odi} w_d^l}{\prod_{k \in \{S,M\}} \left(a_k \times p_k^{\alpha_k}\right)}$$

[31] The constant is used to re-scale the expression to be more simplified, which is defined as $\kappa \equiv \prod_{k \in \{S,M\}} \alpha_k^{-\alpha_k} \times a_k^{-1}$, and $a_k$ are also constant to be explained latter.

where $w_d^l$ denotes the wage of worker of type $l \in \{W, B\}$ in location $d$, defined as $w_d^l \equiv w_d^{l(k)} = w_{d,k}$. Parameter $a_k$ is a constant chosen to eliminate the effect of need $k$ on utility. As all the prices are random variables, after perceiving taste shocks $z_{odi}$, consumer $i$ only cares about their destination specific component of indirect utility in location $d$. We denote this term as $\nu_d^l$ which could be expressed as (derivation is provided in the Appendix):

$$\nu_d^l = B_d \times w_d^l \times \Upsilon_d^{\frac{1}{\theta}} \tag{29}$$

The first exogeneous term[32] in (29) $B_d$ reflects region's capability of supplying final consumption, i.e., how easily a consumer manage to find a supplier to satisfy his demand[33]; the second term $w_d^l$ stands for the income effect, i.e., cities with higher income generate a higher level of utility for both types of workers; lastly, $\Upsilon_d$ captures all the other factors influencing a region's amenity. As discussed in the appendix, all else equal, a higher level of technology $T_d$ and greater probability of matching between buyers and sells $\lambda_{k,d}$ leads to greater value of $\Upsilon_d$ and hence higher level of welfare. On the other hand, higher wages will reduce the attractability of region by lowering $\Upsilon_d$, and this will partly offset the income effect and decreases the welfare. Specifically, if we further assume the $\tilde{\lambda}_{k,d}$ is common to all cities, i.e., $\tilde{\lambda}_{k,d} = \tilde{\lambda}_k$ and smartly choose $a_k = \tilde{\lambda}_k^{\frac{\alpha_k}{\theta(1-\tilde{\gamma})}} \times (1-\tilde{\gamma})^{-\frac{\alpha_k}{\theta(1-\tilde{\gamma})}} \times \Gamma\left(1 - \frac{\alpha_k}{\theta(1-\tilde{\gamma})}\right)$ , the expected value of indirect utility of living in region $d$ can be simplified as

$$\nu_d^l = w_d^l \times \Upsilon_d^{\frac{1}{\theta}} \tag{30}$$

which provides a more simplified welfare expression.

### 5.2.1 Migration Decision Across Regions

Migration is modeled as a once-for-life choice. We model the heterogeneity in the utility that workers obtain from living in different regions following Ahlfeldt, Redding, Sturm, and Wolf (2015). Upon birth, workers learn their idiosyncratic taste of living across all the regions and decide where to work, taking into account their destination specific component in indirect utility of destination $d$, $\nu_d^l$, as well as the migration cost that they will have to incur. The migration cost, denoted as $e_{od}^l$, are both skill-specific $l \in \{B, W\}$ and source-destination specific.

Formally, taking the idiosyncratic component taste across regions as given, $\{z_{odi} | d = 1, 2, ..., N\}$, worker $i$ born in region $o$ chooses to live in region $d$ to maximize the welfare:

$$\max_{d \in \{1,2,..,N\}} \left\{ \frac{z_{odi} \nu_d^l}{e_{od}^l} \right\} \tag{31}$$

where $\nu_d^l = B_d w_d^l \Upsilon_d^{\frac{1}{\theta}}$ denotes the amenity-adjusted expected real wage rate in region $d$ for workers of type $l \in \{B, W\}$, born in $o$ and working in $d$. Note that the migration cost can also be interpreted as

[32] $B_d$ is defined as $B_d \equiv \left[\prod_{k \in \{S,M\}} \tilde{\lambda}_{k,d}^{\alpha_k}\right]^{\frac{1}{\theta(1-\tilde{\gamma})}}$.

[33] For example, it is positively correlated with the development level of retailing technology.

how workers of type $l$ discount income from destination[34], and this cost is similar to the iceberg cost assumption used in international trade literature. For ease of notation, worker $i$ of type $l \in \{B, W\}$, born in region $o$, will move to region $d$ if and only if the migration gives the highest utility:

$$\frac{\nu_d^l z_{odi}}{e_{od}^l} \geq \frac{\nu_g^l z_{ogi}}{e_{og}^l},\ \forall g \in \{1, 2, ..., N\} \tag{32}$$

Following Ahlfeldt, Redding, Sturm, and Wolf (2015), we assume $\{z_{odi} | d = 1, 2, ..., N\}$are drawn from the Fréchet distribution that are independent across the birth regions ($o$) and individuals ($i$)[35]. In particular, we allow that the individual-specific component in worker's taste are correlated across regions: some people may like living in many cities more than others. Specifically, the vector of idiosyncratic taste shocks for any given worker is generated from the following $c.d.f$ of Fréchet

$$F(z_{1i}, z_{2i}, ..., z_{Ni}) = exp\left[-\left(\sum_d z_{di}^{-\epsilon_l}\right)^{1-\rho}\right] \tag{33}$$

where $\epsilon_l$ measures the extent of taste dispersion (dispersion increases as $\epsilon_l$ decreases) for workers of type $l$, $\rho$ reflects the inter-region correlation of taste draws[36]. If $\rho = 0$, taste shocks are uncorrelated across regions while if $\rho = 1$, the tastes shocks are perfectly correlated for that person. Under this assumption, the probability that a worker of type $l$ from origin $o$ moves to destination $d$ (derivation is provided in the Appendix), is:

$$\begin{aligned} \delta_{od}^l &= \mathbf{Pr}(\frac{\nu_d^l z_{di}}{e_{od}^l} \geq \frac{\nu_g^l z_{gi}}{e_{og}^l},\ \forall g) \\ &= \frac{\left(\frac{\nu_d^l}{e_{od}^l}\right)^{\epsilon_l}}{\sum_g \left(\frac{\nu_g^l}{e_{og}^l}\right)^{\epsilon_l}} \end{aligned} \tag{34}$$

Denote $L_d^l$ as the number of workers of type $l$ who are working in $d$, and $\bar{L}_o^l$ as the number of workers of type $l$ who are born in region $o$. Then the number of workers of type $l$ moving from $o$ to $d$ is $L_{od}^l \equiv \delta_{od}^l \bar{L}_o^l$, and we derive the supply of labor of type $l$ in region $d$ as

$$L_d^l = \sum_o L_{od}^l = \sum_o \delta_{od}^l \bar{L}_o^l \tag{35}$$

**Proposition 3.** *For any set of migration cost* $\{e_{od}^l\}$*,there exist a unique set of* $\{\nu_d^l\}$ *(depending on the normalization) such that the model generated number of workers employed in each cities equals the employment in data, i.e,* (35) *is satisfied, where* $\bar{L}_o^l$ *denotes the number of workers who are originally from region o (data),*

[34]We model the migration cost as variable cost for simplicity, while in reality there are both fixed and variable cost when migrating across cities.

[35]For ease of notation, denote $z_{di} = z_{odi}$.

[36]The parametric assumption on distribution is also used by Hsieh, Hurst, Jones, and Klenow (2013); Ahlfeldt, Redding, Sturm, and Wolf (2015); Bryan and Morten (2015).

*and $L^l_d$ is the number of workers who are employed in region d (data), and $\delta^l_{od}$ is the probability that workers from region o to migrate to region d which are generated by model (model).*

Finally, we derive the expected utility of born in $o$ for worker of type $l$, which is denoted as $E(u^l_o)$:

$$E(u^l_o) = \left(M^l_o\right)^{\frac{1}{\epsilon_l}} \Gamma\left[1 - \frac{1}{(1-\rho)\epsilon_l}\right] \tag{36}$$

where $M^l_o \equiv \sum_g \left(v^l_g / e^l_{og}\right)^{\epsilon_l}$ measures the welfare of being born in location $o$. The more connected location $o$ is to the labor markets of other region (smaller $e^l_{og}, \forall g$) and the more attractive the nearby cities are (greater $v^l_g, \forall g$), the higher utility the worker of type $l$ born in region $o$ benefit. Note that, the expected utility in (36) does not depend on the destination region $d$ for workers from the same region whose average welfare will be the same regardless of the location where they live. On the one hand, more attractive destination characteristics directly raise the welfare of a worker given his idiosyncratic taste draw, which increases the expected utility. On the other hand, more attractive destination characteristics attract workers with lower idiosyncratic taste draws, which reduces the average utility. With a Fréchet distribution of taste shocks, these two effects just cancel out one another for workers born in the same place, which only depends on the characteristics of the original region. As the migration is costly, the expected utility do not necessarily equal across regions, and this implies we are also able to capture the policy implications on regional disparity under this framework. It is straightforward to show that the aggregate average welfare is:

$$W = \sum_{l\in\{B,W\}} \sum_{i=1}^{N} m^l_i \times \left(M^l_i\right)^{\frac{1}{\epsilon_l}} \Gamma\left[1 - \frac{1}{(1-\rho)\epsilon_l}\right] \tag{37}$$

where $m^l_i = \bar{L}^l_i / \left(\sum_{l\in\{B,W\}} \sum_{s=1}^{N} \bar{L}^l_s\right)$ is the original share of worker of type $l$ in region $i$.

## 5.3 Equilibrium

We now solve the for the aggregate equilibrium given the structural described above. To start, we solve for the equilibrium of intermediates products, taking labor cost as given, and then turn to the labor market equilibrium.

### 5.3.1 Production Equilibrium

With balanced trade, the total expenditure on final goods is equal to labor income, since there is no profit in this model. For region $n$, we have

$$X^C_n = \sum_{l\in\{B,W\}} w^l_n L^l_n \tag{38}$$

The aggregate production in region $i$ consists of the total revenue from selling the consumption goods and that in supplying the intermediates around the world:

$$Y_i = \sum_{n=1}^{N} \pi_{ni} \left[ X_n^c + \Phi_n^I Y_n \right] \tag{39}$$

where $\Phi_n^I \equiv 1 - \beta_n^L$ denotes the shares of intermediate inputs in final production, and $\beta_n^L$ denotes the overall labor share of in production costs that is defined in (16).Writing (39) in matrix format (the details are listed in appendix), we derive

$$\mathbf{Y} = \mathbf{\Pi}' \left( \mathbf{X}^C + \mathbf{\Phi}^I \mathbf{Y} \right)$$

We can further solve for the total production $\mathbf{Y}$ as

$$\mathbf{Y} = (\boldsymbol{I} - \mathbf{\Pi}' \mathbf{\Phi}^I)^{-1} \mathbf{\Pi}' \mathbf{X}^C \tag{40}$$

where $\boldsymbol{I}$ is the $N \times N$ identity matrix.

### 5.3.2 Labor Market Equilibrium

With balanced trade, the final expenditure $X_i^C$ in region $i$ is given by (38), and the equilibrium for labor of type $l \in \{B, W\}$ in region $i$ implies:

$$w_i^l L_i^l = \beta_i^l Y_i \tag{41}$$

where $\beta_i^l$ denotes the production cost share of labor of type $l$ in region $i$, as given in (15); $L_i^l$ is the endogenous supply of workers of type $l$ in region $i$, provided in (35). These sets of equations for each type of labor $l$ in region $i$ solves the wage $w_i^l$.

### 5.3.3 Definition of Equilibrium

To summarize, the parameters used in the model are the following: preference parameters, including $\{\alpha_M, \alpha_S\}$, spatial frictions, including matching parameters $\tilde{\gamma}$ and $\{\lambda_{M,i}, \tilde{\lambda}_{S,i}, \tilde{\lambda}_{M,i}\}$, migration related parameters $\rho$, $\{\epsilon_l\}$ and $\{e_{od}^l\}$, and trade cost $\{d_{ni}\}$; production technology, including $\{\beta_M, \beta_S, \theta, \phi\}$ and $\{T_i\}$; and initial labor endowment $\{\bar{L}_i^l\}$.

**Definition 1.** *A competitive equilibrium of the economy is defined as a set of prices and allocations such that the following conditions are satisfied:*

1. *The migration decision for workers of both type are optimal, i.e.,* (32) *and* (34) *are satisfied.*
2. *The distribution of labor $L_i^l$, and the final consumption expenditure $X_i^C$ are consistent with the endogenous supply of workers, i.e.,* (35) *and* (38).
3. *The decisions of firm's production are optimal.*
4. *The decisions of worker's consumption are optimal.*
5. *Labor markets and goods market clear* (39) *and* (41).

## 6 Model Calibration

Before using the model to conduct the counterfactual policy experiments, we calibrated the model to the 2007 equilibrium. This section describes the detailed steps in calibrating the main model parameters, such as migration cost, trade costs, technology parameters, as well as matching efficiency parameters. Besides these, we also estimate the aggregate the parameters governing production function ($\beta_S$, $\beta_S$ and $\phi$), and the steps are explained in the appendix.

### 6.1 Estimation of Migration Cost

We parameterize the migration cost of $l \in \{W, B\}$ type worker to move from $o$ to $d$ as:

$$\ln e^l_{od} = \beta^l_1 \times I_1 + \beta^l_2 \times I_1 \times Dist_{od} + \beta^l_3 \times I_2 + \beta^l_4 \times I_2 \times Dist_{od} + \beta^l_5 \times I_3 + \beta^l_6 \times I_3 \times Dist_{od} \quad (42)$$

where $I_1$ to $I_3$ are the mutually exclusive dummy variables. Specifically, $I_1$ indicates if $d$ is neighboring province of $o$ within the same region; $I_2$ indicates if $o$ and $d$ are two different provinces within the same region that do not share the province boundary; $I_3$ is the dummy variable which equals one if $o$ and $d$ are in the different regions; $Dist_{od}$ captures the geographic barriers to migration which is measured as the centroid distance between $o$ and $d$. Dummy variables $I_1$ to $I_3$ captures the structural difference in migration barriers.

We infer the migration cost from the optimal migration decision as shown in equation (34). To do so, we divide $\delta^l_{od}$ by $\delta^l_{oo}$ and takes logs on the ratio to obtain:

$$\ln \frac{\delta^l_{od}}{\delta^l_{oo}} = \epsilon_l \times (\underbrace{\ln \nu^l_d - \ln \nu^l_o}_{\text{Fixed Effects}} - \underbrace{\ln e^l_{od}}_{\text{Migration Cost}} + \underbrace{\ln e^l_{oo}}_{0}) \quad (43)$$

we substitute (42) into (43) and estimated it using the linear regression[37]:

$$\ln \frac{\delta^l_{od}}{\delta^l_{oo}} = \tilde{\beta}^l_1 \times I_1 + \tilde{\beta}^l_2 \times I_1 \times Dist_{od} + \tilde{\beta}^l_3 \times I_2 + \tilde{\beta}^l_4 \times I_2 \times Dist_{od} + \tilde{\beta}^l_5 \times I_3 + \tilde{\beta}^l_6 \times I_3 \times Dist_{od} + \tilde{\nu}^l_d - \tilde{\nu}^l_o \quad (44)$$

Specification (44) provides us with the consistent estimates of $\{\tilde{\beta}^l\}$. However, it also shows clearly that we are not able to estimate $\tilde{\nu}^l_d$ due to multicollinearity, as well as separately identify the preference dispersion $\epsilon_l$ from the migration parameters $\{\beta^l\}$. To recover $\epsilon_l$, $\{\beta^l\}$ and $\{\nu^l_d\}$, we adopt the method of nonlinear least squares to minimize the difference between the model predicted migration flow in 2005 and their counterparts in data. Formally, let $l^l_o$ be the number of workers of type $l$ who are originally from province $o$, and recall that $\delta^l_{od}$ is the model predicted share of workers who move from $o$ to $d$ (equation (34)), then $\delta^l_{od} \times l^l_o$ is the model predicted flow of workers from $o$ to $d$. Denote

[37] Specifically, $\tilde{\beta}^l_j = -\epsilon_l \beta^l_j, \forall j = 1, 2, ..., 6$, $\tilde{\nu}^l_d = \epsilon_l \times \ln \nu^l_d$. We impose the symmetry in origin and destination fixed effects, as implied by the model.

$l^l_{od}$as the number of workers moving from $o$ to $d$ in the data, then we jointly estimate $\epsilon_l$ and $\{\nu^l_d\}$ by minimizing the difference between model predicted flow and that in data:

$$\min_{\epsilon_l,\{\nu^l_d\}} \sum_o \sum_d \left( \ln(\delta^l_{od} \times l^l_o) - \ln l^l_{od} \right)^2 \tag{45}$$

The estimation follows a nested procedure. In the inner loop of the procedure, for any given $\epsilon_l$, we firstly calculate $\beta^l_j = -\tilde{\beta}^l/\epsilon_l$, and use them to recover the migration cost for any province pairs $\{e^l_{od}\}$ according to the formula of (42). Then we solve $\{\nu_d\}$ so that the model migration flow predicted by the model is same as that in the data, i.e., $\sum_o l^l_o \times \delta^l_{od} = \sum_o l^l_{od}$. Once we obtain $\{\nu^l_d\}$ , we evaluate the objective function to calculate the sum of the residuals (which depends on $\epsilon_l$). In the outer loop, we then search over the parameter $\epsilon_l$ to minimize the objective function. Proposition (3) guarantees the feasibility of this approach by showing the existence and uniqueness of the $\{\nu^l_d\}$ for the inner loop.

## 6.2 Joint Estimation of Trade Cost, Regional Technology and Regional Outsourcing Ability

### 6.2.1 Parameterize Trade Cost and Outsourcing Ability

We parameterize the trade cost following the gravity literature in international trade. The first line of (46) specifies the the trade cost between any two provinces within China, and the second line is for the trade cost between any province in China and the rest of the world (ROW):

$$\ln d_{ni} = \begin{cases} \gamma_1 D_1 + \gamma_2 D_1 \times Dist_{ni} + \gamma_3 D_2 + \gamma_4 D_2 \times Dist_{ni} + \gamma_5 D_3 + \gamma_6 D_3 \times Dist_{ni}, & \text{if } n, i \in \mathbf{P} \\ \ln d_{j(i)i} + \kappa_{j(i)}, & \text{if } n \text{ is } ROW \end{cases} \tag{46}$$

where $\mathbf{P}$ denotes the collection of Chinese provinces, and $D_1$ to $D_3$ are the same dummy variables used in estimating the migration cost. Specifically, $D_1$ indicates if $n$ is neighboring region of $i$ within the same region; $D_2$ indicates if $n$ and $i$ are two different provinces within the same region that do not share the province boundary; $D_3$ is the dummy variable which equals one if $n$ and $i$ are in the different regions; $Dist_{ni}$ are the same distance measure. In the second line, $j(i)$ denotes the nearest coastal province to region $i$ and $\kappa_{j(i)}$ are the parameters to be estimated. Particularly, we model the trade cost between any given Chinese province and the ROW as the trade cost between that province and its nearest coastal province[38] ($\ln d_{nj(i)}$), plus a parameter ($\kappa_{j(i)}$) that captures all the factors, such as tariff and institutional quality, which could influence the international trade flow. We allow the international trade cost parameter to differ across these coastal provinces[39].

Next, we follow the assumption that only production tasks can be outsourced ( i.e.$\lambda_{S,i} = 0$ and $\lambda_{M,i} > 0$ for any region $i$), as described in the model. Access to HSR increases $\lambda_{M,i}$, as it becomes more convenient for a buyers to search sellers and supervise them so that the contracted intermediate

[38]There are ten coastal provinces that have ports: Liaoning, Tianjin, Hebei, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong and Guangxi.

[39]The trade cost specification is similar to the ones adopted in Fan (2015) and Ma and Tang (2016).

inputs are proper for production usage. To capture such influence and to quantify the resulting welfare change, we parameterize $\lambda_{M,i}$ as a log linear-function of the number of HSR hubs in that province:

$$\ln \lambda_{M,i} = \begin{cases} a + b \times N_i^{HSR}, & \text{if } i \in \mathbf{P} \\ a_W, & \text{if } i \text{ is } ROW \end{cases} \tag{47}$$

where $a$ denotes the original average outsourcing ability at national level[40] (we use 2007 as the initial time when $N_i = 0$, $\forall i \in \mathbf{P}$), and $b$ denotes the marginal improvement of outsourcing ability by linking additional one city into the HSR network. We assume that the matching efficiency of the ROW is relative stable that is captured by the scalar $a_W$ in (47), during our sample period 2007 to 2015.

