Transport Infrastructure, City Productivity Growth and Sectoral Reallocation: Evidence from China*

Yang Yang[†]

University of California, Los Angeles

This version: January 10, 2017

(Link to the most updated version)

Abstract

This paper examines the impact of highway expansion on aggregate productivity growth and sectoral reallocation between cities in China. To do so, I construct a unique dataset of bilateral transportation costs between Chinese cities, digitized highway network maps, and firm-level census. I first derive and estimate a market access measure for cities in China from 1995 to 2005. I then examine the channels through which the highway infrastructure affected economic outcomes. The results suggest that highways promoted aggregate productivity growth by facilitating entry of new firms and reallocation among existing firms. I estimate the aggregate economic impact of China's national highway system and find that eliminating all highways in China would decrease aggregate productivity by 3.2%. There is also evidence that the national highway system led to a sectoral reallocation between cities in China.

Key Words: Transport infrastructure, trade, highway, productivity, China.
JEL Classification: F10, H54, O18, O40, R10.

1 Introduction

While transport infrastructure projects are among the most expensive investments in the world, many questions remain as to whether and how transport infrastructure influences economic outcomes. Transport infrastructure facilitates interactions between cities by reducing the costs of transportation between them. We learn from trade theories that resource reallocation at both the firm and industry levels generates gains from reductions in trade barriers.[1] A large body of research in international trade has examined how the removal of trade barriers affects industries and firms, as well as the welfare implications of trade liberalization.[2] While international trade barriers have been reduced drastically over the past few decades, domestic transportation costs remain high even in developed countries.[3] An interesting and important question for researchers is the following: Are the channels suggested in the literature operational in the case of reductions in domestic transportation costs? If so, to what extent?

This paper aims to make two contributions to the literature. To the best of my knowledge, this paper is the first to examine the channels through which highways affect firm and aggregate productivity growth.[4] Previous research focuses on whether a road or railroad connection affects GDP or population growth. We still know very little about the channels through which transport infrastructure affects aggregate TFP growth. In this paper, I decompose the change in aggregate productivity resulting from highway connection into four channels: within-firm productivity growth, entry of new firms, reallocation between existing firms and exit of inefficient firms.

Second, this paper adds to the identification and estimation of the effects of transport infrastructure. This paper seeks to obtain reliable estimates of transportation costs between

---

[1] The former channel is usually called industry specialization suggested by classical trade theories such as the Ricardian model and the Heckscher-Ohlin model or the New Economic Geography models pioneered by Krugman (1991). The latter channel refers to resrouce reallocation between firms within industries suggested by more recent trade theories (Melitz 2003, Bernard et al. 2003, and Chaney 2008).

[2] There are many good references on this topic, including Melitz and Trefler (2012), Arkolakis et al. (2012), Costinot and Rodriguez-Clare (2014).

[3] See *World Development Report 2009: Reshaping Economic Geography* (2009) for a detailed account of domestic trade barriers across the world.

[4] Ghani et al. (2016) examine the effects of the Golden Quadrilateral project in India on population, GDP and labor productivity. However, they do not estimate TFP in their paper, nor do they explore the channels of productivity gains.

cities using price quote data from logistics companies. I construct a time-varying instrumental variable to address endogeneity concerns on both the *location* and *timing* of highway placement.[5] This newly developed instrumental variable approach allows me to shed light on the causal impact of highways on economic outcomes.