We adopt a nested procedure to calibrate trade cost, regional technology as well as regional outsourcing ability. Specifically, we start with an initial guess for international parameters ($\{\kappa_{j(i)}\}$, $\lambda_{M,W}$, $T_w$). Conditional on the international parameters, we guess a distribution of regional technology and outsourcing ability ($\{T_i, \lambda_{M,i}\}_{i\in\mathbf{P}}$), with which we solve the model for the expected value of indirect utility of each types of workers ($\nu_i^l$) and the overall regional efficiency ($T_i\Xi_i$). We choose $\{T_i, \lambda_{M,i}\}_{i\in\mathbf{P}}$ such that the model generated expected value of indirect utility and the relative regional efficiency (relative to Xinjiang Uighur Autonomous Region) are closest to the data. For each set of international parameters ($\{\kappa_{j(i)}\}$, $\lambda_{M,W}$, $T_w$), we are able to solve for an optimal intra-national parameters ($\{T_i, \lambda_{M,i}\}_{i\in\mathbf{P}}$). In the outer loop, we search for the international parameters such that the model predicted provincial openness is closest to the data. The algorithm of joint estimation is demonstrated in Figure A.8.

### 6.2.2 Calibrating Intra-national Parameters

Similarly to the migration cost estimation, we infer the internal trade cost from the trade share equation (21). We divide $\pi_{ni}$ by $\pi_{nn}$ and takes logs:

$$\ln \frac{\pi_{ni}}{\pi_{nn}} = \underbrace{(\ln T_i + \ln \Xi_i) - (\ln T_n + \ln \Xi_n)}_{\text{Fixed Effects}} - \underbrace{\theta \ln d_{ni}}_{\text{Internal Trade Cost}} + \underbrace{\theta \ln d_{nn}}_{0} \tag{48}$$

Substituting (46) into (48), we derive (49) that is used to estimate the coefficients governing the internal trade cost[41]:

$$\ln \frac{\pi_{ni}}{\pi_{nn}} = \tilde{\gamma}_1 \times D_1 + \tilde{\gamma}_2 \times D_1 \times Dist_{ni} + \tilde{\gamma}_3 \times D_2 + \tilde{\gamma}_4 \times D_2 \times Dist_{ni} + \tilde{\gamma}_5 \times D_3 + \tilde{\gamma}_6 \times D_3 \times Dist_{ni} + \tilde{T}_i - \tilde{T}_n \tag{49}$$

[40]The specification can be extended to allow heterogeneity across provinces ($a_i$).

[41]Specifically, $\tilde{\gamma}_j = -\theta\gamma_j$, $\forall j = 1, 2, ..., 6$, $\tilde{T}_i = \ln T_i \Xi_i$. We use the extended IO table at provincial level to estimate the internal trade cost, and we assume that the internal trade cost of Tibet is the same with Qinghai Province due to data limitation. We impose the symmetry in sourcing and destination fixed effects, as implied by the model.

For the given value of $\theta$ and using the domestic trade flow at provincial level, we are able to recover $\{\gamma\}$ as well as the internal trade cost $\{\hat{d}_{ni}\}_{n,i\in\mathbf{P}}$, while leaving $\tilde{T}_i$ unidentified due to perfect multicollinearity[42] . To also capitalize the information on technology, we omit the dummy variable of Xinjiang, and the coefficient of each of the other province dummy variables stands for the its relative technology to Xinjiang. Conditional on the international parameters ($\{\kappa_{j(i)}\}$, $\lambda_{M,W}$, $T_w$), we search over $\{T_i,\ \lambda_{M,i}\}_{i\in\mathbf{P}}$ to minimize the difference between model generated expected value of indirect utility and the relative regional efficiency, and their counterpart in the data:

$$\min_{\{T_i,\lambda_{M,i}\}_{i\in\mathbf{P}}} \sum_{l\in\{B,W\}} \sum_{i\in\mathbf{P}} \left(\nu_i^l - \text{the data counterpart}\right)^2 + \sum_{i\in\mathbf{P}/\{Xinjiang\}} \left(\ln \frac{T_i \Xi_i}{T_{XJ}\Xi_{XJ}} - \text{the data counterpart}\right)^2 \tag{50}$$

#### 6.2.3 Calibrating International Parameters

For each set of international parameters ($\{\kappa_{j(i)}\}$, $\lambda_{M,W}$, $T_w$), we are able to calibrate an optimal intra-national parameters ($\{T_i,\ \lambda_{M,i}\}_{i\in\mathbf{P}}$) by solving (50). Next, bringing these parameters into the model, we calculate the trade openness for each province in 2007, which is defined as (Export + Import)/GDP. We search for the international parameters ($\{\kappa_{j(i)}\}$, $\lambda_{M,W}$, $T_w$) so that the model predicted trade openness is closest to the data counterpart:

$$\min_{\{\kappa_{j(i)}\},\lambda_{M,W},T_w} \sum_{i\in\mathbf{P}} [\text{Openness(Model)} - \text{Openness(Data)}]^2$$

#### 6.2.4 Calibrating HSR Effect

Given the calibrated parameters of model as explained above, the last step is to estimate the the parameter reflecting the marginal improvement of outsourcing ability from linking additional one city into the HSR network ($b$). To do so, we follow Dekle, Eaton, and Kortum (2007) and express the change of export share (to the ROW) relative to the initial equilibrium value for each province[43] as (between 2007 to 2015):

$$\hat{\pi}_{Wi} = \frac{\widehat{T_i\Xi_i}\hat{d}_{Wi}^{-\theta}}{\sum_j \pi_{wj}\widehat{T_j\Xi_j}\hat{d}_{Wj}^{-\theta}} \tag{51}$$

where $\pi_{Wi}$ denotes the initial equilibrium import share of ROW from province $i$ (we year 2007 as the initial time), and $\hat{\pi}_{Wi}$ denotes the share change that is computed from the data. During the period 2007 to 2015 when there were massive construction of high-speed rails, we further assume the international trade cost ($\{d_{Wj}\}$) barely changed, and we solve for $\widehat{T_i\Xi_i}$ from the system of equations $\hat{\pi}_{Wi} = \widehat{\Xi}_i / \sum_j \pi_{Wj}\widehat{\Xi}_j$, which provides the moments to calibrate outsourcing ability parameter $b$. We

[42]The exporter and importer fixed effects are symmetric as implied by the model. After adjusting for this, the new dummies sums to zero, which is in perfect multicollinearity.

[43]Let $x'$ be the current value and $x$ be the initial value. The change of $x$ is expressed as $\hat{x} = x'/x$.

estimate $b$ with non-linear least square by minimize the provincial export share change between model and the counterpart in data[44]:

$$\min_{b} \sum_{i \in \mathbf{P}} \left[\hat{\pi}_{Wi}(\text{Model}) - \hat{\pi}_{Wi}(\text{Data})\right]^2$$

So far, we have managed to estimate the migration cost $\{e^{l}_{od}\}_{l\in\{B,W\}}$, taste dispersion parameter$\{\epsilon_l\}_{l\in\{B,W\}}$, internal and international trade cost$\{d_{ij}\}$, regional technology efficiency $\{T_i\}$, as well as the parameters governing regional outsourcing ability for manufacturing tasks $\{\lambda_{M,i}\}$. With these ingredients on hand, we are able to evaluate the welfare gains from the construction of HSR, as well as carry out some counterfactual experiments. In the next section, we present our estimation results following the procedures as explained above.

# 7 Estimating Result

## 7.1 Migration and Trade Cost

Table 2 presents the point estimates for the migration cost of skilled and unskilled workers and internal trade cost. All the specifications fit the migration flow and internal trade flow well, as indicated by the $R^2$. As expected, both migration cost and internal trade costs increase with distance. The magnitude of point estimates indicates the substantial migration cost in China for both types of workers. The migration costs to nearby provinces sharing a common border within the same region are 360 log points for the unskilled labor and 380 for the skilled labor. As migration takes place with longer distance, it incurs larger migration cost: the additional costs of moving to a province within the same region that does not share a common border are 57 and 36 log points for unskilled and skilled labor, and moving to the different region costs an additional 51 log points for unskilled and 18 log points for the skilled. The point estimates of the structural changes on migration costs also indicate the unskilled labor are likely to move to the nearby area while the skilled have a comparative advantage in moving to distant area. The continuous geographic distances have nonlinear effects on the migration costs. For instance, when the origin and destination are in different regions, the marginal cost of moving an extra 1,000 kilometers is 43 log points, which is smaller than that where the origin and destination are in the same regions. Such pattern holds for both types of labors, but the coefficients are larger for the unskilled than the skilled workers.

The estimates for domestic trade costs are reported in the third column of Table 2. The specification fits data well with $R^2$ of 0.97. According to the table, transporting goods to nearby provinces that share a common border incurs trade costs by about 90 log points. As products transportation takes place with longer distance, it incurs larger trade cost: the additional trade costs to a province within the same region that does not share a common border are 17 log points, and shipping to the different region costs an additional 6 log points. Geographic distances unambiguously increases trade cost:

[44] In the appendix, we explains how we calculate provincial export share $\{\pi_{Wi}\}$.

for trading partners from provinces within the same region and sharing a common border, each additional 1,000 kilometers raises the trade cost by 36 log points; for trading partners from two provinces within the same region without a common border, the marginal trade cost of each 1,000 kilometers is 11 log points; for trading partners from two different regions, the impacts are the smallest with magnitude of 8 log points per 1,000 kilometers. Overall, the point estimates imply that the internal trade costs are substantial and increases with both institutional and geographic distance[45]. The estimates of migration cost and internal trade cost are similar to the ones obtained by Fan (2015).

Table 2: Estimation Result of Internal Trade and Migration Costs

| | Migration Cost | | Trade Cost |
|---|---|---|---|
| | *Unskilled* | *Skilled* | |
| $I_1$(Common Border, Same Region) | 3.598*** | 3.798*** | 0.906*** |
| | (0.173) | (0.180) | (0.062) |
| $I_2$(No Common Border, Same Region) | 4.172*** | 4.158*** | 1.079*** |
| | (0.234) | (0.247) | (0.081) |
| $I_3$(Different Regions) | 4.682*** | 4.344*** | 1.137*** |
| | (0.065) | (0.068) | (0.021) |
| $Dist \times I_1$(Common Boarder, Same Region) | 0.924*** | 0.709*** | 0.361*** |
| | (0.234) | (0.251) | (0.092) |
| $Dist \times I_2$(No Common Boarder, Same Region) | 0.577*** | 0.403** | 0.116* |
| | (0.181) | (0.201) | (0.069) |
| $Dist \times I_3$(Different Regions) | 0.433*** | 0.440*** | 0.083*** |
| | (0.031) | (0.033) | (0.011) |
| Observations | 860 | 793 | 841 |
| R-squared | 0.981 | 0.978 | 0.965 |

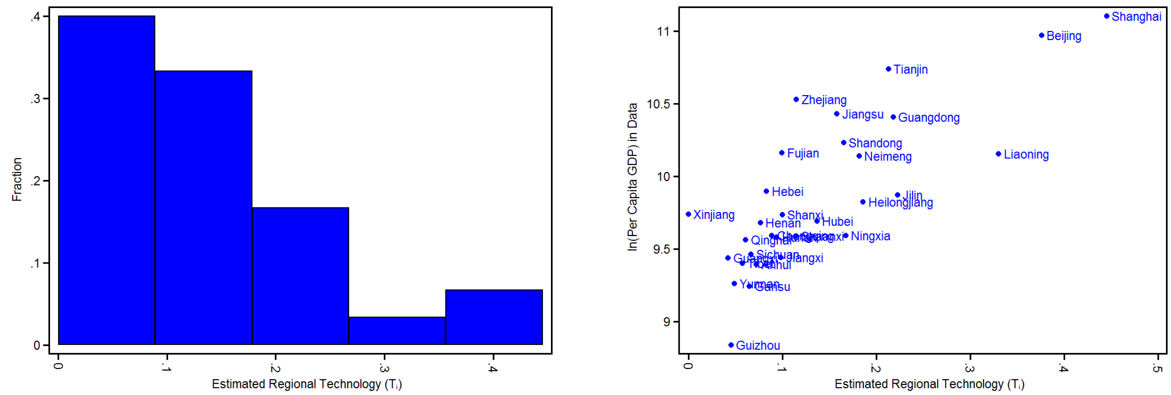
Notes: This table reports the estimates of domestic migration and trade costs. Distance is measured as the centroid distance between provinces (in 1000 km); Robust standard errors are clustered at the group level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

## 7.2 Welfare and Productivity in the Initial Equilibrium

Figure 5 displays the estimates for technology parameters. Panel (a) shows the distribution of the estimated technology parameter $T_c$ at the province level. According to our estimation, the top five provinces with high technology parameters are Shanghai, Beijing, Liaoning, Jilin, and Guangdong, and the bottom five provinces are Tibet, Yunnan, Guizhou, Guangxi, and Xinjiang. We further compare our estimated technology parameters to the provincial per capita GDP in 2007, and the latter has not been used in the calibration. We display the relationship between them in Panel (b). The graphs

[45]Institutional distance is reflected by the dummy variables in the regression

clearly exhibit a positive correlation between $T_i$ and average income, which indicates our estimation method performs reasonably, and our estimates are reliable from this point of view.




(a) Distribution of $T_i$ (b) Estimated $T_i$ and Per Capita GDP in Data

Figure 5: Estimated Technology Parameters $T_i$ (2007)

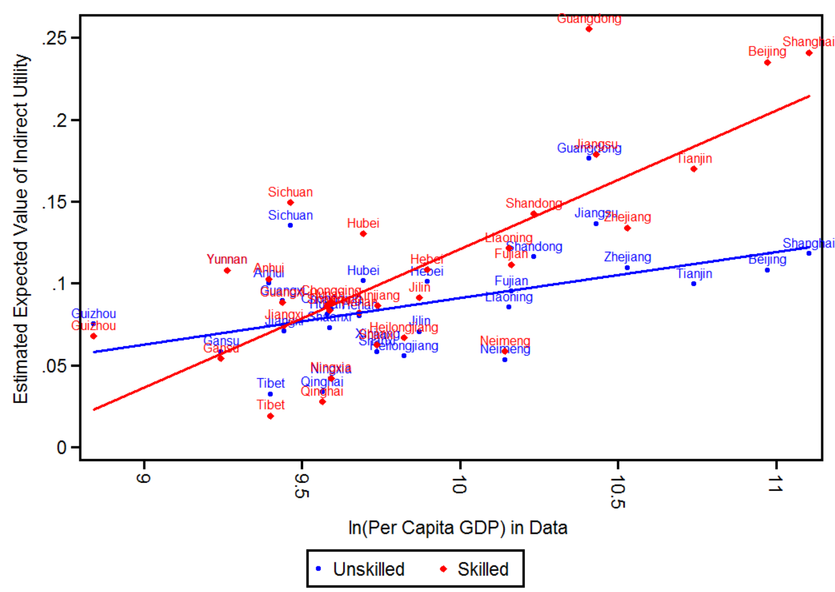



Figure 6: Estimated Expected Indirect Utility by Skill Types $v_d^l$ (2005 Census)

We display the estimated expected indirect utility of each location for labors of both types, defined as equation (30), in Figure 6. The plot shows the relationship between the indirect utility and 2007 provincial per capita GDP in logs. For both skilled and unskilled labors, living in locations with a higher level of development generates higher utility level. According to our estimation, the top five most popular provinces are Guangdong, Shanghai, Beijing, Jiangsu and Tianjin for the skilled labors. The unskilled labor benefits the most if living in Guangdong, Jiangsu, Sichuan Shanghai, and Shandong provinces. Notably, the skilled labors are more sensitive to the regional development, and their elasticity of indirect utility with respect to per capita GDP is 0.72 which is roughly twice as much as that for the unskilled labors (0.33).

Following the procedures as explained in Section 6, the full list of other calibrated parameters is reported in Table 3. Before the placement of HSR, the average matching efficiency within China is 0.11. The average semi-elasticity of matching efficiency with respect to the number of HSR hubs for the none-municipality province is 0.07. Among municipalities, Beijing gains the most along with Chongqing and Tianjin. Notably, we source three parameters from the external literature, which is the shape parameter of the Fréchet distribution for firm productivity $\theta$, the shape parameter of the distribution for worker productivity $\phi$, as well as inter-region correlation of taste draws $\rho$[46].

Table 3: Calibrated Parameters for Counterfactual Analysis

| Parameter | Value | Data Target/Source |
|---|---|---|
| Production Parameters | | |
| $\beta_M$ | 0.60 | China Statistical Yearbook (2007) |
| $\beta_S$ | 0.40 | China Statistical Yearbook (2007) |
| | | |
| International Parameters | | |
| $T_{ROW}$ | 0.001 | Provincial openness |
| $\lambda_{M,ROW}$ | 0.122 | Provincial openness |
| $\kappa$ | 0.167 | Provincial openness |
| | | |
| Other Parameters | | |
| $\theta$ | 4 | Simonovska and Waugh (2011) |
| $\rho$ | 0 | Tombe and Zhu (2015) |
| $\epsilon_{unskill}$ | 1.585 | Migration flow |
| $\epsilon_{skill}$ | 1.366 | Migration flow |
| $\phi$ | 3 | Eaton et al (2016) |
| $\gamma$ | $(\theta-\phi)/\theta$ | Determined by $\theta$ and $\phi$ |
| | | |
| Matching Efficiency Parameters | | |
| $\lambda_{M,i}, i \in \mathbf{P}$ | 0.11 | Indirect utility and technology difference |
| $b_{Other}$ | 0.07 | Change of export share $\pi_{Wi}$ (2007-2015) |
| $b_{Beijing}$ | 2.92 | Change of export share $\pi_{Wi}$ (2007-2015) |
| $b_{Shanghai}$ | 0.31 | Change of export share $\pi_{Wi}$ (2007-2015) |
| $b_{Tianjin}$ | 1.81 | Change of export share $\pi_{Wi}$ (2007-2015) |
| $b_{Chongqing}$ | 2.84 | Change of export share $\pi_{Wi}$ (2007-2015) |

Notes: $\lambda_{M,i} \equiv exp(a)$ is the matching parameter in the initial period when $N_i^{HSR} = 0, \forall i \in \mathbf{P}$, which is assumed to be same across provinces in China.

[46] In practice, one can calibrate $\phi$ using the aggregate production function as shown in Proposition 2, which is skipped due to data limitation. In the sensitivity test, we study how these outsourced parameters could affect the variables of interest.

# 8 Counterfactual Analysis

## 8.1 Aggregate Impact

In this section, we use the model as the laboratory to conduct a sequence of policy experiments to study the impact of HSR connection on the overall welfare and inequality, as well as its interaction with the domestic frictions. To begin with, we keep all model parameters at the calibrated values and close the improvement in the matching efficiency due to HSR connection, by decreasing the $b$ in (47) to zero. We compute the welfare expressed in (37) as China moves from the equilibrium without HSR to the equilibrium with HSR, which is reported in Table 4.

Table 4: Aggregate Impact of HSR Placement

| | NO HSR | HSR | Percentage Change |
|---|---|---|---|
| Unskilled Labor | 0.6434 | 0.6409 | -0.39% |
| Skilled Labor | 10.5913 | 10.7342 | 1.35% |
| Overall | 1.1907 | 1.1962 | 0.46% |

Notes: The number listed in the table stands for the computed welfare level in the given scenario, accroding to equation (37).

As shown in the table, HSR connection increases China's overall welfare by 0.46%, but the welfare gains do not accrue to everyone in the economy equally. The welfare of skilled labor rises by 1.35% while that of unskilled decreases by 0.39%. The drops in welfare level for the unskilled workers attributes to the lower costs of intermediate task brought from HSR connection, as explained in Proposition 2. An improvement in matching efficiency $\lambda_{k,i}$ decreases the average costs of using intermediate tasks, which can be regarded as the composite input biased technology shocks that reduce the demand for unskilled workers that decreases the welfare of unskilled labors.

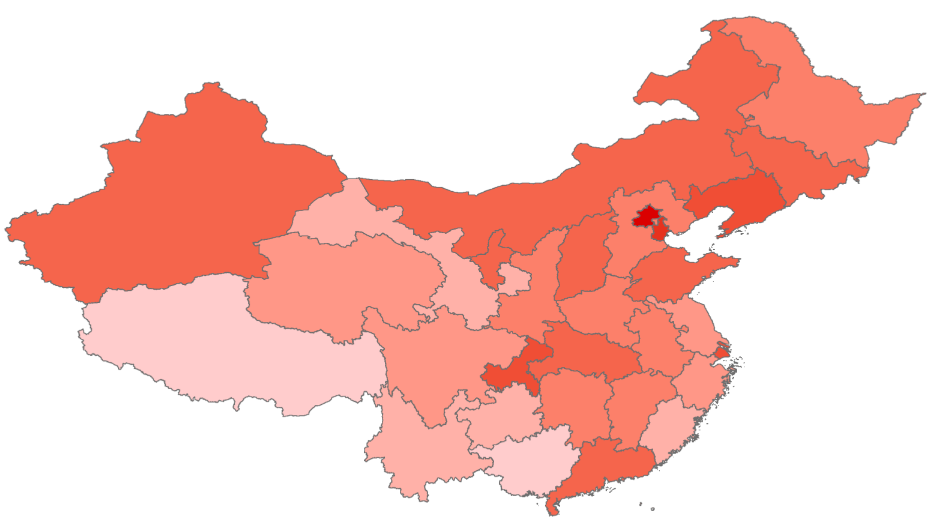


Figure 7: Regional Gains From HSR

Figure 7 displays the geographic distribution of welfare gains from HSR across provinces (the details refer to Table A.5 in the appendix.). Hereafter, we use the cold color (blue) to indicate negative changes in our variables of interest, and the hot color (red) to represent positive changes. As shown in the figure, the gains from HSR are unevenly distributed across regions. Based on the estimation, the municipalities gain the most from being linked to the HSR network. Notably, Beijing gains the most followed by Tianjin, Chongqing, and Shanghai. Welfare benefits of the inner land provinces are quite limited, which could be potentially reasoned by their relative abundance in the unskilled labors and that they are comparatively "good at" producing manufacturing tasks. An improvement in the matching with manufacturing tasks suppliers unproportionally motives the firms from regions with high labor costs to outsourcing their tasks[47], and this brings larger gains for them as well.