Evaluating the impact of a transport infrastructure project is not a straightforward exercise. One approach is to estimate or calibrate all model parameters and simulate the effects of changes in policies on economic outcomes. This structural approach has been successfully implemented in a few recent papers (Redding 2016; Nagy 2015).[6] Another approach is to use (plausibly exogenous) variation in treatment intensity within a country to examine the *relative* effects across regions (Autor et al. 2013; Donaldson 2016). With such cross-regional variation, the standard reduced-form analysis can be implemented to evaluate the effects of the project. For this paper, I first adopt the latter approach to estimate all the direct and indirect impact of highways on productivity. I then consider the counterfactual scenario of eliminating all highways to evaluate the aggregate impact of China's national highway system.[7]

I first estimate road transportation costs between any two prefecture-level cities in China. Most early studies in the literature use distance as a proxy for transportation costs (Hanson 2005; Redding and Venables 2004). A few recent papers combine digitized maps of transport routes with selected parameters to estimate transportation costs (Baum-Snow et al. 2015; Donaldson and Hornbeck 2016). I estimate transportation costs between any two prefecture-level cities in China with price quote data and digitized maps of China's highway network. This approach is novel in the literature and improves on the methods mentioned above because I use the actual prices that firms have to pay to transport goods from the origin to the destination to measure proximity to suppliers and consumers for 339 prefecture-level cities in China.

A market access measure is derived from a multi-city trade model with heterogeneous

---

[5]Whereas most studies use historical routes or planned routes as instruments for actual road construction, my instrument is time-varying and is constructed from engineering and network theory.

[6]The estimated impact, however, hinges heavily on choices of functional forms and parameters.

[7]It is interesting to consider and compare other counterfacdtual scenarios. For example, would an alternative highway system generate larger economic gains? What would happen to aggregate productivity if the railway system expanded instead of highways?

firms. I estimate city market access from 1995 to 2005 using estimated bilateral transportation costs.[8] Donaldson and Hornbeck (2016) point out that the "market access" approach addresses the methodological challenge of estimating aggregate treatment effects with considerable treatment spillover effects. In this paper, treatment intensity differs across cities even though the objective of the Chinese government was to build an integrated national highway network system. I exploit such variation across cities to identify the aggregate effects of highway infrastructure.

I use firm-level data to examine how a city's production efficiency responds to an expansion of highway network. I use the standard procedures proposed in the industrial organization literature (Olley and Pakes 1996; Levinsohn and Petrin 2003; Ackerberg et al. 2015 and Wooldridge 2009) to estimate firm TFP. Although there is a large and growing literature in international trade on firm adjustments after trade liberalization, there are very few papers that have looked at firm-level adjustments and its aggregate implications after a large-scale transportation infrastructure project. I find that average firm TFP increased as a city gained market access, which is consistent with models such as Melitz (2003) and Chaney (2008). I do not find strong evidence for within-firm productivity growth.

To evaluate the aggregate economic impact of China's national highway system, I choose the scenario of removing all highways as the baseline counterfactual.[9] I calculate the counterfactual trade costs between all city-pairs and cities' counterfactual market access in absence of highways. Based on the estimated impact of market access on productivity and the calculated decline in cities' market access in the counterfactual scenario, I estimate that eliminating all highways in China would decrease aggregate productivity by 3.2%. If we allow highways to influence the distribution of population across cities, the estimated impact of highways on productivity is 3.8%. The counterfactual analysis suggests a sizable effect of highway infrastructure on aggregate TFP.

---

[8]Before 2000, long-distance freight was shipped primarily by rail, whereas short-distance freight was carried on local roads. However, the railway system changed very little over the period that I study. Thus, I do not expect the rail network to be a strong driver for growth after 2000. Banerjee et al. (2012) examine the effects of railways on local economic outcomes and they find railway had weak positive effects on the level of GDP, but not on GDP growth. In my empirical specification that I will discuss in detail below, I include the railway network as a control variable.

[9]Donaldson and Hornbeck (2016) develop a general methodology for evaluating the aggregate economic impact of transport network projects.

I then conduct a decomposition exercise to examine the channels through which the highway infrastructure promoted productivity growth in China. I follow Haltiwanger (1997) and decompose changes in aggregate TFP at the industry and city level into four components: within-firm productivity growth of continuing firms, the reallocation of market shares from less-productive continuing firms to more-productive continuing firms, the entry of productive firms and the exit of inefficient firms. The decomposition exercise suggests that the entry of new and productive firms contributed most to TFP gains. The reallocation among large incumbents and exit of inefficient firms also contributed to TFP growth.[10] I find similar effects of access to the international market through domestic transportation cost reductions.[11] These findings are broadly in line with the findings in the international trade literature (Brandt et al. 2016).