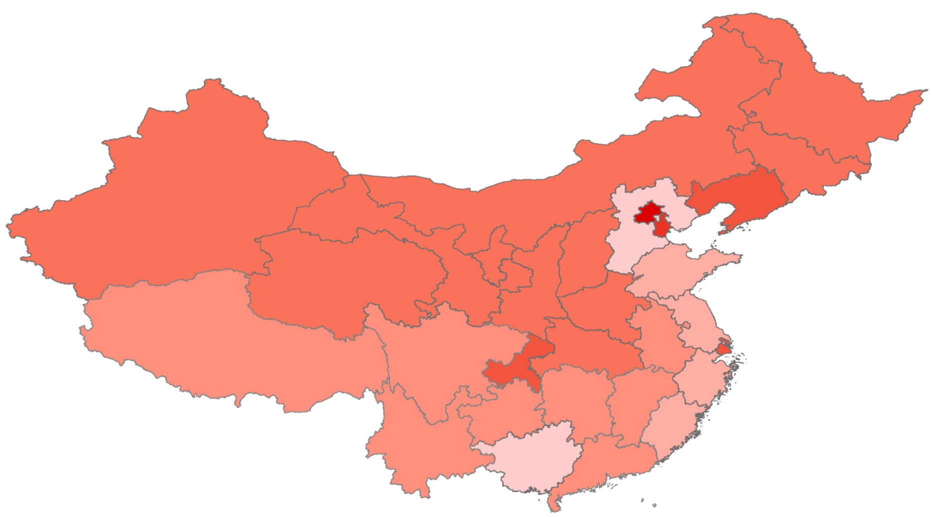


Figure 8: Technology Gains ($\widehat{T_i \Xi_i}$) From HSR

Figure 8 displays the equivalent productivity gains from being linking to the HSR (the details refer to Table A.6 in the appendix.). Similar to the welfare gains, the municipalities gain the most from being linked to the HSR network, i.e., Beijing gains the most followed by Tianjin, Chongqing, and Shanghai. Different from the welfare gains pattern, the regional technology benefits are also substantial for the inner lands. This is because firms of these regions also raise their outsourcing incentives (though not much), and they have better access to these cheap labors since they are located close to the nearby regions with low unskilled labor costs. Consequently, China's export as the share of rest of the world's total consumption increases by 0.56% due to the productivity gains by HSR connection. Table 5 reports the model performance in explaining changes of export by regimes, at

[47]Though firms from inner land provinces also have such incentive, which, however, will be much smaller than the firms from regions with high labor cost (e.g. municipalities), as their costs of performing manufacturing tasks with local unskilled labors have been already low.

the provincial level, between 2007 and 2015, according to which, HSR projects explains around 50% of export expansion in that period. It explains much less on the changes of the export in processing regime, as expected. These evidence suggests that the investment in infrastructure such as high-speed railways could effectively help to release the suppressing trend in export resulted from the rising labor costs in China.

Table 5: Export Expansion due to HSR Construction

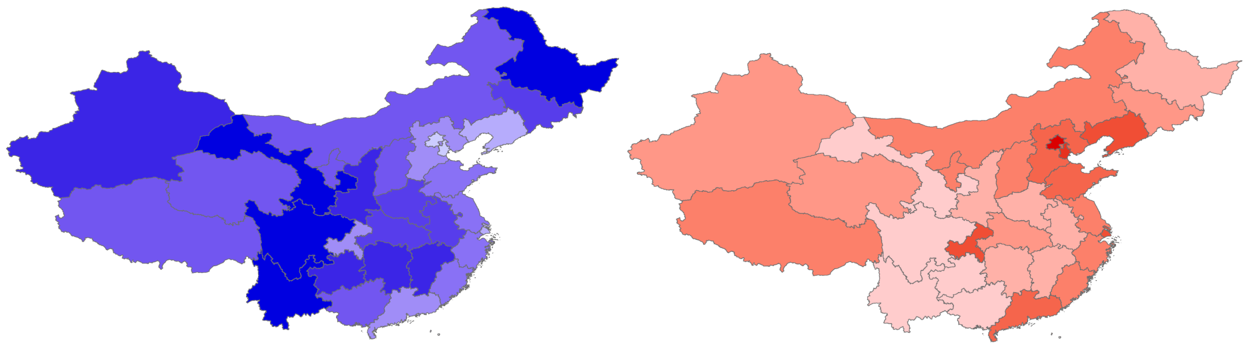
| Dept. var $\Delta Export$ (2007 - 2015) | Ordinary Export | Processing Export |
|---|---|---|
| Model Fit | 52% | 14% |

Notes: The table reports the $R^2$ as the indicator of model fit in explaining the expansion of export at the provincial level (including the municipalities) between 2007 and 2015.

As the welfare impacts are quite different between the skilled and unskilled labors indicated in Table 4, we explore in detail the welfare gains by skill types across geographic locations, as presented in Figure 9. Panel (a) shows the welfare changes across regions for the unskilled labors and Panel (b) for the skilled labors. As expected, the welfare level of unskilled labors universally declines, and that of the skilled labors rises. The overall patterns are explained by the HSR-skill complementarity in production, while the heterogeneity across regions attributes to the specialization induced by HSR. Connection to HSR raises the relative demand for skilled labors due to outsourcing, which is most evident in the coastal regions (also include the municipalities) where the average labor cost is high. This, in turn, motivates the skilled labor to move to the coastal areas, and drives the unskilled labors to the inner land area where the decline in their relative demand (due to outsourcing) is less severe compared to coastal regions[48]. With the inflow of skilled labor to coastal regions and inflow of the unskilled to the inner land, coastal regions enhance the incentive to outsourcing manufacturing tasks[49] and require more skilled labors, while the inner land provinces undergo a further drop in the relative return to the unskilled. Consequently, the skilled labors gain the most in the coastal, while unskilled have the most welfare loss in the inner land.

[48]The rising demand (from the coastal regions) for intermediate manufacturing tasks increases the relative demand for the unskilled labors in inner land.

[49]The inflow of skilled labor to the coastal area makes a firm to cost less on hiring skilled labor and will increase their production and require more skilled labor input. In the meanwhile, the inflow of the unskilled labors to the inner land increases the relative supply of unskilled labor there, making their return decline further.



(a) Unskilled Worker (b) Skilled Worker

Figure 9: Welfare Changes by Skill Types

The welfare implications for the skilled and unskilled labors found in the previous part are also consistent with the labor market outcomes of HSR. Figure 10 plots the change of employment share of skilled labor(with/without HSR placement) against a region's initial endowment of the skilled population. As clearly shown in the figure, regions that are relatively abundant in skilled labors are to attract more skilled labors if they are connected to the high-speed railways. This result is consistent with the specialization pattern as found in Lin (2017) that a better intercity passenger transportation promotes the cross-city knowledge exchanges and ultimately shifts cities connected to HSR towards skilled and communication intensive productions.

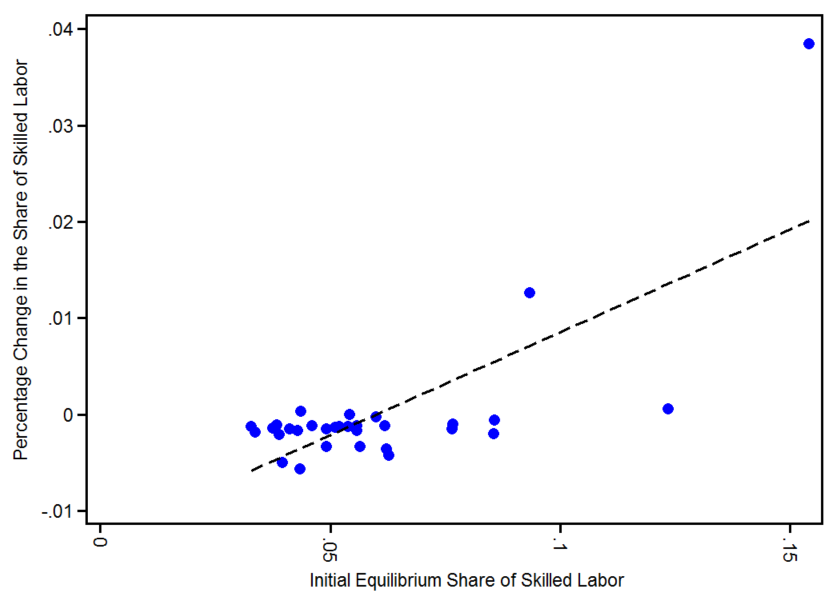



Figure 10: Regional Specialization Due to HSR

The specialization pattern is attributed to the rising return to skills as displayed in Figure 11. Connection to HSR raises the relative demand for skilled labors due to outsourcing, which is most evident in the coastal regions where the average labor cost is high. Accordingly, there is a negative gradient in changes in skill premium with respect to regions' distance to the coast on average. The finding is in line with the reduced-form evidence in Han, Liu, and Zhang (2012), and with the structural studies by Fan (2015) that attributes it to the different degree of international trade liberalization. Different from theirs, our paper provides a brand new channel emphasizing the role of regional outsourcing in explaining this active interaction between the return to skills and the geographic dimension.

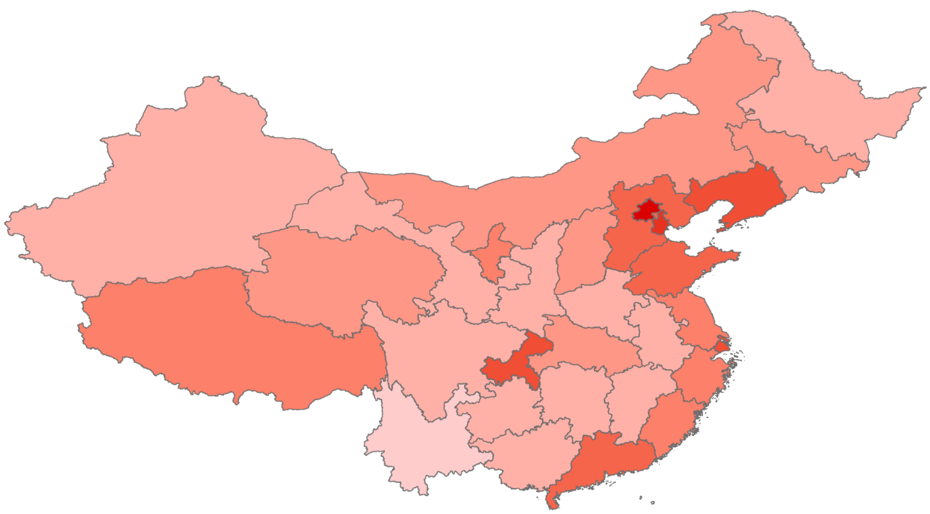


Figure 11: Changes of Skill Premium Across Regions

## 8.2 Distributional Impact

The substantially heterogeneous welfare impacts on workers of different types suggest that HSR connection might have a significant influence on inequality, between workers with different skills, and among workers of the same type but from different regions. To study the distributional effects, we compute the Theil index to measure the overall inequality in real wages in China. We further decompose the Theil index into between-region and within-region component and into between-skill and within-skill component to examine the impacts of HSR placement on each component.

Panel A of Table 6 presents the Theil inequality and its decomposition based on geographic dimension. The first row is the equilibrium without HSR placement and the second row is for the benchmark economy. In both cases, the within-region component constitutes more than 90% of the

Table 6: Distributional Impacts of High-Speed Railways

| | | | |
|---|---|---|---|
| A. Inequality by Regions | | | |
| | Within Region | Between Region | Theil Index |
| NO HSR | 0.517 | 0.046 | 0.563 |
| HSR | 0.524 | 0.047 | 0.571 |
| Increase (%) | 1.29% | 2.39% | 1.38% |
| Relative Contribution (%) | 85.90% | 14.10% | 100.00% |
| B. Inequality by Worker Types | | | |
| | Within Group | Between Group | Theil Index |
| NO HSR | 0.039 | 0.525 | 0.563 |
| HSR | 0.036 | 0.535 | 0.571 |
| Increase (%) | -6.49% | 1.96% | 1.38% |
| Relative Contribution (%) | -32.05% | 132.05% | 100.00% |

Notes: Panel A reports the decomposition of Theil index, into within- and between-region components in economy with and without HSR. The last row reports the relative contributions of the two components to the increase in the aggregate inequality after being linked to the HSR. Panel B repeats the decomposition over different working types.

overall inequality in China, while the between-region account for a very limited proportion. As reported in the third row of Panel A, when moving from the economy without HSR to the one with HSR, the overall inequality rises by 1.38%, and both between-region and within-region components inequalities rise. Though the between-region component contributes only about 8% of the inequality, its relative contribution to the increases is 14%, and the within-region contributes to the remaining 86% of the rise in the total inequality. Panel B presents the decomposition results based on skill dimension. Consistent with the message delivered by Panel A, the average difference between skilled and unskilled labors accounts for the majority of the overall inequality. HSR connection widens this gap while decreasing the disparities among workers of the same type. We learn from these numbers that the difference between the skilled and unskilled labor constitutes the primary proportion of the overall inequality in this economy; HSR connection widens such disparity substantially, due to the substitutability of unskilled laborers with intermediate tasks; it is also accompanied by the as well as the declining inequality among labors of the same type as well as the rising disparities across regions.

### 8.3 Welfare and Inequality: the Role of Labor Mobility

So far, we have learned that HSR placement brings the positive welfare benefits, at the cost of rising overall inequality which is undesired by the policy makers. In this section, we explore how the potential reforms on Hukou system aiming to reduce migration costs could help release the side effect[50]. To do so, we keep all model parameters at the calibrated values in the equilibrium with HSR and additionally decrease the migration costs for both types of workers by some proportion as listed

[50]It is motivated by the recent calls for the reform of Hukou system, which aims to heal the urban disease by reducing the migration costs embedded in institutional constraints (see Lu and Xia (2017) for details). On the other hand, as many countries along China are constructing (or constructing more) high-speed railways in the short future, understanding the role played by labor mobility bears policy significance, which is helpful to the government to make better policy to release the long-run side effect induced by HSR.

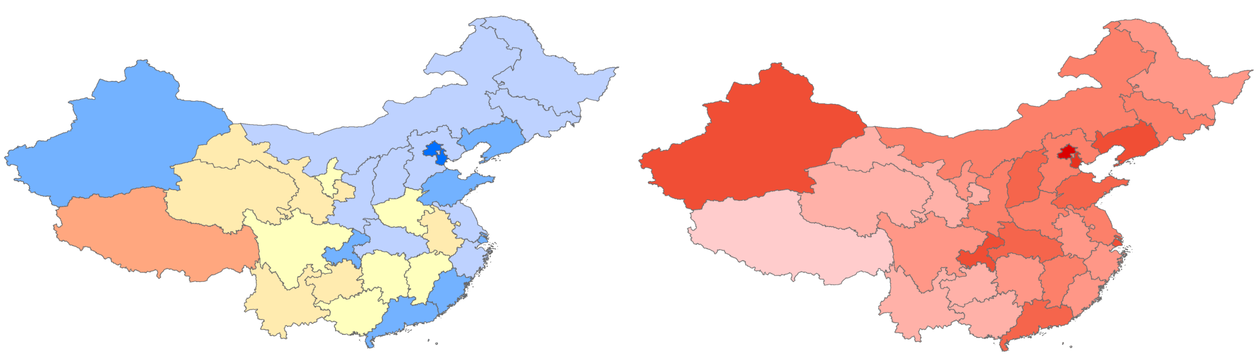
in the first row of Table 7 (e.g. 0 means no reduction in the migration costs). We calculate the welfare changes relatively to the HSR economy without any reduction in migration costs, as well as the corresponding overall inequality, which is reported in the second and third row, respectively. As clearly displayed in the table, less restriction on the labor mobility generates the substantial additional welfare gains as expected, and it also helps reduce the overall inequality.

Table 7: Impact of Additional Migration Costs Reduction on Welfare and Inequality

| Reduction in Migration Cost (%) | 0 | 20 | 40 | 60 |
|---|---|---|---|---|
| Additional Welfare Gains (%) | 0 | 0.59 | 1.69 | 4.24 |
| Inequality (Theil Index) | 0.5711 | 0.5698 | 0.5677 | 0.5640 |

Notes: The welfare gains is calculated as the percentage change of welfare relative to that in the HSR economy without migration cost reduction.

To see how welfare would be improved for labors of different types, we simulate the welfare associated with HSR placement and with the additional migration cost reduction[51], and compute the relative welfare changes compared to the initial equilibrium (no HSR nor migration cost reduction), which is presented in Figure 13. First, the unskilled labors have positive gains in some regions of the inner land (recall that they universally suffer the welfare loss as in Figure 9). In the meanwhile, the skilled labors still have positive gains but with smaller dispersion across regions. The numerical exercise suggests that the government should consider the future reform to reduce internal migration cost as an effective method to alleviate the rising inequality induced HSR. Besides that, the potential gains of future reform on Hukou system remains substantial.



(a) Unskilled Worker

(b) Skilled Worker

Figure 12: Welfare Changes by Skill Types with Reduction in Migration Cost of Unskilled Worker

[51]We set the reduction as 5%.

Above analysis considers free labor mobility as an additional adjustment margins to alleviate the side effect brought from the HSR connection[52]. We now turn to the examination of the interaction between the labor mobility and HSR placement. To do so, we simulate the welfare and inequality changes when moving from a none-HSR economy to the one with HSR, while keeping internal migration cost constant between the two scenarios. Then we gradually decrease the migration costs for both types of workers and plot the corresponding relative welfare and inequality changes in Figure 13. We use red color to denote the welfare changes and blue color for the changes on inequality. Welfare gains from being connected to HSR network are found to be smaller in the economy with lower migration costs. The intuition is as follow, when labor could freely mobile across regions, there would be no substantial difference in the labor costs, which implies firm's incentive to outsourcing will be small. As a result, the placement of HSR would not provide many benefits. In turn, it indicates that HSR would be especially beneficial in the long run for countries (like China) which has a large dispersion in regional development or have a high internal migration cost.