To evaluate the effects of a transport infrastructure project, we need to deal with the threat to identification posed by endogenous highway assignment. I employ an instrumental variable approach to address the potential endogeneity of both the *location* and *timing* of highway assignment. To deal with I construct a *time-varying* least-cost path spanning tree network to instrument for the expansion of China's highway network over the period 1995–2005. I borrow from engineering and network theory to construct the hypothetical network. I then construct the market access measure with the least-cost path network and use it as an instrument for the actual city market access. The results from two-stage least squares (2SLS) regressions confirm my findings from the baseline OLS regressions. I perform various robustness checks to test the validity of my results.

I also examine the heterogeneous effects of the national highway system on the specialization pattern of industries. I examine different dimensions of industry characteristics that have been emphasized by the classical trade theories such as Heckscher–Ohlin model and the New Economic Geography models. In particular, I examine the differential effects of market access on industries with different transportation costs, capital intensities, and product differentiation. I find that industries with larger transportation costs and higher capital in-

---

[10]The exit of inefficient firms made the smallest contribution to TFP gains, which suggests that exit decisions may be subject to other economic and political factors.

[11]Similar to the results from the firm-level regressions, I do not find strong evidence for within-firm productivity gains from an increase in domestic market access. I do find that a reduction in transportation costs to ports promoted within-firm TFP growth.

tensity tend to concentrate in locations with better highway access. These findings improve our understanding of are consistent with New Economic Geography models and complement empirical studies in urban economics that investigate the geographic concentration of industries (Frye 2014).

This paper is related to several strands of literature in urban and spatial economics. First, this research is related to a recent and rapidly-growing empirical literature on the economic impacts of transport infrastructure. Donaldson (2016) is a seminal paper that investigates the effects of India's railroad construction in the colonial era on trade, prices and income. Ghani et al. (2016), Alder (2015) and Asturias et al. (2015) focus on the effects of the Golden Quadrilateral highway project in India. For China, Faber (2014)[12] and Baum-Snow et al. (2015) look at the effects of China's highway construction on local GDP and population growth. Banerjee et al. (2012) examines the effect of railway network on per capita GDP levels and growth in China. Second, this paper is also related to classic theories of international trade and new economic geography theory pioneered by Krugman (1991). Hanson (2005), Redding and Venables (2004), Hanson and X. (2004) and Donaldson and Hornbeck (2016) are influential empirical studies that connect the models to the data tightly.

My paper also expands the literature on trade barriers and firm-level adjustments. Recent literature has been increasingly focusing on how firms respond to trade liberalization. This trend originates from theoretical work by Melitz (2003) and Bernard et al. (2003), which introduce a new margin of gain from trade–reallocation of resources from the less productive firms to the more productive firms. Recent papers in urban economics emphasize the sorting of firms across space (Behrens et al. 2014; Gaubert 2014. Lileeva and Trefler 2010 and Bustos 2011 examine the effects of trade liberalization on firm productivity and exporting behavior. It is little known, however, how domestic market integration through expansions of transportation infrastructure affects production efficiency of firms. This study departs from the literature and aims to shed light on how productivity responds to a reduction in trade costs due to highway connection.

The remainder of the article is organized as follows. Section 2 briefly discusses the

---

[12]Contrary to Faber (2014), in this paper I treat a city center along with its periphery as one metropolitan area. I do not investigate how the spatial organization of production within each metropolitan area.

background of the construction of China's national highway system. Section 3 presents a multi-city trade model with heterogeneous firms to illustrate how transportation infrastructure affects firms and industries. Section 4 describes the data I use and how I construct transportation costs and market access at the city level and industry level from the data. Section 5 discusses the instrumental variable and the identification strategy I use. Section 6 presents and discusses my findings on aggregate TFP. Section 7 discusses findings on the spatial reallocation of industries. The last section concludes.