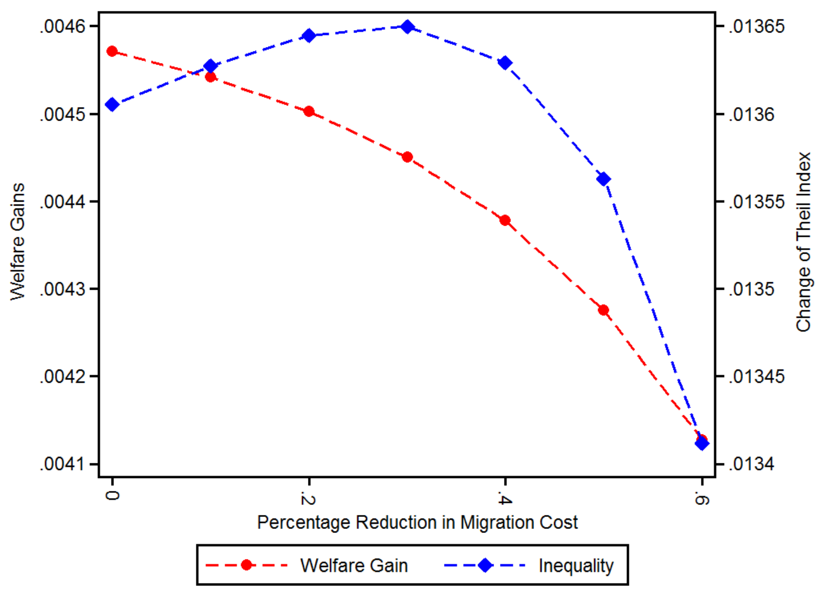
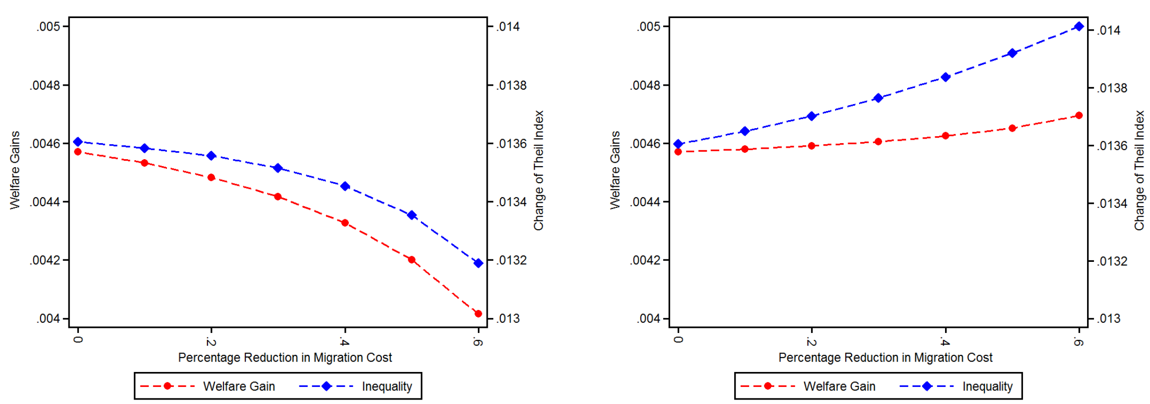



Figure 13: Impact of Reduction in Migration Cost on Welfare Gains and Inequality

Surprisingly, when we reduce the migration costs (for both types of workers), the overall inequality initially rises and then declines, exhibiting a non-monotone pattern. To understand the inverted-U relationship between labor mobility and inequality, we carry out the similar quantitative exercise but with the reduction in migration cost of only one particular type of workers in each experiment. The

[52]The previous quantitative exercise studies how the economy would evolve in the long run for the future policy change, while hereafter we analyze the implications of China's past labor market integration on welfare, which is backward looking.

results are presented in Figure 14. Panel (a) displays the relative changes of welfare and inequality when we only reduce the migration costs of the unskilled labor, and Panel (b) shows the results when we only reduce the migration costs of skilled labor. When the unskilled labors become more mobile, the incentive of outsourcing would be smaller, since a firm could easily hire a local worker to perform the manufacturing tasks and it would incur the similar production costs compared to outsourcing to other regions. In the absence of outsourcing motivation, the shrink in the relative demand for unskilled workers would be quite limited, so is the benefits of being connecting to HSR. This helps explain the pattern as found in Panel (a). In contrast, when only reducing the migration cost of the skilled labors, there remains a substantial difference in the labor cost of using unskilled labor, in which case firms still have the incentive to outsourcing manufacturing tasks. Similar to the benchmark case, the relative demand for unskilled (skilled) labor declines (rises) after being connected to HSR. In fact, as skilled labors become more mobile, they become more willing to move to regions where their relative return is high. The rising skill supply decreases the absolute cost of the skilled labor, which in turn leads firms to expand production and enhance the incentive to outsourcing. The complementary relationship between skills and HSR raises the overall welfare mainly through the increased relative demand and increased relative return to skills, which is consistent with the rising overall inequality as found in Panel (b). The inverted-U relationship in Figure 13 is then jointly determined by these two opposite forces, which sources from reductions in the migration costs of the unskilled and skilled labors, respectively[53].




(a) Reduction in *Unskilled* Migration Cost (b) Reduction in *Skilled* Migration Cost

Figure 14: Impact of Reduction in Migration Cost on Welfare Gains and Inequality By Skill Type

The above analysis suggests that reforms in the internal labor markets would generate different outcomes when the policy instruments target on the different groups of workers.[54]. Specifically, we

[53]The slope of welfare change with respect to migration cost reduction also becomes flattered in Figure13 than the one when we only reduce the migration costs of unskilled labor.

[54]The heterogeneous outcomes differ our paper from the previous study such as Fan (2015), who found the interaction between domestic labor mobility and welfare gains from international trade to be universally negative, regardless of

expect the overall inequality decline but also accompanied by the shrink in the welfare gains from HSR placement, if policy aims to promote the mobility of the low skilled workers who are highly exposed to outsourcing. On the other hand, if the policy is primarily to boost the mobility among high-educated groups, we expect to benefit more gains from the construction of HSR, but at the cost of the rising inequality. Choosing the proper policy instruments would be important which should be based on policymakers' objective of reforms.

## 9 Model Extension: Universal Gains from HSR

A crucial implication of our benchmark model is that the HSR-induced productivity gains are biased towards the skilled labor. A reduction in intermediate inputs prices leads firms to substitute away from unskilled labors for the manufacturer task (elasticity is greater than unity), given that intermediate inputs use final output and the aggregate production technology is considered as equivalent to a roundabout production function according to Proposition 2. In this section, we extend our benchmark model such that, in representative aggregation, the production of intermediate inputs also uses unskilled labors in addition to the final outputs of other firms[55]. In the extended framework, we show that there would be positive welfare (productivity) gains that benefit both types of workers.

We assume a firm $j$ in location $i$ can produce a quantity of output $Q_i(j)$ by combining two intermediate tasks (service and manufacturing tasks, which are denoted as $S$ and $M$, respectively) according to the new production function

$$Q_i(j) \quad = \quad z_i(j) \prod_{k \in \{S,M\}} b_k^{-1} \left[ \frac{(m_{k,i}(j)/\zeta_k)^{\zeta_k} \left(L_{k,i}/(1-\zeta_k)\right)^{1-\zeta_k}}{\beta_k} \right]^{\beta_k} \tag{52}$$

where $z_i(j)$ stands for producer $j$'s Hicks-neutral productivity, $m_{k,i}(j)$ denotes the input of task $k$, $b_k$ is a constant to be specified later, and $\beta_k$ is the Cobb-Douglas share of task $k$, which satisfy that $\beta_k > 0$ and $\beta_S + \beta_M = 1$. Different from the previous production function, to transform task $k$ to output, it also requires labor appropriate for that task, and we assume labor ($L_{k,i}$) and task ($m_{k,i}(j)$) are aggregated under Cobb-Douglas technology with $1 - \zeta_k$ denotes within-task labor share. Particularly, the extension will be the same with our benchmark if $\zeta_k$ is chosen at unity (the proof is provided in Appendix).

With the rest of the model structure unchanged, the aggregate production function of region $i$ under the extended model could be expressed as:

$$Q_i = \{L_{S,i}\}^{\beta_S} \left\{ \left[ \left( \psi \left(L_{1M,i}\right)^{\frac{\phi}{\phi+1}} + (1-\psi) \left(I_{M,i}\right)^{\frac{\phi}{\phi+1}} \right)^{\frac{\phi+1}{\phi}} \right]^{\zeta_M} \left[L_{2M,i}\right]^{1-\zeta_M} \right\}^{\beta_M} \tag{53}$$

whether the object is unskilled or skilled labor.

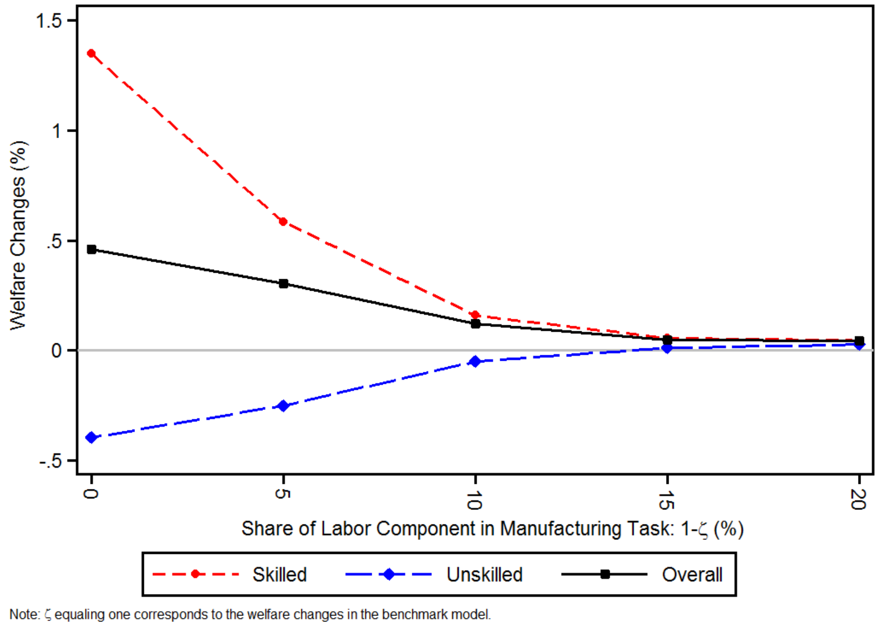
[55]One could consider the extension as that the manufacturing task is fulfilled by using some unskilled workers to assemble the intermediate inputs. Therefore, at the aggregate level, there is a positive reservation demand for unskilled labors as long as the demand for manufacturing inputs are positive, whereas there is no reservation demand for unskilled labors in our benchmark model.

where $L_{S,i}$ denotes skilled labors employed in fulfilling service task; $L_{1M,i}$ are unskilled labors employed in producing manufacturing inputs (subject to outsourcing); $L_{2M,i}$ denotes the unskilled labors employed for "assembling" in manufacturing task (not subject to outsourcing); $I_{M,i}$ are the intermediates used for producing manufacturing inputs, and $\psi \equiv 1/\left[1+\Gamma(1+1/\phi)^{\phi/(1+\phi)}\right]$. The new term $L_{2M,i}$ captures the reservation demand for the unskilled labors. Distinct from Proposition 2, production of final output would always require the positive amount of unskilled labors with a share no less than $\beta_M(1-\zeta_M)$, no matter how low the prices of manufacturing intermediate inputs are. Whereas in the benchmark model, the unskilled labors could be demanded at zero when the costs of manufacturing intermediate inputs are limiting to zero. As the representative firms will always demand unskilled workers above the reservation amount, a continuous reduction (due to HSR connection) in intermediate costs from high level would ultimately generate positive gains for both skilled ($L_{S,i}$) and unskilled ($L_{2M,i}$), regardless of the loss from the shrinking demand of the other unskilled labors ($L_{1M,i}$).The ultimate gains or losses for the overall unskilled workers would depend on $\zeta_M$, the assembling labor share within manufacturing task.

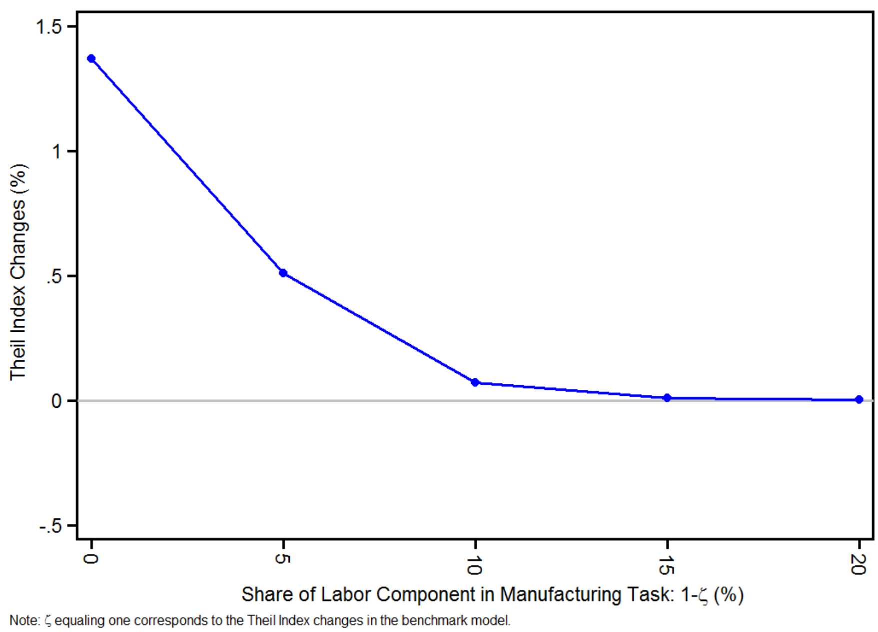
To shed light on the role of reservation demand for unskilled labors in affecting the distribution of welfare gains, we simulate the extended model keeping all calibrated parameters unchanged except for the new parameter $\zeta_M$. Figure 15 displays the HSR induced welfare and inequality changes with respect to the share of reservation demand for the unskilled in production[56]. The top panel shows the results on welfare. When reservation labor share for the unskilled increases, the welfare of the unskilled improves. As clearly demonstrated in the graph, when $1-\zeta_M$ reaches to some modestly large number (about 13.75%), there are positive gains for both unskilled and skilled labors. In contrast, the gains for the skilled become less pronounced, as a matter of fact that the potential outsourceable bundle in production accounts for a smaller share as $1-\zeta_M$ becomes larger. Consequently, the overall gains from outsourcing (gains from HSR) shrink as $\zeta_M$ decreases. This also implies the HSR induced inequality becomes less pronounced, as denoted in the bottom panel.

In this section, we offer an extension to our benchmark model by introducing the reservation demand for particular types of labor in production. We show with simulations that these reservation demands are crucial in adjusting the distribution of welfare gains by HSR. As there is reservation demand for unskilled labors, a continuous reduction in intermediate prices would eventually generate sufficient productivity gains to compensate the loss in demand for the unskilled due to outsourcing, such that HSR would positively influence the welfare for both types of workers. However, incorporating such feature would require a more detailed data where we could distinguish the unskilled labors used for "assembling" (not subject to outsourcing) from the ones for producing manufacturing material inputs (subject to outsourcing). Constrained by the data limitation, we leave the exercise of complete calibration for the future study.

[56]A zero labor share, i.e., $\zeta_M = 1$, corresponds to the benchmark model.




(a) *Welfare Changes*




(b) *Inequality (Theil Index) Changes*

Figure 15: Welfare and Inequality Redux: Share of Labor Component in Manufacturing Task (1-$\zeta_k$)

# 10 Conclusion

This paper studies the aggregate and distributional effects of the connection to high-speed railway on an economy with internal trade costs and migration costs. We make two contributions to our understanding of the impacts of large transportation infrastructure projects, in the context of an enormous expansion of high-speed rails in China. Taking advantage of the rapid expansion as plausible exogenous shocks that improves firm-to-firm matching efficiency across regions over time, our first contribution is to identify the causal relationship between HSR connection and exporting performance in case of China. We find the connection to HSR significantly promotes a region's exports. Besides the direct impact, we also find the positive spillovers of HSR, whose effect is stronger in areas closer to HSR hubs.

Our second contribution is to shed light on the mechanisms at work by relating the HSR-driven regional outsourcing to the HSR-driven increases in welfare and inequality. To do so requires a calibrated, general equilibrium model of trade with many regions and with producer-supplier linkage. We extend the work of Eaton, Kortum, and Kramarz (2016) to construct such a unified model and structurally estimated the key parameters using the auxiliary model equations. Our quantitative exercise reveals that the construction of HSR between 2007 and 2015 increased China's overall welfare by 0.46%, but was also associated with an increase in national inequality. However, the rising inequality could be alleviated by the reform to reduce internal migration cost. Gains from HSR are larger when labor migration costs are higher, implying HSR projects are an ideal policy for countries like China which feature high internal migration barriers. On top of it, we find labor mobilities for workers of different types play the distinct (reverse) roles in affecting the HSR-induced welfare and inequality changes, which also suggests that adopting the proper policy instruments would be important.

In an extended framework, we introduce the reservation demand for particular types of labor in production. We show with simulations that reservation demands are crucial in adjusting the distribution of welfare gains by HSR. A continuous reduction in intermediate prices would eventually generate sufficient productivity gains to compensate the loss in demand for the unskilled due to outsourcing, such that HSR would positively influence the welfare for both types of workers.

Despite the abundance of studies on the goods-shipping transportation infrastructure, minimal attention has been drawn to the passenger-shipping transportation infrastructure that reduces the cost of moving people. Our paper is notably distinct from the studies on the effects of highway/ traditional railroads development and the ones on the partial effect of HSR. This paper abstracts from some important aspects of the real world that could affect the impacts of HSR connection. For instance, to allow the service tasks to be outsourceable might reduce both aggregate and distributional effects. Incorporating the important features such as service outsourcing, heterogeneity in outsourcibility across industries, to distinguish the unskilled labors for "assembling" from the ones for producing manufacturing material inputs, as well as the dynamic effects into current framework will be in the future research.

# References


AHLFELDT, G. M., S. J. REDDING, D. M. STURM, AND N. WOLF (2015): "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 83(6), 2127–2189.

ALVAREZ, F., AND R. E. LUCAS (2007): "General equilibrium analysis of the Eaton–Kortum model of international trade," *Journal of monetary Economics*, 54(6), 1726–1768.

AMITI, M., AND J. KONINGS (2007): "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," *American Economic Review*, 97(5), 1611–1638.

ANTRAS, P., T. C. FORT, AND F. TINTELNOT (2014): "The margins of global sourcing: theory and evidence from US firms," Discussion paper, National Bureau of Economic Research.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRÍGUEZ-CLARE (2012): "New trade models, same old gains?," *The American Economic Review*, 102(1), 94–130.

BANERJEE, A., E. DUFLO, AND N. QIAN (2012): "On the road: Access to transportation infrastructure and economic growth in China," Discussion paper, National Bureau of Economic Research.

BERNARD, A. B., A. MOXNES, AND Y. U. SAITO (2015): "Production networks, geography and firm performance," Discussion paper, National Bureau of Economic Research.

BØLER, E. A., A. MOXNES, AND K. H. ULLTVEIT-MOE (2012): "Technological change, trade in intermediates and the joint impact on productivity," .

BRYAN, G., AND M. MORTEN (2015): "Economic development and the spatial allocation of labor: Evidence from Indonesia," .

CAI, J., AND A. SZEIDL (2016): "Interfirm Relationships and Business Performance," Discussion paper, National Bureau of Economic Research.

CALIENDO, L., AND F. PARRO (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 82(1), 1–44.

CALIENDO, L., F. PARRO, E. ROSSI-HANSBERG, AND P.-D. SARTE (2014): "The impact of regional and sectoral productivity changes on the US economy," Discussion paper, National Bureau of Economic Research.

CHARNOZ, P., C. LELARGE, AND C. TREVIEN (2016): "Communication Costs and the Internal Organization of Multi-Plant Businesses: Evidence from the Impact of the French High-Speed Rail," Discussion paper, Institut National de la Statistique et des Etudes Economiques, DESE.

COSTINOT, A., AND A. RODRIGUEZ-CLARE (2013): "Trade theory with numbers: Quantifying the consequences of globalization," Discussion paper, National Bureau of Economic Research.

CRISTEA, A. D. (2011): "Buyer-seller relationships in international trade: Evidence from US States' exports and business-class travel," *Journal of International Economics*, 84(2), 207–220.

DEKLE, R., J. EATON, AND S. KORTUM (2007): "Unbalanced Trade," *The American Economic Review*, 97(2), 351–355.

DI GIOVANNI, J., A. A. LEVCHENKO, AND J. ZHANG (2014): "The global welfare impact of China: trade integration and technological change," *American Economic Journal: Macroeconomics*, 6(3), 153–183.

DONALDSON, D. (2016): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*.

DURANTON, G., P. M. MORROW, AND M. A. TURNER (2014): "Roads and Trade: Evidence from the US," *Review of Economic Studies*, 81(2), 681–724.

EATON, J., AND S. KORTUM (2002): "Technology, geography, and trade," *Econometrica*, 70(5), 1741–1779.

EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An anatomy of international tade: Evidence from Frech firms," *Econometrica*, 79(5), 1453–1498.

——— (2016): "Firm-to-Firm Trade: Imports, exports, and the labor market," Discussion paper, Research Institute of Economy, Trade and Industry (RIETI).

EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2015): "Competition, markups, and the gains from international trade," *The American Economic Review*, 105(10), 3183–3221.

FABER, B. (2014): "Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System," *Review of Economic Studies*, 81(3), 1046–1070.

FAN, J. (2015): "Internal geography, labor mobility, and the distributional impacts of trade," *Working Paper*.

FEENSTRA, R. C., AND G. H. HANSON (1996): "Globalization, outsourcing, and wage inequality," *The American Economic Review*, 86(2), 240.

GIROUD, X. (2013): "Proximity and investment: Evidence from plant-level data," *The Quarterly Journal of Economics*, 128(2), 861–915.

GOLDBERG, P. K., A. K. KHANDELWAL, N. PAVCNIK, AND P. TOPALOVA (2010): "Imported intermediate inputs and domestic product growth: Evidence from India," *The Quarterly Journal of Economics*, 125(4), 1727–1767.

HALPERN, L., M. KOREN, AND A. SZEIDL (2015): "Imported inputs and productivity," *The American Economic Review*, 105(12), 3660–3703.

Han, J., R. Liu, and J. Zhang (2012): "Globalization and wage inequality: Evidence from urban China," *Journal of International Economics*, 87(2), 288–297.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2013): "The allocation of talent and us economic growth," Discussion paper, National Bureau of Economic Research.