2   Background

China has shown greater interest in investing in its domestic infrastructure than have many other large developing countries over the past two decades. Some argue that large investments in transport infrastructure contributed to China's "growth miracle". China experienced a period of rapid highway expansion in the late 90s to mid-2000s, from virtually no highway to an integrated national highway network. In 1992, the Chinese State Council approved the construction of the "7-5" network. The objectives stated by the government were to connect all provincial capitals and cities with an urban population above 500,000 through a National Trunk Highway System (NTHS) by 2020. The plan outlined the construction of 12 trunk highway roads, including five longitudinal roads and seven latitudinal roads. Most of the projects, however, were completed during a 6-year period from 1998 to 2003. For example, Faber (2014) estimates that 81% of NTHS opened to traffic between mid-1997 and end of 2003. The completion of that project marked the beginning of an integrated domestic market supported by a national highway system.[13] Figure 3 shows the original plan for the NTHS.[14]

The drastic expansion of Chinese cities' access to the international market, especially in coastal areas, is equally pronounced. This expansion stirred the movement of goods, services

---

[13]The majority (76%) of the NTHS is expressways, the rest are highways.

[14]there were many local constructions of highways which did not belong the the National Trunk Highway System. When we construct the measure of market access, I use both the national trunk roads and local highways. When I construct my instrumental variable, I only use the trunk roads, as the construction of those trunk roads were much less subject to local economic conditions.

and people across cities. As shown in Figure 4, the total value of inter-provincial trade as a percentage of GDP increased from 82% to 108% for the period from 2002 to 2007. The rapid expansion of the highway network is also evident in the change in average travel distance for different transport modes. Road transportation experienced the largest increase in average travel distance among all transport modes. In addition to the reduction in transportation costs from highway expansion, China also experienced a massive trade liberalization during the same period of time. The country's WTO accession at the end of 2001 represented a large milestones for China's effort to liberalize trade and domestic market. The importance of international trade increased thereafter until the outbreak of the global financial crisis. Although WTO accession and associated policy changes were largely a national experiment, the impact of the trade reform differs across cities. This national shock allows us to use the city-level data to assess the effects of international market access on industrial clustering and production efficiency from a reduction in domestic transportation costs[15].

3 A Multi-City Trade Model with Heterogeneous Firms

3.1 Preferences

Consumer preferences in city k are defined over the consumption of goods produced in sectors $s \in 0, 1, ...S$:

$$U_k = Q_0^{\beta_0} \prod_{s=0}^{S} Q_{i,s}^{[\sigma_s/(\sigma_s-1)]\beta_s}, \quad \sum_{s=0}^{S} \beta_s = 1, \beta_s \geq 0. \tag{1}$$

$$Q_{i,s} = \left[\int_{\omega \in \Omega_s} q_s(\omega)^{(\sigma_s-1)/\sigma_s} d\omega \right]^{\sigma_s/(\sigma_s-1)}, \quad \sigma_s > 1, s \geq 1. \tag{2}$$

A homogeneous good 0 is produced in every city, is freely traded and is used as a numeraire. It is produced under constant returns to scale and one unit of labor in city k can

---

[15]In the regressions below, I control for industry-year fixed effects to account for effects stemming from tariff and other industry-specific policy changes

produce w_k units of good 0. Its price is set to 1 so the wage of city k is w_k.[16] In this case, the wage in each city is exogenously determined by the city's efficiency in producing the numeraire good.