Lin, Y. (2017): "Travel costs and urban specialization patterns: Evidence from China high speed railway system," *Journal of Urban Economics*, 98, 98–123.

Lu, M., and Y. Xia (2017): "Migration in the People's Republic of China," Discussion paper, Asian Development Bank.

Ma, L., and Y. Tang (2016): "Geography, Trade, and Internal Migration in China," *Working Paper*.

Michaels, G. (2008): "The effect of trade on the demand for skill: Evidence from the interstate highway system," *The Review of Economics and Statistics*, 90(4), 683–701.

Michaels, G., S. J. Redding, and F. Rauch (2011): "Technical Note: An Eaton and Kortum (2002) Model of Urbanization and Structural Transformation," Discussion paper, Mimeo, London School of Economics and Princeton University.

MRC (2008): "Mid-to-Long Term Railway Network Plan," *Report of The Ministry of Railways of China*.

Ou, J., B. Richard, Y. Jin, and N. Zhou (2014): "High-Speed Railways in China: Trafic Analysis," *The World Bank Report: China Transport Note Series*.

Qin, Y. (2017): "No county left behind? The distributional impact of high-speed rail upgrades in China," *Journal of Economic Geography*, 17(3), 489–520.

Ramondo, N., A. Rodríguez-Clare, and M. Saborío-Rodríguez (2016): "Trade, domestic frictions, and scale effects," *The American Economic Review*, 106(10), 3159–3184.

Redding, S. J., and M. A. Turner (2014): "Transportation costs and the spatial organization of economic activity," Discussion paper, National Bureau of Economic Research.

Tombe, T., X. Zhu, et al. (2015): "Trade, migration and productivity: A quantitative analysis of china," *Manuscript, University of Toronto*.

Wright, G. C. (2014): "Revisiting the employment impact of offshoring," *European Economic Review*, 66, 63–83.

# Theoretical Appendix

## Proof of Proposition 1

The probability that firm $j$ in country $i$ is not able to encounter any suppliers with price lower than $c$ for task $k$ is $exp(-\rho_{k,i}(c))$ by Poisson distribution. The cost distribution of outsourcing the task, given parameter restriction $\gamma = (\theta - \phi)/\theta$, is expressed as

$$G^o_{k,i}(c) \quad = \quad 1 - exp(-v_{k,i}c^{\phi})$$

where $v_{k,i} \equiv \frac{\theta}{\phi}\lambda_{k,i}Y_i^{\phi/\theta}$. The the probability that labor cost is no less than $c$ ($w_{k,i}/q_{k,i}(j) \geq c$) is $F(w_{k,i}/c)$, and the cost distribution of hiring labor is thus

$$G^l_{k,i}(c) = 1 - exp(-w_{k,i}^{-\phi}c^{\phi})$$

If the cost of using labor to fulfill task $k$ is $p$, then the probability that hiring labor costs less than outsourcing is equal to

$$\begin{aligned} Pr(c_o \geq p) &= 1 - G^o_{k,i}(p) \\ &= exp(-v_{k,i}p^{\phi}) \end{aligned}$$

Then we integrate this for all possible $p$ under density $dG^l_{k,i}(p)$ to obtain

$$\begin{aligned} \int_0^{+\infty} exp(-v_{k,i}p^{\phi})dG^l_{k,i}(p) &= -w_{k,i}^{-\phi}\int_0^{+\infty} \phi p^{\phi-1}exp\left[-(v_{k,i} + w_{k,i}^{-\phi})p^{\phi}\right]dp \\ &= -\frac{w_{k,i}^{-\phi}}{w_{k,i}^{-\phi} + v_{k,i}}\int_0^{+\infty} dexp\left[-(v_{k,i} + w_{k,i}^{-\phi})p^{\phi}\right] \\ &= \frac{w_{k,i}^{-\phi}}{\Xi_{k,i}} \end{aligned}$$

where $\Xi_{k,i} = v_{k,i} + w_{k,i}^{-\phi}$. The probability of fulfilling task $k$ by outsourcing is then $v_{k,i}/\Xi_{k,i}$.

## Proof of Proposition 2

The proof follows Eaton, Kortum, and Kramarz (2016). We only focus on the manufacturing tasks, since service tasks uses the skilled workers only. Given the price distribution of the for an intermediate to fulfill manufacturing task in country $i$ is $G^o_{M,i}(c) = 1 - exp(-v_{M,i}p^{\phi})$, the average unit cost of using intermediate input across firms in region $i$ is

$$\bar{p}_{M,i} = \int_0^{+\infty} p dG^o_{M,i}(p)$$
$$= \int_0^{+\infty} \phi \times v_{M,i} p^{\phi} \times exp\left(-v_{M,i}p^{\phi}\right) dp$$
$$= \int_0^{+\infty} \left(\frac{x}{v_{M,i}}\right)^{1/\phi} \times exp(-x)dx$$
$$= (v_{M,i})^{-1/\phi}\,\Gamma\left(1+\frac{1}{\phi}\right)$$

The aggregate labor share in the production of task $M$ can be rewritten as (by written $v_{M,i}$ as function of average price $\bar{p}_{M,i}$):

$$\beta_i^{l(M)} = \beta_M \times \frac{w_{M,i}^{-\phi}}{v_{M,i} + w_{M,i}^{-\phi}}$$
$$= \beta_M \times \frac{w_{M,i}^{-\phi}}{\left[\bar{p}_{M,i}/\Gamma\left(1+\frac{1}{\phi}\right)\right]^{-\phi} + w_{M,i}^{-\phi}}$$

We now turn to the production function shown in Proposition 2. For each manufacturing task in region $i$, the representative firm could hire unskilled labor $L_{M,i}$ at blue-collar wage rate $w_i^B$, and purchase the composite intermediate task $I_{M,i}$ at the average price $\bar{p}_{M,i}$. The first order condition of cost minimization of manufacturing task yields:

$$\frac{w_{M,i}L_{M,i}}{\bar{p}_{M,i}I_{M,i}} = \Gamma\left(1+\frac{1}{\phi}\right)^{-\phi} \times \left(\frac{w_{M,i}}{\bar{p}_{M,i}}\right)^{-\phi}$$
$$= \frac{w_{M,i}^{-\phi}}{\left[\bar{p}_{M,i}/\Gamma\left(1+\frac{1}{\phi}\right)\right]^{-\phi}}$$

Then it is straightforward that the labor share generated by the aggregate production function is same with the one implied by the assumptions on technology in model.

### Derivation of Cost Distribution $\mu_{ni}(c)$

$$\begin{aligned}
\mu_{ni}(c) &= E\left\{\mu_i^Z\left(\frac{d_{ni}}{c}\prod_{k\in\{S,M\}}\frac{c_{k,i}^{\beta_k}}{b_k}\right)\right\} \\
&= T_i d_{ni}^{-\theta} c^{\theta} \prod_{k\in\{S,M\}} \int_0^{+\infty} b_k^{\theta} c_k^{-\theta\beta_k} dG_{k,i}(c_k) \\
&= T_i d_{ni}^{-\theta} c^{\theta} \prod_{k\in\{S,M\}} b_k^{\theta}\left(\Xi_{k,i}\right)^{\theta\beta_k/\phi} \int_0^{+\infty} x^{-\theta\beta_k/\phi} exp(-x)dx \\
&= T_i d_{ni}^{-\theta} c^{\theta} \prod_{k\in\{S,M\}} b_k^{\theta}\left(\Xi_{k,i}\right)^{\theta\beta_k/\phi}\,\Gamma\left(1-\frac{\theta\beta_k}{\phi}\right) \\
&= T_i \Xi_i d_{ni}^{-\theta} c^{\theta}
\end{aligned}$$

where we define $x_k \equiv \Xi_{k,i} c_k^{\phi}$ in the third equality. The last equality is derived by assuming $b_k = \Gamma(1 - \frac{\theta\beta_k}{\phi})^{-\frac{1}{\theta}}$, which require $\theta\beta_k/\phi < 1$ in order to make $\mu_{ni}(c)$ meaningful. We define $\Xi_i \equiv \prod_{k\in\{S,M\}} (\Xi_{k,i})^{\theta\beta_k/\phi}$ to obtain the last equality.

### Algorithm of Computing Y

Y is denoted as the vector of $N \times 1$, whose $n$th element is defined as (20). We denote the function form of (20) as $\mathrm{Y} = f(\mathrm{Y})$. To show there is a fixed point, we transform the fixed point problem in form of $\ln \mathrm{Y} = F(\ln \mathrm{Y})$,where $F(x) \equiv \ln f(exp(x))$. There exists a unique fixed point of $F$ if it is a contraction. $F(x)$ is derived as [57],

$$\begin{aligned}
F_n(x) &= \ln \left\{ \sum_i T_i d_{ni}^{-\theta} \prod_{k\in\{S,M\}} \left[ \frac{\theta}{\phi} \lambda_{k,i} exp(\frac{\phi}{\theta} x) + w_{k,i}^{-\phi} \right]^{\theta\beta_k/\phi} \right\} \\
&= \ln \left\{ \sum_i T_i d_{ni}^{-\theta} w_{S,i}^{-\theta\beta_k} \left[ \frac{\theta}{\phi} \lambda_{M,i} exp(\frac{\phi}{\theta} x) + w_{M,i}^{-\phi} \right]^{\theta\beta_M/\phi} \right\} \\
&= \ln \left\{ \sum_i exp \left( \ln \left( T d_{ni}^{-\theta} \right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M}{\phi} \ln \left[ \frac{\theta}{\phi} \lambda_{M,i} exp(\frac{\phi}{\theta} x) + w_{M,i}^{-\phi} \right] \right) \right\}
\end{aligned}$$

where the second equality comes from the assumption that the service task cannot be outsourced, $\lambda_{S,i} = 0$.

To apply the Blackwell's contraction mapping theory, we need to verify both monotonicity and discounting properties for $F(x)$ to confirm the sufficient conditions. For monotonicity, let $x$ and $y$ satisfy that $x \leq y$, as all the exogenous parameters are positive then

$$\begin{aligned}
F_n(x) &= \ln \left\{ \sum_i exp \left( \ln \left( T d_{ni}^{-\theta} \right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M}{\phi} \ln \left[ \frac{\theta}{\phi} \lambda_{M,i} exp(\frac{\phi}{\theta} x) + w_{M,i}^{-\phi} \right] \right) \right\} \\
&\leq \ln \left\{ \sum_i exp \left( \ln \left( T d_{ni}^{-\theta} \right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M}{\phi} \ln \left[ \frac{\theta}{\phi} \lambda_{M,i} exp(\frac{\phi}{\theta} y) + w_{M,i}^{-\phi} \right] \right) \right\} \\
&= F_n(y)
\end{aligned}$$

For discounting, for $\forall a > 0$

[57] $F_n(x)$ denote the $n$th element of vector $F(x)$.

$$
\begin{aligned}
F_n(x+a) &= \ln\left\{\sum_i exp\left(\ln\left(Td_{ni}^{-\theta}\right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i}exp(\frac{\phi}{\theta}a)\times exp(\frac{\phi}{\theta}x) + w_{M,i}^{-\phi}\right]\right)\right\} \\
&= \ln\left\{\sum_i exp\left(\ln\left(Td_{ni}^{-\theta}\right) - \theta\beta_S \ln w_{S,i} + \beta_M a + \frac{\theta\beta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i}exp(\frac{\phi}{\theta}x) + exp(-\frac{\phi}{\theta}a)w_{M,i}^{-\phi}\right]\right)\right\} \\
&= \beta_M a + \ln\left\{\sum_i exp\left(\ln\left(Td_{ni}^{-\theta}\right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i}exp(\frac{\phi}{\theta}x) + exp(-\frac{\phi}{\theta}a)w_{M,i}^{-\phi}\right]\right)\right\} \\
&\leq \beta_M a + \ln\left\{\sum_i exp\left(\ln\left(Td_{ni}^{-\theta}\right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i}exp(\frac{\phi}{\theta}x) + w_{M,i}^{-\phi}\right]\right)\right\} \\
&= \beta_M a + F_n(x)
\end{aligned}
$$

where the inequality comes from that $exp(-\frac{\phi}{\theta}a) < 1$. As $\beta_M < 1$, discounting is verified.

Therefore, we are able to compute the fixed point ($\ln \Gamma$) by iterating on:

$$\ln \mathrm{Y}_{t+1} = F(\ln \mathrm{Y}_t)$$

with the initial value of $\ln \mathrm{Y}_0 = 0$. After we obtain $\ln \mathrm{Y}$, we are able to calculate $\mathrm{Y}$. The result directly implies that each element of $\ln \mathrm{Y}$ increases with technology parameter $T_i$ anywhere, decreases with any task-specific labor cost $w_{k,i}$ in any region, and increases with the task-specific arrival rate of price quotes $\lambda_{k,i}$ in any region, though the comparative statics only capture the partial effect and do not predict the general equilibrium outcomes.

## Derivation of Expected Indirect Utility $v_d^l$

Given taste shocks $z_{odi}$, the expected indirect utility is expressed as $v_d^l = \prod_{k\in\{S,M\}} \int w_d^l / \left(\prod_{k\in\{S,M\}} \left(a_k \times p_k^{\alpha_k}\right)\right) dp_k$ (since the prices are independently drawn), which could be derived as:

$$
\begin{aligned}
v_d^l &= w_d^l \prod_{k\in\{S,M\}} a_k^{-1} \int_0^{+\infty} p_k^{-\alpha_k} d\tilde{G}(p_k) \\
&= w_d^l \prod_{k\in\{S,M\}} a_k^{-1} \int_0^{+\infty} p_k^{-\alpha_k} \tilde{v}_{k,d}\theta(1-\tilde{\gamma}) p_k^{\theta(1-\tilde{\gamma})-1} exp\left[-\tilde{v}_{k,d} p_k^{\theta(1-\tilde{\gamma})}\right] dp_k \\
&= w_d^l \prod_{k\in\{S,M\}} a_k^{-1} \tilde{v}_{k,d}^{\frac{\alpha_k}{\theta(1-\tilde{\gamma})}} \Gamma\left(1 - \frac{\alpha_k}{\theta(1-\tilde{\gamma})}\right) \\
&= B_d \times w_d^l \times \mathrm{Y}_d^{\frac{1}{\theta}}
\end{aligned}
$$

where we require $\alpha_k < \theta(1-\tilde{\gamma})$ to make the Gamma function well-defined, and choose $a_k = (1-\tilde{\gamma})^{-\frac{\alpha_k}{\theta(1-\tilde{\gamma})}} \times \Gamma\left(1 - \frac{\alpha_k}{\theta(1-\tilde{\gamma})}\right)$ to eliminate the effect of need $k$ on utility. In the third equality, we use the transformation $x_{k,d} \equiv \tilde{v}_{k,d} p^{\theta(1-\tilde{\gamma})}$.

## Derivation of Migration Share

The migration share from $o$ to $d$ is denoted as $\delta_{od}$ that is derived as:

$$
\begin{aligned}
\delta^l_{od} &= \mathbf{Pr}\Big(\frac{\nu^l_d z_{di}}{e^l_{od}} \geq \frac{\nu^l_g z_{gi}}{e^l_{og}},\ \forall g\Big) \\
&= \mathbf{Pr}\Big(z_{gi} \leq \frac{\nu^l_d/e^l_{od}}{\nu^l_g/e^l_{og}} z_{di},\ \forall g\Big) \\
&= \int_0^{+\infty} \left[ \int\!\!\cdots\!\!\int_{z_{gi}\leq \frac{\nu^l_d/e^l_{od}}{\nu^l_g/e^l_{og}} z_{di}, \forall g\neq d} \frac{\partial^n F(\mathbf{z}_{-d}, z_d)}{\partial z_1 \partial z_2 ... \partial z_N} d\mathbf{z}_{-d} \right] dz_d \\
&= \int_0^{+\infty} \left\{ \int\!\!\cdots\!\!\int_{z_{gi}\leq \frac{\nu^l_d/e^l_{od}}{\nu^l_g/e^l_{og}} z_{di}, \forall g\neq d,j} \left[ \int_0^{\frac{\nu^l_d/e^l_{od}}{\nu^l_j/e^l_{oj}}} \left( d\frac{\partial^{n-1} F(\mathbf{z}_{-d-j}, z_j, z_d)}{\underbrace{\partial z_1 \partial z_2 ... \partial z_g ... \partial z_N}_{g\neq d,j}} / dz_j \right) dz_j \right] d\mathbf{z}_{-d-j} \right\} dz_d \\
&= \int_0^{+\infty} \left\{ \int\!\!\cdots\!\!\int_{z_{gi}\leq \frac{\nu^l_d/e^l_{od}}{\nu^l_g/e^l_{og}} z_{di}, \forall g\neq d,j} \left[ \frac{\partial^{n-1} F(\mathbf{z}_{-d-j}, \frac{\nu^l_d/e^l_{od}}{\nu^l_j/e^l_{oj}} z_{di}, z_d)}{\underbrace{\partial z_1 \partial z_2 ... \partial z_g ... \partial z_N}_{g\neq d,j}} \right] d\mathbf{z}_{-d-j} \right\} dz_d \\
&\ldots\ldots \\
&= \int_0^{+\infty} \frac{\partial F(\frac{\nu^l_d/e^l_{od}}{\nu^l_1/e^l_{o1}} z_{di}, ..., \frac{\nu^l_d/e^l_{od}}{\nu^l_N/e^l_{oN}} z_{di})}{\partial z_d} dz_d
\end{aligned}
$$

where $\mathbf{z}_{-d}$ denotes the vector of $\{z_{gi}\}$ which does not include $z_{di}$, and the same notation meaning for $\mathbf{z}_{-d-j}$.

Given the formula of $F$, it follows that

$$
\begin{aligned}
\delta^l_{od} &= \int_0^{+\infty} \left[ \sum_g \left( \frac{\nu^l_d/e^l_{od}}{\nu^l_g/e^l_{og}} \right)^{-\epsilon_l} z_{di}^{-\epsilon_l} \right]^{-\rho} (1-\rho)\epsilon_l z_{di}^{-\epsilon_l-1} exp\left\{ -z_{di}^{-\epsilon_l(1-\rho)} \times \left[ \sum_g \left( \frac{\nu^l_d/e^l_{od}}{\nu^l_g/e^l_{og}} \right)^{-\epsilon_l} \right]^{1-\rho} \right\} dz_{di} \\
&= \int_0^{+\infty} \frac{1}{\sum_g \left( \frac{\nu^l_d/e^l_{od}}{\nu^l_g/e^l_{og}} \right)^{-\epsilon_l}} \times d exp\left\{ -z_{di}^{-\epsilon_l(1-\rho)} \times \left[ \sum_g \left( \frac{\nu^l_d/e^l_{od}}{\nu^l_g/e^l_{og}} \right)^{-\epsilon_l} \right]^{1-\rho} \right\} \\
&= \frac{\left(\nu^l_d/e^l_{od}\right)^{\epsilon_l}}{\sum_g \left(\nu^l_g/e^l_{og}\right)^{\epsilon_l}}
\end{aligned}
$$

## Derivation of Expected Utility

The joint probability of utility for worker of type $l$ from city $o$ is denoted by vector $\mathbf{u}^l_o = \{u^l_{od} | d = 1, 2, ...N\}$ where $u^l_{od} \equiv \frac{z_{odi}\nu^l_d}{e^l_{od}}$. Then the joint distribution of utility is derived as

$$
\begin{aligned}
U_o^l(\mathbf{u}_o^l) &= \mathbf{Pr}\left(u_{o1}^l \le u_1, ..., u_{oN}^l \le u_N\right) \\
&= \mathbf{Pr}\left(..., z_{odi} \le \frac{e_{od}^l u_d}{\nu_d^l}, ...\right) \\
&= exp\left\{-\left[\sum_d \left(\frac{e_{od}^l u_d}{\nu_d^l}\right)^{-\epsilon_l}\right]^{1-\rho}\right\}
\end{aligned}
$$

Therefore, the distribution of $u_o^l \equiv \max_{d\in\{1,2,..,N\}}\left\{\frac{z_{odi}\nu_d^l}{e_{od}^l}\right\}$, which is denoted as $\bar{U}_o^l(u)$:

$$
\begin{aligned}
\bar{U}_o^l(u) &= \mathbf{Pr}\left(u_o^l \le u\right) \\
&= \mathbf{Pr}\left(u_{od}^l \le u,\ \forall d\right) \\
&= U_o^l(u, u, ..., u) \\
&= exp\left\{-\left[\sum_d \left(\frac{e_{od}^l u}{\nu_d^l}\right)^{-\epsilon_l}\right]^{1-\rho}\right\} \\
&= exp\left\{-\left[\sum_d \left(\frac{e_{od}^l}{\nu_d^l}\right)^{-\epsilon_l}\right]^{1-\rho} u^{-(1-\rho)\epsilon_l}\right\} \\
&= exp\left[-\left(M_o^l\right)^{1-\rho} u^{-(1-\rho)\epsilon_l}\right]
\end{aligned}
$$

where $M_o^l \equiv \sum_d \left(\frac{e_{od}^l}{\nu_d^l}\right)^{-\epsilon_l}$. Then the expected utility born in city $o$ is then

$$
\begin{aligned}
E(u_o^l) &= \int_0^{+\infty} u d\bar{U}_o^l(u) \\
&= \int_0^{+\infty} \left(M_o^l\right)^{1-\rho}(1-\rho)\epsilon_l u^{-(1-\rho)\epsilon_l} exp\left[-\left(M_o^l\right)^{1-\rho} u^{-(1-\rho)\epsilon_l}\right] du \\
&= \left(M_o^l\right)^{\frac{1}{\epsilon_l}} \int_0^{+\infty} x^{-\frac{1}{(1-\rho)\epsilon_l}} exp\left(-x\right) dx \\
&= \left(M_o^l\right)^{\frac{1}{\epsilon_l}} \Gamma\left[1 - \frac{1}{(1-\rho)\epsilon_l}\right]
\end{aligned}
$$

where the derivation of the third equality uses the transformation $x = \left(M_o^l\right)^{1-\rho} u^{-(1-\rho)\epsilon_l}$ and the last equality comes from the definition of Gamma function.