3.2   Production

Each firm uses labor to produce a variety. I assume firms pay a fixed cost of f_{kj}^s units of labor and an iceberg variable trade cost τ_{kj}^s to serve market n, and $\tau_{kk}^s = 0$ without loss of generality. Each firm uses labor to produce a variety. Costs are specified in terms of origin's labor. For a firm with productivity φ, the total amount of labor required to produce q units of a variety and sell them to market n is:

$$c_k^s = f_{kj}^s + \frac{w_k^s \tau_{kj}^s q^s}{\varphi}. \tag{3}$$

A representative consumer in city k solves a constrained maximization problem, and the relative demand between any two varieties can be expressed as follows:[17]

$$\frac{q^s(\omega_1)}{q^s(\omega_2)} = \left(\frac{p^s(\omega_1)}{p^s(\omega_2)}\right)^{-\sigma_s}. \tag{4}$$

Multiply both sides by $p^s(\omega_1)$ and rearrange:

$$\beta_s Y_k = \int_0^N p^s(\omega_1)q(\omega_1)d\omega_1 = p(\omega_2)^\sigma q(\omega_2) \int_0^N p(\omega_1)^{1-\sigma_s}d\omega_1 \tag{5}$$

If we define a price index of sector s in city j to be $P_j^s = \left(\int_0^N p(\omega_1)^{1-\sigma_s}d\omega_1\right)^{\frac{1}{1-\sigma_s}}$, then city j's Marshallian demand for any variety ω produced in city k is

$$q_{kj}^s(\omega) = \beta_s Y_j \left(P_j^s\right)^{\sigma_s-1} p_{kj}^s(\omega)^{-\sigma_s}, \tag{6}$$

Solving a firm's profit maximization problem, we arrive at the standard expression for

[16]I assume labor is perfectly mobile across sectors so w_k^s, the wage in sector s in city k, should be equal to w_k for all sectors in city k.

[17]Note that we are using CES utility function, and from (5) it is immediate that $\frac{d(q(\omega_1)/q(\omega_2))/(q(\omega_1)/q(\omega_2))}{d(U_{\omega_1}/U_{\omega_2})/(U_{\omega_1}/U_{\omega_2})} =$ $\frac{d(q(\omega_1)/q(\omega_2))/(q(\omega_1)/q(\omega_2))}{d(p(\omega_1)/p(\omega_2))/(p(\omega_1)/p(\omega_2))} = \frac{-dln(q(\omega_1)/q(\omega_2))}{-dln(p(\omega_1)/p(\omega_2))} = \sigma_s$

the firm's optimal price as its marginal cost multiplied by a constant markup:

$$p_{kj}^s(\varphi) = \frac{\sigma_s}{\sigma_s - 1} \frac{w_k^s \tau_{kj}^s}{\varphi}. \tag{7}$$

Equilibrium revenue of the firm is given by

$$r_{kj}^s(\varphi) = \beta_s Y_k P_s^{\sigma_s - 1} p_{kj}^s(\varphi)^{1 - \sigma_s}, \tag{8}$$

Profit is given by

$$\pi_{kj}^s(\varphi) = \frac{r_{kj}^s(\varphi)}{\sigma_s} - w_k^s f_{kj}^s = \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} w^{1 - \sigma_s} \beta_s Y_j P_s^{\sigma_s - 1} \varphi^{\sigma_s - 1} - w_k^s f_{kj}^s, \tag{9}$$

Productivity Cutoff.—After observing its productivity, a firm decides on whether to exit or to sell in that market. We can pin down the export cutoff productivity $\overline{\varphi}$ by the zero-profit condition:

$$\overline{\varphi}_{kj}^s = \lambda_s \left(\frac{f_{kj}^s}{\beta_s Y_j} \right)^{1/(\sigma_s - 1)} \frac{w_k \tau_{kj}^s}{P_j^s}, \tag{10}$$

where λ_s is a constant[18] and f_{kj}^s is fixed cost[19]. A special case is a firm selling to its domestic market. In that case, $\tau_{kk} = 1$ and the firm will exit if its productivity is lower than the survival productivity.

$$\overline{\varphi}_k^s = \lambda_s \left(\frac{f^s}{\beta_s Y_k} \right)^{1/(\sigma_s - 1)} \frac{w_k}{P_k^s}. \tag{11}$$

3.3 Quantitative Predictions

In order to generate predictions from the model, I make a few simplifying assumptions. I first assume firm productivity is drawn from a common Pareto distribution in each city.

$$g_s(\varphi) = \gamma_s \varphi_{min}^{\gamma_s} \varphi^{-(\gamma_s + 1)}, \quad G_s(\varphi) = 1 - \left(\frac{\varphi_{min}}{\varphi} \right)^{\gamma_s} \tag{12}$$

---

[18] $\lambda_s = (\sigma_s/\beta_s)^{1/(\sigma_s - 1)}(\sigma_s/(\sigma_s - 1))$

[19] For simplicity, below I assume fixed cost of exporting is independent of origins and destinations, or $f_{kj}^s = f^s$.

where $\varphi_{min} > 0$ is the lower bound of the support of the productivity distribution and γ_s is the shape parameter (lower values of the shape parameter correspond to greater dispersion in productivity). Without loss of generality, I assume productivity is distributed over $[1, +\infty]$.[20]

We also assume that the total mass of potential entrants in city j in each differentiated sector is proportional to $w_j L_j$, similar to Eaton and Kortum (2002) and same as Chaney (2008). We abandon the free entry condition in the Melitz model to simplify the analysis for now, although imposing the free entry condition will not change predictions from the model. Since we do not impose free entry, firms produce positive profits that need to be redistributed to the workers. Following Chaney (2008), I assume a global fund collects profits from all firms and redistribute them to workers in units of the numeraire good. Each worker owns w_i shares of the fund.

The price index P_j^s takes the following form:

$$P_j^s = \left(\sum_{k=1}^{N} w_k L_k \int_{\overline{\varphi}_{kj}^s}^{\infty} \left(\frac{\sigma_s}{\sigma_s - 1} \frac{w_k \tau_{kj}^s}{\varphi} \right)^{1-\sigma_s} dG_s(\varphi) \right)^{1/(1-\sigma_s)} \tag{13}$$

dividends per share, d, is defined as

$$d = \frac{\sum_{s=1}^{S} \sum_{k=1}^{N} w_k L_k \left(\int_{\overline{\varphi}_{kj}^s}^{\infty} \pi_{kl}^s(\varphi) dG_s(\varphi) \right)}{\sum_{n=1}^{N} w_n L_n}, \tag{14}$$

$$Y_j = (1 + d) \times w_j L_j \tag{15}$$

To unburden us from the heavy notation, I ignore the industry subscript s from now on. The other industries are analogous. Given the productivity cutoff $\overline{\varphi}$,

$$P_j = \kappa_1 Y_j^{1-\sigma} \Theta_j; \tag{16}$$

where $\Theta_j^{-\gamma} = \sum_{k=1}^{K} (Y_k/Y) \times (w_k \tau_{kj})^{-\gamma} \times f^{-[\gamma/(\sigma-1)-1]}$. This term is very similar to the "multilateral resistance variable" coined by Anderson and Van Wincoop (2003).

---

[20]If we assume the Pareto distribution is bounded from above, we will need to assume that $\gamma_s > \sigma_s - 1$ for firm size to be finite.

If a firm's productivity is larger than the cutoff $\overline{\varphi}_{kj}$, then export value is given by

$$x_{kj} = \kappa_2 \times \left(\frac{Y_j}{Y}\right)^{(\sigma-1)/} \times \left(\frac{\Theta_j}{w_k \tau_{kj}}\right)^{\sigma-1} \times \varphi^{\sigma-1}. \tag{17}$$

plugging this into (10), we have

$$\overline{\varphi}_{kj} = \kappa_3 \times \left(\frac{Y}{Y_j}\right)^{1/\gamma} \times \left(\frac{w_k \tau_{kj}}{\Theta_j}\right) \times f^{1/(\sigma-1)} \tag{18}$$