## Proof of Proposition 3

To start, we begin by constructing the vector of function $\mathbf{F}(\boldsymbol{\nu}^l)$ such that the $d$th element is defined as

$$
F_d(\boldsymbol{\nu}^l) = L_d^l - \sum_o \delta_{od}^l \bar{L}_o^l, \quad \delta_{od}^l = \frac{\left(\frac{\nu_d^l}{e_{od}^l}\right)^{\epsilon_l}}{\sum_g \left(\frac{\nu_g^l}{e_{og}^l}\right)^{\epsilon_l}} \tag{54}
$$

Next, we need to show the following:

1. $\mathbf{F}(\boldsymbol{\nu}^l)$ is continuous in $\boldsymbol{\nu}^l$;

This is verified by the function properties of (54).

2. $\mathbf{F}(\boldsymbol{\nu}^l)$ is homogeneous of degree zero in $\boldsymbol{\nu}^l$;

For $\forall\lambda > 0$,$F_d(\lambda\boldsymbol{\nu}^l) = L_d^l - \sum_o \left[\left(\frac{\lambda\nu_d^l}{e_{od}^l}\right)^{\epsilon_l} / \sum_g \left(\frac{\lambda\nu_g^l}{e_{og}^l}\right)^{\epsilon_l}\right]\bar{L}_o^l = L_d^l - \sum_o \left[\left(\frac{\nu_d^l}{e_{od}^l}\right)^{\epsilon_l} / \sum_g \left(\frac{\nu_g^l}{e_{og}^l}\right)^{\epsilon_l}\right]\bar{L}_o^l = F_d(\boldsymbol{\nu}^l)$. The homogeneity of degree zero is verified.

3. $\sum_d F_d(\boldsymbol{\nu}^l) = 0,\ \forall\boldsymbol{\nu}^l \in R_+^N$;

$\sum_d F_d(\boldsymbol{\nu}^l) = \sum_d L_d^l - \sum_d\sum_o \delta_{od}^l\bar{L}_o^l = \sum_d L_d^l - \sum_o\sum_d \delta_{od}^l\bar{L}_o^l = \sum_d L_d^l - \sum_o \bar{L}_o^l = 0$, where the last equality comes from the fact that $\sum_d \delta_{od}^l = 1\ \forall o$.

4. $\mathbf{F}(\boldsymbol{\nu}^l)$ exhibits gross substitution in $\boldsymbol{\nu}^l$.

We can check this property by computing derivatives

$$\frac{dF_d(\boldsymbol{\nu}^l)}{d\nu_d^l} = -\sum_o \frac{d\delta_{od}^l}{d\nu_d^l}\bar{L}_o^l = -\sum_o \frac{\epsilon_l\left(\nu_d^l\right)^{\epsilon_l-1}\left(e_{od}^l\right)^{-\epsilon_l}\left[\sum_{g\neq d}\left(\frac{\nu_g^l}{e_{og}^l}\right)^{\epsilon_l}\right]}{\left[\sum_g\left(\frac{\nu_g^l}{e_{og}^l}\right)^{\epsilon_l}\right]^2}\bar{L}_o^l < 0$$

On the other hand, for region $s \neq d$

$$\frac{dF_d(\boldsymbol{\nu}^l)}{d\nu_s^l} = \sum_o \frac{d\delta_{od}^l}{d\nu_d^l}\bar{L}_o^l = \sum_o \frac{\left(\frac{\nu_g^l}{e_{og}^l}\right)^{\epsilon_l}\epsilon_l\left(\nu_s^l\right)^{\epsilon_l-1}\left(e_{os}^l\right)^{-\epsilon_l}}{\left[\sum_g\left(\frac{\nu_g^l}{e_{og}^l}\right)^{\epsilon_l}\right]^2}\bar{L}_o^l > 0$$

By the homogeneity conditions, we follow Michaels, Redding, and Rauch (2011) by restricting the solution to the unit simplex $\Delta \equiv \{\boldsymbol{\nu}^l|\sum_d \nu_d^l = 1\}$ and construct $\mathbf{F}^+(\boldsymbol{\nu}^l)$ on $\Delta$ such that $\mathbf{F}^+(\boldsymbol{\nu}^l) = max\{0,\mathbf{F}(\boldsymbol{\nu}^l)\}$, and $\mathbf{H}(\boldsymbol{\nu}^l) = \left[\boldsymbol{\nu}^l+\mathbf{F}^+(\boldsymbol{\nu}^l)\right] / \sum_d\left[\nu_d^l + F_d^+(\boldsymbol{\nu}^l)\right]$ .Then $\mathbf{H}(\boldsymbol{\nu}^l)$ is the continuous function mapping from $\Delta$ to $\Delta$, which implies there exist a fixed points by Brouwer's fixed point theorem. To summarize, condition (1) and (2) guarantee the existence and (3) and (4) guarantee the uniqueness. The detailed proof steps refer to Michaels, Redding, and Rauch (2011). Given the bilateral migration cost, we can solve (54) for $\boldsymbol{\nu}^l$.

### Equilibrium Condition in Matrix Format

We write (39) in matrix form as:

$$\mathbf{Y} = \mathbf{\Pi}'\left(\mathbf{X}^C + \mathbf{\Phi}^I\mathbf{Y}\right)$$

where

$$\mathbf{Y}=\begin{bmatrix} Y_1\\ Y_2\\ \cdot\\ \cdot\\ \cdot\\ Y_N \end{bmatrix},\quad \mathbf{X}^C=\begin{bmatrix} X_1^C\\ X_2^C\\ \cdot\\ \cdot\\ \cdot\\ X_N^C \end{bmatrix},\quad \mathbf{\Phi}^I=\begin{bmatrix} \Phi_1^I & 0 & ... & 0 & 0\\ 0 & \Phi_2^I & ... & 0 & 0\\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot\\ 0 & 0 & ... & \Phi_{N-1}^I & 0\\ 0 & 0 & ... & 0 & \Phi_N^I \end{bmatrix}$$

and

$$\mathbf{\Pi}=\begin{bmatrix} \pi_{11} & \pi_{12} & ... & \pi_{1,N-1} & \pi_{1,N}\\ \pi_{21} & \pi_{22} & ... & \pi_{2,N-1} & \pi_{2,N}\\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot\\ \pi_{N-1,1} & \pi_{N-1,2} & ... & \pi_{N-1,N-1} & \pi_{N-1,N}\\ \pi_{N,1} & \pi_{N,2} & ... & \pi_{N,N-1} & \pi_{N,N} \end{bmatrix}$$

### Extension: Proof of Equivalence ($\zeta_k \to 1^-$ )

The proof is completed by showing that $\lim_{\zeta_k \to 1^-} \left(1/(1-\zeta_k)\right)^{1-\zeta_k} = 1$.

$$\begin{aligned} \lim_{\zeta_k \to 1^-} \left(1/(1-\zeta_k)\right)^{1-\zeta_k} &= \lim_{x\to 0^+} (1/x)^x \\ &= \lim_{x\to 0^+} exp\left(-x\ln x\right) \\ &= \lim_{x\to 0^+} exp\left(-\frac{\ln x}{1/x}\right) \\ &= \lim_{x\to 0^+} exp\left(\frac{1/x}{1/x^2}\right) \\ &= \lim_{x\to 0^+} exp\left(x\right) = 1 \end{aligned}$$

where the first equality uses the transformation $x \equiv 1-\zeta_k$; and the fourth equality applies L'Hopital's rule given that it is an indeterminate form of type $\frac{\infty}{\infty}$.

### Extension: Solution to the Extended Model

Since firm-to-firm matching for intermediate and the competition between unskilled labors and intermediate inputs for fulfilling manufacturing tasks remain unchanged, the cost distribution of $m_{k,i}(j)$ would be the same as the one in the benchmark:

$$G_{k,i}(c) = 1 - exp\left(-\Xi_{k,i}c^{\phi}\right)$$

where $\Xi_{k,i} \equiv v_{k,i} + w_{k,i}^{-\phi}$ and $v_{k,i} \equiv \frac{\theta}{\phi}\lambda_{k,i}Y_i^{\phi/\theta}$. Denote the cost of $m_{k,i}(j)$ as $c_{k,i}$, so the unit cost of this composite bundle $\left(m_{k,i}(j)/\zeta_k\right)^{\zeta_k}\left(L_{k,i}/(1-\zeta_k)\right)^{1-\zeta_k}$ would be $c_{k,i}^m = \left(c_{k,i}\right)^{\zeta_k}\left(w_{k,i}\right)^{1-\zeta_k}$. The cost distribution of such composite bundle is expressed as:

$$
\begin{aligned}
G^m_{k,i}(c) &= \mathbf{Pr}\left(c^m_{k,i} \leq c\right) \\
&= \mathbf{Pr}\left(c_{k,i} \leq c^{1/\zeta_k}\left(w_{k,i}\right)^{-(1-\zeta_k)/\zeta_k}\right) \\
&= G_{k,i}\left(c^{1/\zeta_k}\left(w_{k,i}\right)^{-(1-\zeta_k)/\zeta_k}\right) \\
&= 1 - exp\left(-\Xi_{k,i}\left(w_{k,i}\right)^{-\phi(1-\zeta_k)/\zeta_k} c^{\phi/\zeta_k}\right)
\end{aligned}
$$

$G^m_{k,i}(c)$ denotes the new cost distribution for task $k$ in region $i$ ($G^m_{k,i}(c)$ will be coincide with $G_{k,i}(c)$ when $\zeta_k$ is unity), with which we derive the measure of firms from $i$ to $n$ with costs less than $c$:

$$
\begin{aligned}
\mu_{ni}(c) &= E\left\{\mu^Z_i\left(\frac{d_{ni}}{c}\prod_{k\in\{S,M\}}\frac{\left(c^m_{k,i}\right)^{\beta_k}}{b_k}\right)\right\} \\
&= T_i d_{ni}^{-\theta} c^\theta \prod_{k\in\{S,M\}} \int_0^{+\infty} b_k^\theta c_k^{-\theta\beta_k} dG^m_{k,i}(c_k) \\
&= T_i d_{ni}^{-\theta} c^\theta \prod_{k\in\{S,M\}} b_k^\theta \left(\Xi_{k,i}\right)^{\theta\beta_k\zeta_k/\phi} \int_0^{+\infty} x^{-\theta\beta_k\zeta_k/\phi} exp(-x)dx \\
&= T_i d_{ni}^{-\theta} c^\theta \prod_{k\in\{S,M\}} b_k^\theta \left(\Xi_{k,i}\right)^{\theta\beta_k\zeta_k/\phi} \left(w_{k,i}\right)^{-(1-\zeta_k)\theta\beta_k} \Gamma\left(1-\frac{\theta\beta_k\zeta_k}{\phi}\right) \\
&= T_i \Xi_i d_{ni}^{-\theta} c^\theta
\end{aligned}
$$

where we define $x_k \equiv \Xi_{k,i}\left(w_{k,i}\right)^{-\phi(1-\zeta_k)/\zeta_k} c^{\phi/\zeta_k}$ in the third equality. The last equality is derived by choosing $b_k = \Gamma(1-\frac{\theta\beta_k\zeta_k}{\phi})^{-\frac{1}{\theta}}$, which require $\theta\beta_k\zeta_k/\phi < 1$ in order to make $\mu_{ni}(c)$ meaningful (this restriction is less demanding than the one used in our benchmark given that $\zeta_k \leq 1$). We define $\Xi_i \equiv \prod_{k\in\{S,M\}}\left(\Xi_{k,i}\right)^{\theta\beta_k\zeta_k/\phi}\left(w_{k,i}\right)^{-(1-\zeta_k)\theta\beta_k}$ to obtain the last equality. Finally the measure of firms active in region $n$ is:

$$
\mu_n(c) = \sum_i \mu_{ni}(c) = \Upsilon_n c^\theta
$$

where $\Upsilon_n$ solves the following fixed point problem:

$$
\Upsilon_n = \sum_i T_i d_{ni}^{-\theta} \prod_{k\in\{S,M\}} \left(\frac{\theta}{\phi}\lambda_{k,i} w_{k,i}^{-(1-\zeta_k)\phi/\zeta_k} \Upsilon_i^{\phi/\theta} + w_{k,i}^{-\phi/\zeta_k}\right)^{\theta\beta_k\zeta_k/\phi}
$$

Following the proof method explained in our benchmark model, we can solve the vector of $\Upsilon_n$ by solving the fixed point of the following system of equations (proof is provided in the subsequent section in Appendix).

Similar to Proposition 2, the above assumptions over technology are consistent with an extended aggregate production function for region $i$, which satisfies the function form:

$$Q_i = \{L_{S,i}\}^{\beta_S} \left\{ \left[ \left( \psi \left(L_{1M,i}\right)^{\frac{\phi}{\phi+1}} + (1-\psi) \left(I_{M,i}\right)^{\frac{\phi}{\phi+1}} \right)^{\frac{\phi+1}{\phi}} \right]^{\zeta_M} \left[L_{2M,i}\right]^{1-\zeta_M} \right\}^{\beta_M}$$

where $L_{S,i}$ denotes skilled labors employed in fulfilling service task; $L_{1M,i}$ are unskilled labors employed in producing manufacturing inputs (subject to outsourcing); $L_{2M,i}$ denotes the unskilled labors employed for the labor component in manufacturing task (not subject to outsourcing); $I_{M,i}$ are the intermediates used for producing manufacturing inputs, and $\psi \equiv 1/\left[1 + \Gamma(1 + 1/\phi)^{\phi/(1+\phi)}\right]$.

Similar to our benchmark model, labor share in producing intermediate inputs for manufacturing task in region $i$ remains $w_{M,i}^{-\phi}/\Xi_{M,i}$ . The aggregate share of labor of type $l$ in the total production cost in region $i$ is

$$\beta_i^{l(k)} = \beta_k \zeta_k \frac{w_{k,i}^{-\phi}}{\Xi_{k,i}} + \beta_k \left(1 - \zeta_k\right)$$

Consequently, the overall labor share of in the production in region $i$ is consequently written as:

$$\begin{aligned} \beta_i^L &= \sum_{k \in \{S,M\}} \beta_i^{l(k)} \\ &= \beta_S + \beta_M \left(1 - \zeta_M + \zeta_M \frac{w_{M,i}^{-\phi}}{\Xi_{M,i}}\right) \end{aligned}$$

Next, the general equilibrium conditions are characterized as following (same with the ones in benchmark model). With the balanced trade, the total expenditure on final goods is equal to labor income, since there is no profit in this model. For region $n$, we have

$$X_n^C = \sum_{l \in \{B,W\}} w_n^l L_n^l$$

The aggregate production in region $i$ consists of the total revenue from selling the consumption goods and that in supplying the intermediates around the world:

$$Y_i = \sum_{n=1}^{N} \pi_{ni} \left[X_n^c + \Phi_n^I Y_n\right]$$

where $\Phi_n^I \equiv 1 - \beta_n^L$ denotes the shares of intermediate inputs in final production, and $\beta_n^L$ denotes the overall labor share of in production costs that is defined above. The equilibrium for labor of type $l \in \{B, W\}$ in region $i$ implies:

$$w_i^l L_i^l = \beta_i^l Y_i$$

where $\beta_i^l$ denotes the production cost share of labor of type $l$ in region $i$. $L_i^l$ is the endogenous supply of workers of type $l$ in region $i$, provided in (35). These sets of equations for each type of labor $l$ in region $i$ solves the wage $w_i^l$.

**Extension: Proof of Fixed Point of** $Y$

To show there is a fixed point, we transform the fixed point problem in form of $\ln Y = F(\ln Y)$, where $F(x) \equiv \ln f(exp(x))$.

$$\begin{aligned} F_n(x) &= \ln\left\{\sum_i T_i d_{ni}^{-\theta} \prod_{k\in\{S,M\}} \left[\frac{\theta}{\phi}\lambda_{k,i} w_{k,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}x) + w_{k,i}^{-\phi/\zeta_k}\right]^{\theta\beta_k\zeta_k/\phi}\right\} \\ &= \ln\left\{\sum_i T_i d_{ni}^{-\theta} w_{S,i}^{-\theta\beta_S} \left[\frac{\theta}{\phi}\lambda_{M,i} w_{M,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}x) + w_{M,i}^{-\phi/\zeta_k}\right]^{\theta\beta_M\zeta_M/\phi}\right\} \\ &= \ln\left\{\sum_i exp\left(\ln\left(Td_{ni}^{-\theta}\right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M\zeta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i} w_{M,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}x) + w_{M,i}^{-\phi/\zeta_k}\right]\right)\right\} \end{aligned}$$

where the second equality comes from the assumption that the service task cannot be outsourced, $\lambda_{S,i} = 0$. To apply the Blackwell's contraction mapping theory, we need to verify both monotonicity and discounting properties for $F(x)$ to confirm the sufficient conditions.

For monotonicity, let $x$ and $y$ satisfy that $x \leq y$, as all the exogenous parameters are positive then

$$\begin{aligned} F_n(x) &= \ln\left\{\sum_i exp\left(\ln\left(Td_{ni}^{-\theta}\right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M\zeta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i} w_{M,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}x) + w_{M,i}^{-\phi/\zeta_k}\right]\right)\right\} \\ &\leq \ln\left\{\sum_i exp\left(\ln\left(Td_{ni}^{-\theta}\right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M\zeta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i} w_{M,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}y) + w_{M,i}^{-\phi/\zeta_k}\right]\right)\right\} \\ &= F_n(y) \end{aligned}$$

For discounting, for $\forall a > 0$

$$\begin{aligned} F_n(x+a) &= \ln\left\{\sum_i exp\left(\ln\left(Td_{ni}^{-\theta}\right) - \theta\beta_S \ln w_{S,i} + \frac{\theta\beta_M\zeta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i} w_{M,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}a) \times exp(\frac{\phi}{\theta}x) + w_{M,i}^{-\phi/\zeta_k}\right]\right)\right\} \\ &= \ln\left\{\sum_i exp\left(... + \zeta_M\beta_M a + \frac{\theta\beta_M\zeta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i} w_{M,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}x) + exp(-\frac{\phi}{\theta}a) w_{M,i}^{-\phi/\zeta_k}\right]\right)\right\} \\ &= \zeta_M\beta_M a + \ln\left\{\sum_i exp\left(... + \frac{\theta\beta_M\zeta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i} w_{M,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}x) + exp(-\frac{\phi}{\theta}a) w_{M,i}^{-\phi/\zeta_k}\right]\right)\right\} \\ &\leq \zeta_M\beta_M a + \ln\left\{\sum_i exp\left(... + \frac{\theta\beta_M}{\phi}\ln\left[\frac{\theta}{\phi}\lambda_{M,i} w_{M,i}^{-\phi(1-\zeta_k)/\zeta_k} exp(\frac{\phi}{\theta}x) + w_{M,i}^{-\phi/\zeta_k}\right]\right)\right\} \\ &= \zeta_M\beta_M a + F_n(x) \end{aligned}$$

where the inequality comes from that $exp(-\frac{\phi}{\theta}a) < 1$. As $\beta_M < 1$ and $\zeta_M \leq 1$, discounting is verified. So there exists a fixed point.

Table A.1: Model Parameters and Description

| Symbol | Description |
|---|---|
| Consumer Preference | |
| $\alpha_M$ | Consumer's expenditure share on Manufacturing goods |
| $\alpha_S$ | Consumer's expenditure share on Service goods |
| Production Technology | |
| $\beta_M$ | Cost share on Manufacturing task |
| $\beta_S$ | Cost share on Service task |
| $T_i$ | Endowments of technology in region $i$ |
| $\theta$ | Shape parameter of the Frèchet distribution for firm productivity |
| $\phi$ | Shape parameter of the distribution for worker productivity |
| $\bar{L}_i^B$, $\bar{L}_i^W$ | Endowments of Blue ($B$) and White ($W$) collar labor in region $i$ |
| Matching Technology | |
| $\gamma$ | Shape parameter of matching density between firms and firms |
| $\lambda_{M,d}$ | Efficiency parameter for firm-firm matching in Manufacturing sector of region $d$ |
| $\lambda_{S,d}$ | Efficiency parameter for firm-firm matching in Service sector of region $d$ |
| $\tilde{\gamma}$ | Shape parameter of matching density between firms and consumers |
| $\tilde{\lambda}_{M,d}$ | Efficiency parameter for firm-consumer matching in the Manufacturing sector of region $d$ |
| $\tilde{\lambda}_{S,d}$ | Efficiency parameter for firm-consumer matching in the Service sector of region $d$ |
| Trade Cost Parameter | |
| $d_{ni}$ | Iceberg trade cost from $i$ to $n$ |
| Migration Parameter | |
| $\rho$ | Inter-region correlation of taste draws |
| $\epsilon_B$, $\epsilon_W$ | Taste dispersions of Blue ($B$) and White ($W$) collar workers |
| $e_{od}^B$, $e_{od}^W$ | Migration cost of moving from $o$ to $d$ for Blue ($B$) and White ($W$) collar workers |

Notes: Parameters used to reduce model complexity, such as $a_k$ and $b_k$ for $k \in \{M, S\}$, are not listed in the table.

# Empirical Appendix

## Construction of the Least-Cost HSR Network

The section describes the method of constructing the least-cost path networks of HSR that is used as the instrument variables for being connected to HSR in reality. The construction method follows Faber (2014) and we differ in that the objectives of HSR is to link all the provincial capitals on a single network, as stated in the official report.

Firstly, we adopt the the construction cost function as used in Faber (2014)[58],

$$c_i = 1 + Slope_i + 25 \times Built_i + 25 \times Water_i + 25 \times Wetland_i$$

where $c_i$ is the incurred cost if the HSR crosses a pixel of land $i$, $Slope_i$ stands for the average slope gradient of land $i$, $Built_i$ is a dummy variable that equals one if land $i$ is covered by any industrial, mining and residential building, $Water_i$and $Wetland_i$ indicates whether the pixel is covered by water or wetland. This simple cost function implies that the high costs are associated crossing built structures, water area as well as long and steep routes. We use the satellite sensing data on elevation, land cover and land use to compute the cost ($c_i$) for each pixel. In the calculation, we reclassify the elevation, land cover and land use grids to $1km \times 1km$ grid cells. Figure A.1 graphically demonstrate the construction cost surface. The color ranges from white (the very high cost of crossing) to the black (the very low cost of crossing).

[58]Despite the fact that the cost model in Faber (2014) is for high way system, we believe that the construction cost of HSR is qualitatively similar.

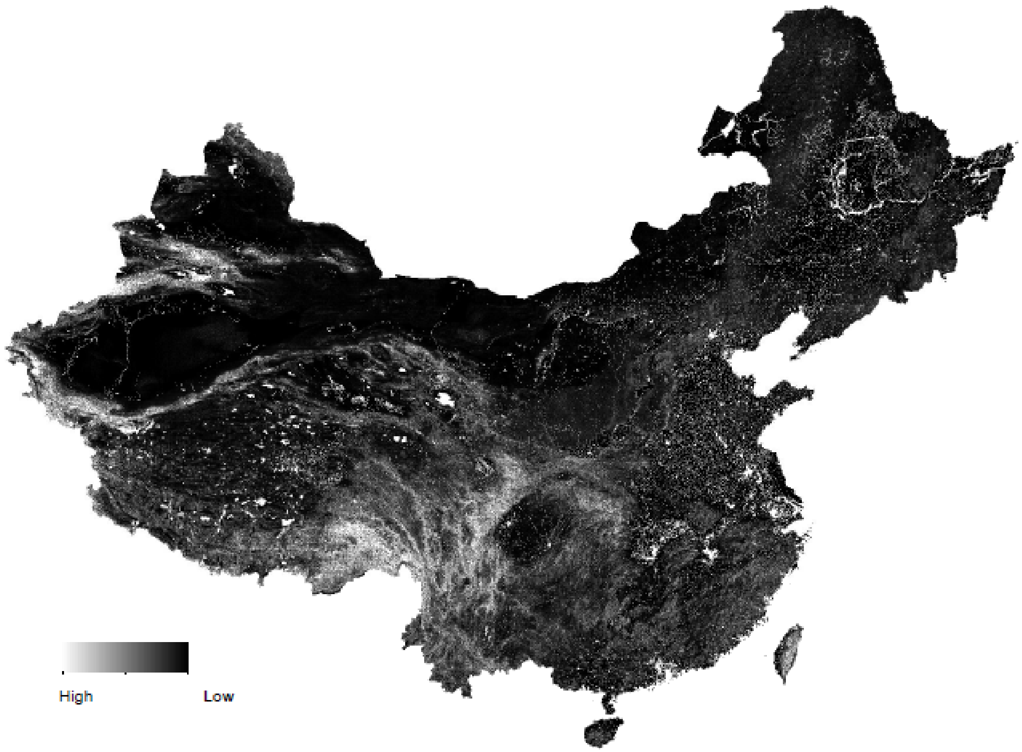
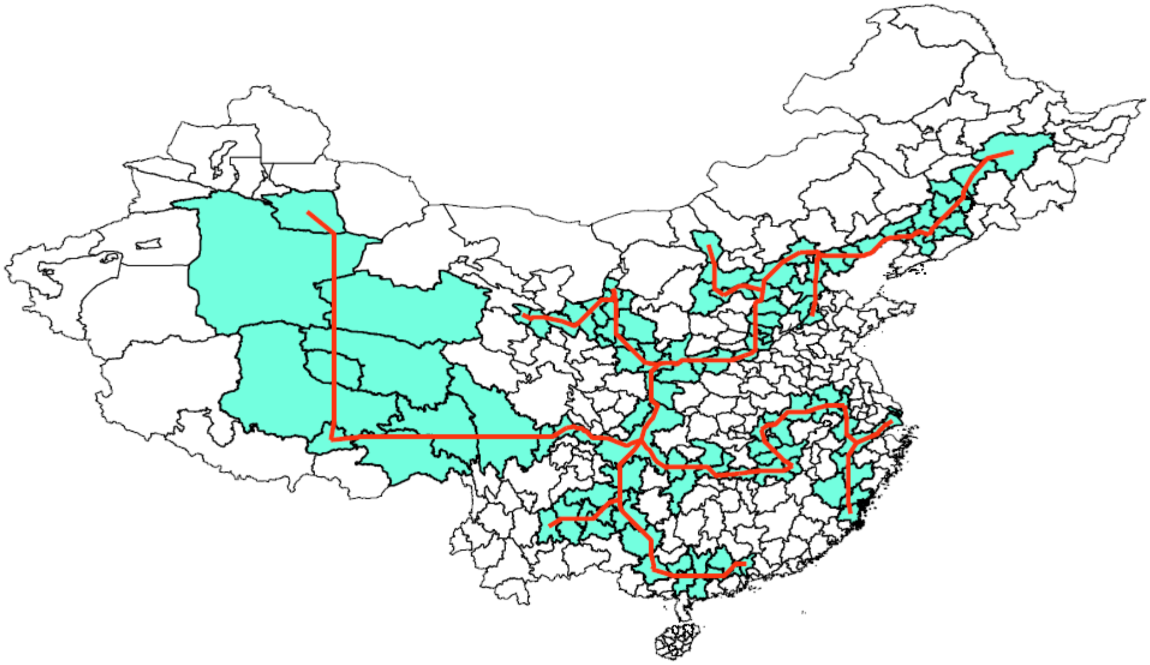


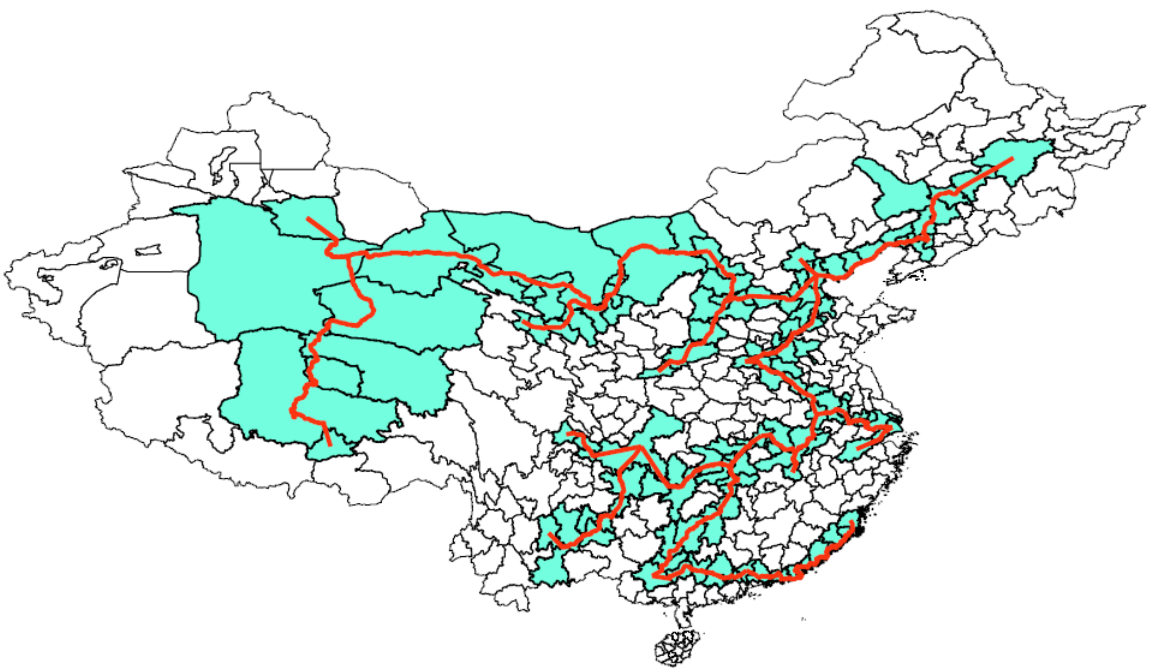

Figure A.1: The Estimated Construction Cost of high-speed Rails in China

After obtaining the constructing cost raster as the input, we then proceed to construct the least cost HSR paths between the 30 *capital cities*[59]. This generates 435 ($\frac{30*29}{2}$) possible bilateral paths between all pairs of the targeted city nodes[60]. We then extract the aggregate construction cost of each bilateral path and use it as the input to compute the Prim's minimum spanning tree algorithm.

[59]The 31 *capital cities* cover 21 provinces excluding Hainan and Taiwan, 5 autonomous regions as well as 4 municipalities.
[60]The detailed information on this procedure refer to Faber (2014).



(a) The LCP based on Euclidean Distance



(b) The LCP based on Landscape Slope

Figure A.2: Alternative Instruments to being Connected to High-Speed Rails

Table A.2: Event Study of HSR Connection on Export Performance

| Dept. var | Period 2006 - 2014 | | | | Period 2009-2014 | |
|---|---|---|---|---|---|---|
| $\ln y_{ct} - \ln y_{c0}$ | (1) | (2) | (3) | (4) | (5) | (6) |
| 3 Years Prior to Initial Connection | 0.029 | 0.055 | 0.076 | 0.133 | 0.125 | 0.062 |
| | (0.078) | (0.074) | (0.096) | (0.103) | (0.123) | (0.094) |
| 2 Years Prior to Initial Connection | 0.012 | 0.090 | 0.060 | 0.190 | 0.131 | 0.130 |
| | (0.089) | (0.093) | (0.116) | (0.136) | (0.149) | (0.135) |
| 1 Year Prior to Initial Connection | 0.031 | 0.149 | 0.066 | 0.252 | 0.161 | 0.190 |
| | (0.111) | (0.131) | (0.141) | (0.186) | (0.173) | (0.168) |
| Initial Connection | 0.061 | 0.229 | 0.108 | 0.354 | 0.227 | 0.303 |
| | (0.117) | (0.157) | (0.153) | (0.221) | (0.221) | (0.225) |
| 1 Year After Initial Connection | 0.200* | 0.374** | 0.248 | 0.513** | 0.364 | 0.459* |
| | (0.112) | (0.158) | (0.147) | (0.225) | (0.215) | (0.225) |
| 2 Years After Initial Connection | 0.221** | 0.433** | 0.263* | 0.577** | 0.402* | 0.529* |
| | (0.107) | (0.178) | (0.142) | (0.262) | (0.212) | (0.256) |
| 3 Years After Initial Connection | 0.145 | 0.384** | 0.130 | 0.494* | 0.250 | 0.440 |
| | (0.099) | (0.159) | (0.124) | (0.243) | (0.204) | (0.254) |
| Observations | 2,177 | 922 | 1,912 | 720 | 1,269 | 474 |
| R-squared | 0.325 | 0.326 | 0.311 | 0.305 | 0.281 | 0.218 |
| Group | 267 | 113 | 236 | 90 | 233 | 88 |
| None-Capital Cities | NO | NO | YES | YES | YES | YES |
| Refined Sample | NO | YES | NO | YES | NO | YES |

Notes: Each regression controls for city fixed effects. *None-Capital Cities* refers to the cities that are not provincial capital *Refine sample* refers to the cities that are either constructed or planned to connect HSR by 2015. Other controls include per capita GDP, population, average rideship and internet coverage at prefecture level. For regressions using year 2006 to 2014, we also include financial crisis shock dummy variable. Robust standard errors are clustered at the province level and reported in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.3: Spillover Effect of HSR Connection on Export Performance

| Dept. var<br>$\ln y_{ct} - \ln y_{c0}$ | (1)<br>60 km | (2)<br>120 km | (3)<br>180 km | (4)<br>240 km | (5)<br>300 km |
|---|---|---|---|---|---|
| 3 Years Prior to Initial Connection | 0.019 | 0.020 | 0.018 | -0.038 | -0.079** |
| | (0.057) | (0.034) | (0.030) | (0.036) | (0.034) |
| 2 Years Prior to Initial Connection | 0.000 | 0.050 | 0.034 | -0.048 | -0.081* |
| | (0.066) | (0.039) | (0.036) | (0.039) | (0.043) |
| 1 Year Prior to Initial Connection | -0.006 | 0.040 | 0.052 | -0.062 | -0.080 |
| | (0.082) | (0.056) | (0.065) | (0.050) | (0.061) |
| Initial Connection | 0.030 | 0.069 | 0.024 | -0.058 | -0.104 |
| | (0.085) | (0.059) | (0.072) | (0.074) | (0.066) |
| 1 Year After Initial Connection | 0.170* | 0.151*** | 0.092 | -0.005 | -0.069 |
| | (0.085) | (0.054) | (0.060) | (0.070) | (0.066) |
| 2 Years After Initial Connection | 0.203** | 0.226*** | 0.195** | 0.112* | 0.079 |
| | (0.090) | (0.079) | (0.073) | (0.063) | (0.055) |
| 3 Years After Initial Connection | 0.132 | 0.211** | 0.175** | 0.140** | 0.117* |
| | (0.092) | (0.095) | (0.077) | (0.064) | (0.058) |
| Observations | 2,177 | 2,177 | 2,177 | 2,177 | 2,177 |
| Number of Group | 267 | 267 | 267 | 267 | 267 |
| R-squared | 0.325 | 0.327 | 0.327 | 0.327 | 0.329 |

Notes: Each regression controls for city fixed effects. Other controls include per capita GDP, population, average ridership and internet coverage at prefecture level. Financial crisis shock dummy variable is included. Robust standard errors are clustered at the province level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Placebo Test: HSR Impact on Ordinary and Processing Export

| Dept. var | Ordinary Export | | | | | | Processing Export | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\ln y_{ct} - \ln y_{c0}$ | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| 3 Years Prior to Initial Connection | 0.040 | 0.067 | 0.063 | 0.111 | 0.119 | 0.050 | -0.093* | -0.076 | -0.081 | -0.039 | -0.080 | -0.082 |
| | (0.084) | (0.078) | (0.099) | (0.106) | (0.114) | (0.092) | (0.053) | (0.050) | (0.095) | (0.089) | (0.100) | (0.094) |
| 2 Years Prior to Initial Connection | 0.004 | 0.088 | 0.035 | 0.154 | 0.120 | 0.113 | 0.003 | 0.033 | -0.003 | 0.080 | 0.015 | 0.026 |
| | (0.090) | (0.091) | (0.108) | (0.128) | (0.141) | (0.135) | (0.049) | (0.046) | (0.105) | (0.100) | (0.117) | (0.111) |
| 1 Year Prior to Initial Connection | 0.033 | 0.160 | 0.051 | 0.225 | 0.169 | 0.192 | 0.004 | 0.044 | -0.001 | 0.097 | 0.011 | 0.032 |
| | (0.110) | (0.126) | (0.133) | (0.176) | (0.170) | (0.174) | (0.052) | (0.061) | (0.104) | (0.108) | (0.094) | (0.097) |
| Initial Connection | 0.082 | 0.261 | 0.140 | 0.370 | 0.266 | 0.335 | -0.045 | 0.009 | -0.095 | 0.024 | -0.060 | -0.022 |
| | (0.121) | (0.155) | (0.151) | (0.220) | (0.223) | (0.235) | (0.062) | (0.063) | (0.106) | (0.096) | (0.071) | (0.073) |
| 1 Year After Initial Connection | 0.266** | 0.450*** | 0.322** | 0.569** | 0.443** | 0.531** | -0.017 | 0.036 | -0.082 | 0.044 | -0.024 | 0.017 |
| | (0.111) | (0.152) | (0.141) | (0.219) | (0.205) | (0.224) | (0.095) | (0.089) | (0.148) | (0.124) | (0.095) | (0.102) |
| 2 Years After Initial Connection | 0.225** | 0.451** | 0.280** | 0.580** | 0.421** | 0.542** | 0.078 | 0.140 | -0.080 | 0.055 | 0.007 | 0.058 |
| | (0.105) | (0.167) | (0.136) | (0.253) | (0.200) | (0.248) | (0.111) | (0.084) | (0.159) | (0.109) | (0.117) | (0.108) |
| 3 Years After Initial Connection | 0.137 | 0.390** | 0.179 | 0.527** | 0.296 | 0.479* | 0.057 | 0.130 | -0.173 | -0.012 | -0.101 | -0.035 |
| | (0.091) | (0.145) | (0.123) | (0.235) | (0.195) | (0.249) | (0.108) | (0.095) | (0.142) | (0.145) | (0.122) | (0.140) |
| Observations | 2,177 | 922 | 1,912 | 720 | 1,269 | 474 | 1,761 | 864 | 1,497 | 662 | 990 | 434 |
| R-squared | 0.303 | 0.322 | 0.298 | 0.320 | 0.277 | 0.243 | 0.240 | 0.253 | 0.246 | 0.206 | 0.173 | 0.141 |
| Group | 267 | 113 | 236 | 90 | 233 | 88 | 213 | 106 | 183 | 83 | 181 | 81 |
| Post 2008 | NO | NO | NO | NO | YES | YES | NO | NO | NO | NO | YES | YES |
| None-Capital | NO | NO | YES | YES | YES | YES | NO | NO | YES | YES | YES | YES |
| Refined Sample | NO | YES | NO | YES | NO | YES | NO | YES | NO | YES | NO | YES |

Notes: Each regression controls for city fixed effects. *Post 2008* refers to the sample between 2009 and 2014. *None-Capital Cities* refers to the sample that are not provincial capital cities. *Refine sample* refers to the sample that are either constructed or planned to connect HSR by 2015. Other controls include per capita GDP, population, average rideship and internet coverage at prefecture level. For regressions using year 2006 to 2014, we also include financial crisis shock dummy variable. Robust standard errors are clustered at the province level and reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Welfare Gains From HSR by Provinces

| Province | Welfare Gains (%) | Province | Welfare Gains (%) | Province | Welfare Gains (%) |
|---|---|---|---|---|---|
| Shanghai | 0.36% | Shanxi | 0.10% | Hunan | 0.07% |
| Yunan | 0.04% | Guangdong | 0.17% | Gansu | 0.05% |
| Neimeng | 0.10% | Guangxi | 0.01% | Fujian | 0.06% |
| Beijing | 12.02% | Xinjiang | 0.12% | Tibet | 0.01% |
| Jilin | 0.11% | Jiangsu | 0.07% | Guizhou | 0.05% |
| Sichuan | 0.06% | Jiangxi | 0.08% | Liaoning | 0.31% |
| Tianjin | 2.22% | Hebei | 0.05% | Chongqing | 0.50% |
| Ningxia | 0.11% | Henan | 0.07% | Shaanxi | 0.08% |
| Anhui | 0.07% | Zhejiang | 0.06% | Qinghai | 0.06% |
| Shandong | 0.12% | Hubei | 0.11% | Heilongjiang | 0.08% |

Table A.6: Changes of Regional Productivity From HSR by Provinces

| Province | $T_i\Xi_i$ Changes (%) | Province | $T_i\Xi_i$ Changes (%) | Province | $T_i\Xi_i$ Changes (%) |
|---|---|---|---|---|---|
| Shanghai | 5.02% | Shanxi | 4.63% | Hunan | 4.49% |
| Yunan | 4.44% | Guangdong | 4.52% | Gansu | 4.59% |
| Neimeng | 4.66% | Guangxi | 4.12% | Fujian | 4.27% |
| Beijing | 33.40% | Xinjiang | 4.58% | Tibet | 4.45% |
| Jilin | 4.61% | Jiangsu | 4.24% | Guizhou | 4.46% |
| Sichuan | 4.51% | Jiangxi | 4.43% | Liaoning | 4.94% |
| Tianjin | 11.79% | Hebei | 4.09% | Chongqing | 5.32% |
| Ningxia | 4.54% | Henan | 4.56% | Shaanxi | 4.57% |
| Anhui | 4.45% | Zhejiang | 4.25% | Qinghai | 4.55% |
| Shandong | 4.35% | Hubei | 4.58% | Heilongjiang | 4.61% |

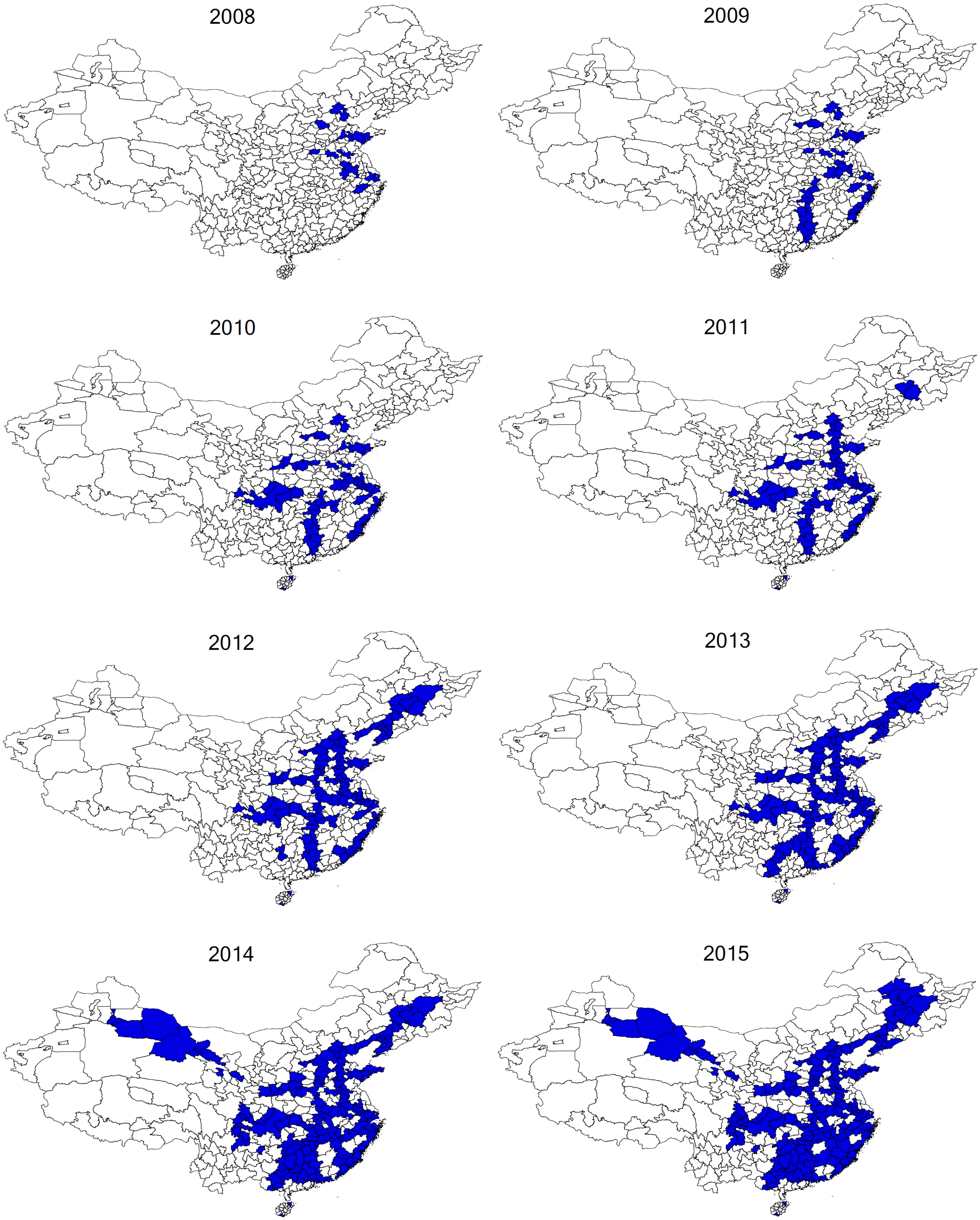



Figure A.3: Evolution of HSR Expansion from 2008 to 2015

*Note: Above figures display the evolution of prefectures that are connected to HSR network from 2008 to 2015. The area in blue are the prefectures being connected by the end of that year.*

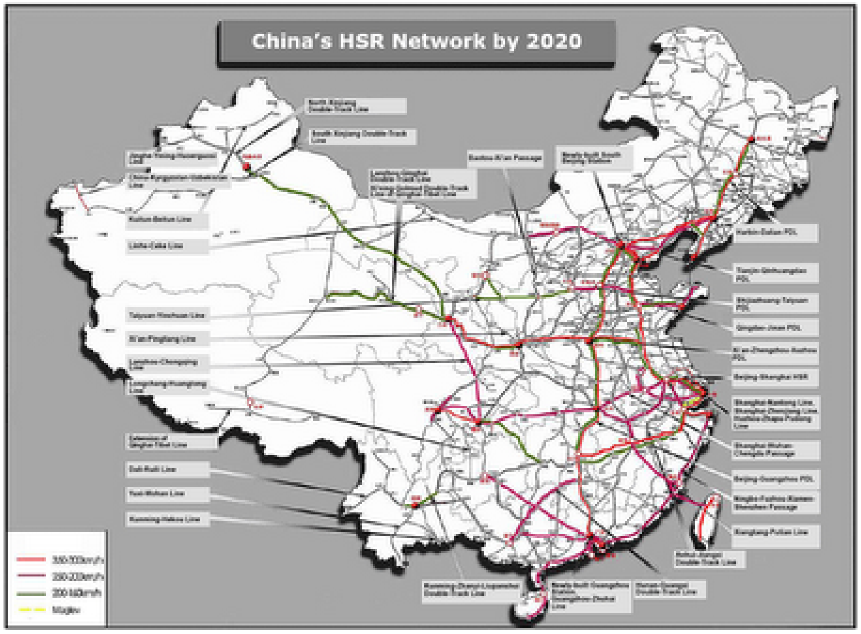



Figure A.4: Planned high-speed Railway Network by 2020

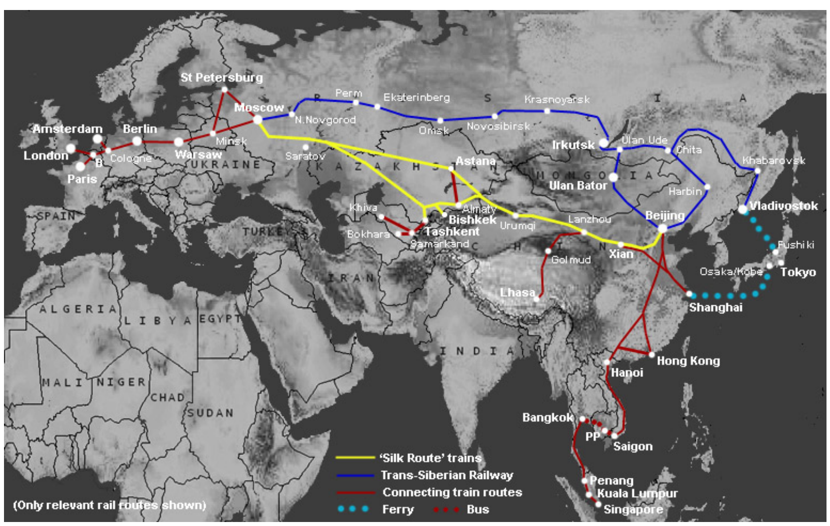
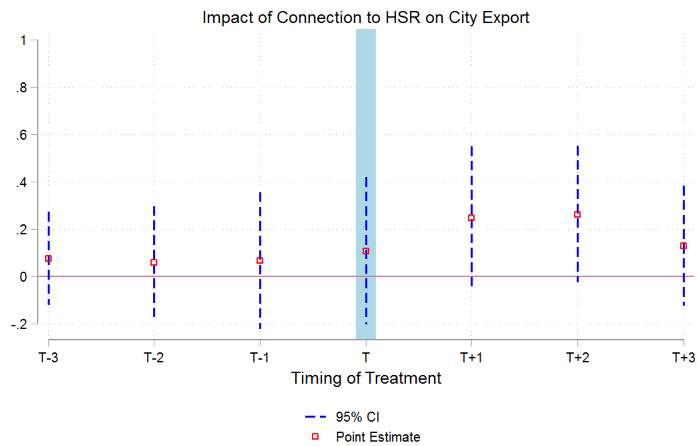
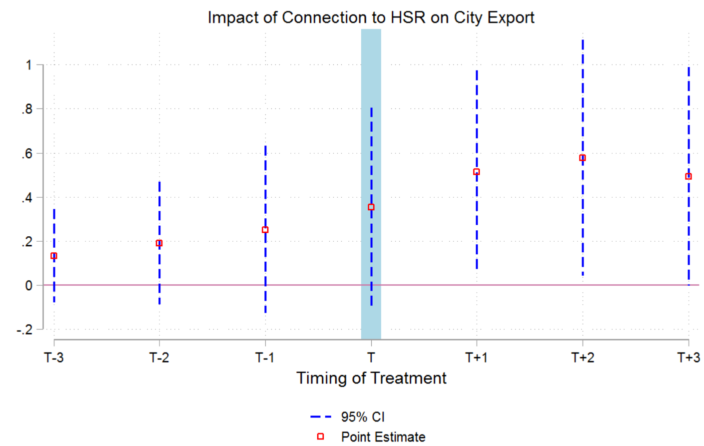



Figure A.5: New Silk Road high-speed Railway Routes

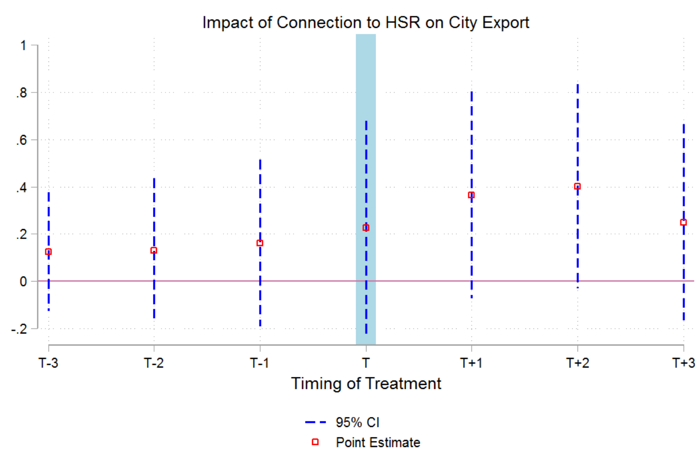
*Note: Three alternative routs are been considered and are marked with color. The figure source refer to: http://www.industrytap.com/reviving-the-silk-road-connecting-chinas-high-speed-rail-to-europe/1278*



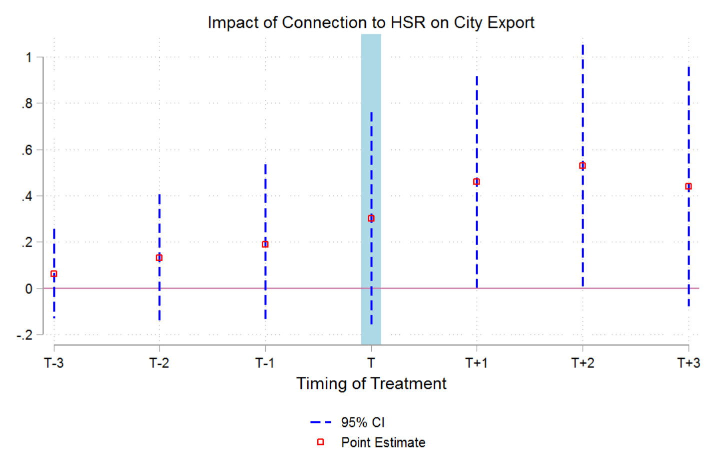

(a) Exclusion of Province Capital Cities: All Sample (2006 - 2014)




(b) Exclusion of Province Capital Cities: Refined Sample (2006 - 2014)

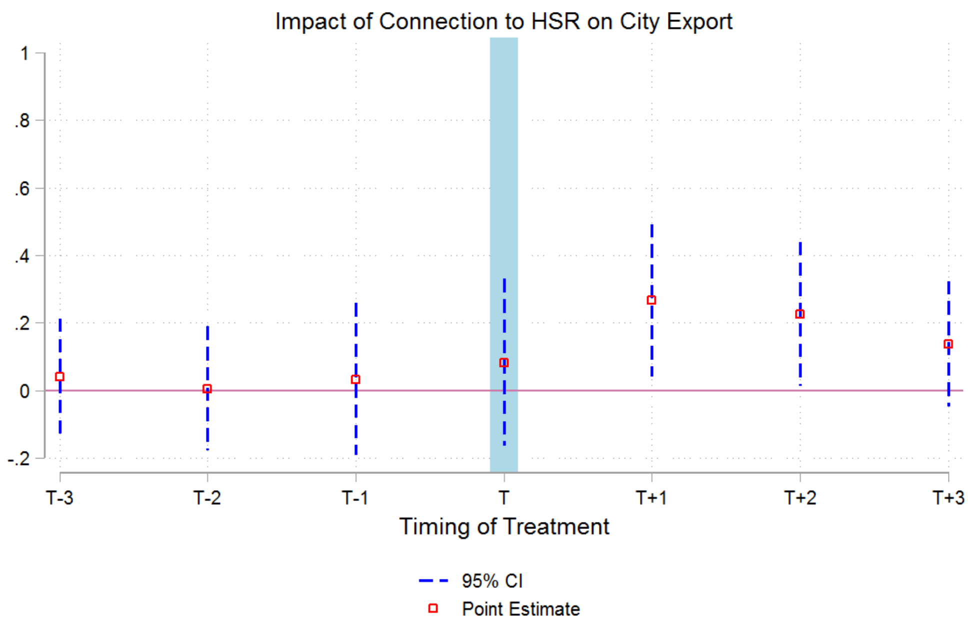



(c) Exclusion of Province Capital Cities: All Sample (2009 - 2014)

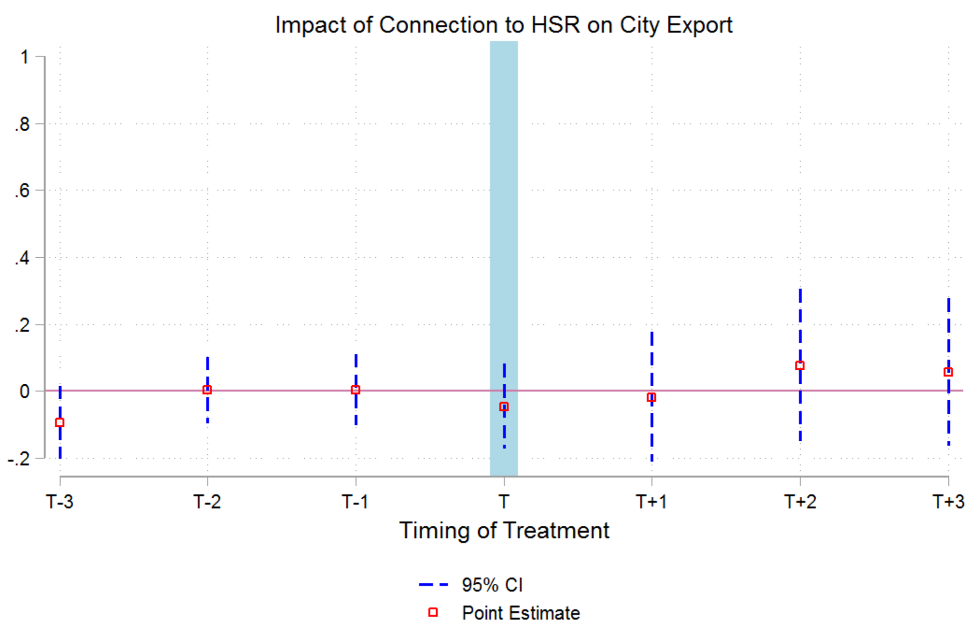



(d) Exclusion of Province Capital Cities: Refined Sample (2009 - 2014)

Figure A.6: Robustness Checks of the Event Study of HSR Connection




(a) Ordinary Export




(b) Processing Export

Figure A.7: Placebo Test: Ordinary and Processing Export

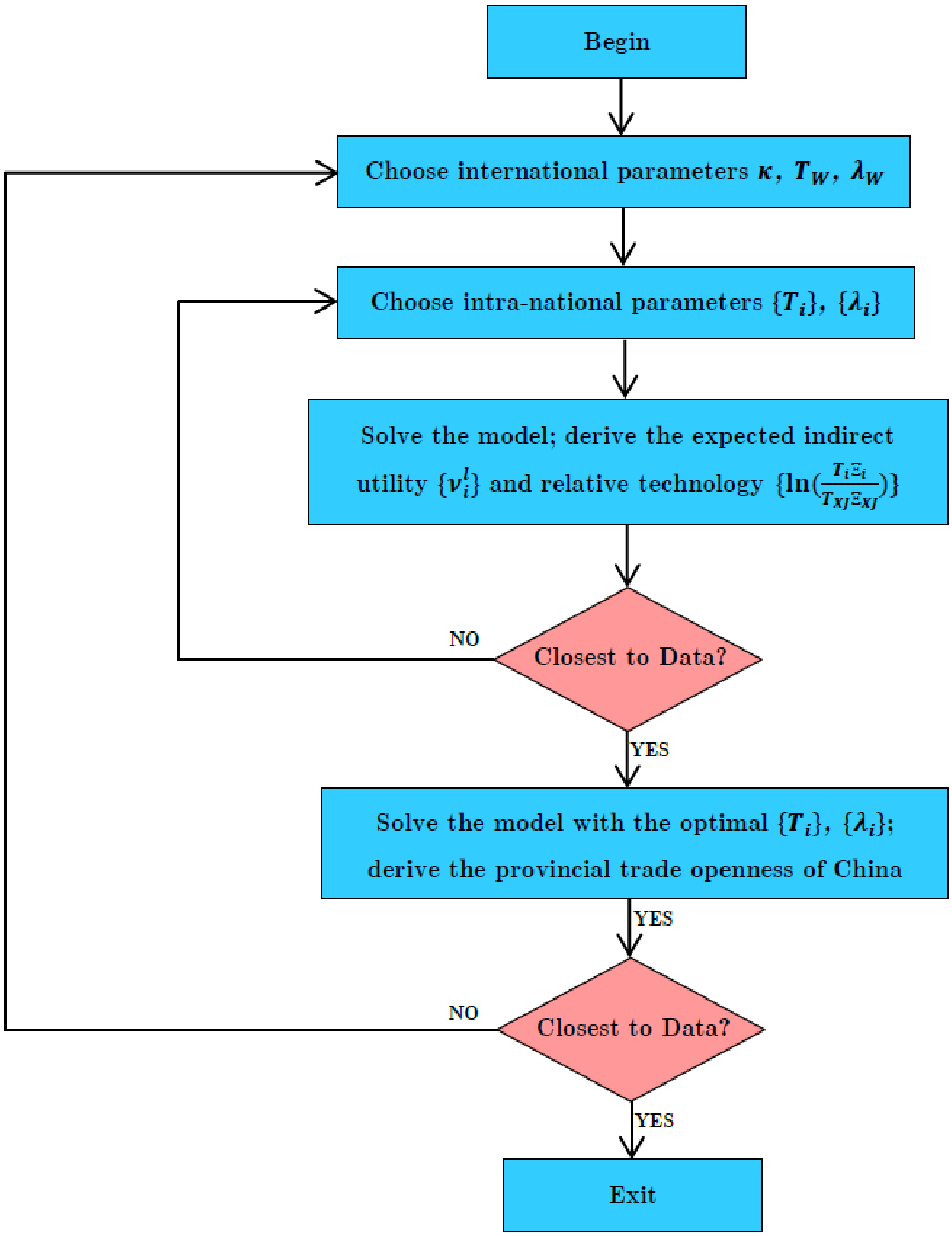



Figure A.8: Flow Chart of Calibration Algorithm: International and Intra-national Parameters