Productivity cutoff in city k is given by

$$\overline{\varphi}_k = \kappa_4 \times \frac{w_k}{Y_k^{1/\gamma}} \times \left(\sum_{j=1}^{N} \frac{w_j^{-\gamma} Y_j}{\tau_{kj}^{\gamma}}\right)^{1/\gamma} \tag{19}$$

The last term on the right resembles the market potential term coined by Harris (1954) and market access term in Redding and Venables (2004) and Donaldson and Hornbeck (2016). More specifically, a city that can reach big cities (large Y_k) in a cheap way (small τ_{kj}) tends to have higher productivity cutoff, meaning that the average firm productivity in the city tends to be higher. Later we introduce the data and methods to estimate this market access term.[21]

The average firm productivity in a city is defined as the following

$$\overline{TFP}_k = \int_{\overline{\varphi}_k}^{\infty} \varphi \, dG_k(\varphi) = \frac{\gamma}{\gamma-1} \overline{\varphi}_k = \frac{\gamma}{\gamma-1} \kappa_4 \times \frac{w_k}{Y_k^{1/\gamma}} \times \left(\sum_{j=1}^{N} \frac{w_j^{-\gamma} Y_j}{\tau_{kj}^{\gamma}}\right)^{1/\gamma} \tag{20}$$

The aggregate productivity in a city is defined as a revenue-weighted TFP:

$$\widehat{TFP}_k = \int_{\overline{\varphi}_k}^{\infty} \varphi s(\varphi) \, dG_k(\varphi) = \frac{\gamma_s}{\gamma-\sigma} \overline{\varphi}_k^{\sigma} = \frac{\gamma}{\gamma-\sigma} \kappa_4^{\sigma} \times \left(\frac{w_k}{Y_k^{1/\gamma}}\right)^{\sigma} \times \left(\sum_{j=1}^{N} \frac{w_j^{-\gamma} Y_j}{\tau_{kj}^{\gamma}}\right)^{\sigma/\gamma} \tag{21}$$

where κ_1, κ_2, κ_3, κ_4 are all constants, and $s(\varphi)$ is the share of a firm's revenue in total industry revenue in city k.

---

[21]In the Appendix, we introduce a closely related concept of market access that is adopted in Redding and Venables (2004) and Donaldson and Hornbeck (2016). The results remain largely unchanged with their definitions.

As mentioned above, the term $\sum_{j=1}^{N} \frac{w_j^{-\gamma} Y_j}{\tau_{kj}^{\gamma}}$ is closely related to the "market access" term that I will construct in Section 4. The derivations from the model show that firms located in cities with larger market access tend to have higher productivity on average, and cities with larger market access tend to have higher aggregate productivity. If a transport infrastructure project reduces the transportation costs between a city and its neighbors hence increases the market access of a city *relative to* other cities, we should expect to see a *relative* increase in the city's production efficiency.[22] In the model, the efficiency gain is a combination of the entry of new and productive firms, the exit of inefficient firms and the reallocation across firms. In Section 6, I will examine the relative importance of each channel.

4   Data and Measurement

There are mainly four data sources I use for this project. First, I use digitized maps of China's highway network system for the period from 1992 to 2015 in combination with local non-highway, railway and waterway maps.[23] The vector data representing China's expressways, highways and local paved roads was built up over many years from a wide variety of published road atlases. That compilation of information on the current expressway routes has then been coded to indicate the type and status of expressways that were depicted in road atlases. The early road atlases from the last century have maps that are quite small scale, whereas the better atlases from more recent years have much more detail. As a result, information in older atlases were coded onto the routes shown in more recent ones to avoid any discrepancy over time. I use these digitized highway maps to construct least cost path between any two cities for 1995, 2001, and 2005. The construction of least cost paths and choice of years

---

[22]In the model, I assume labor is immobile. If labor is perfectly mobile, then we should not expect to see a positive relationship between city market access and production efficiency in the data because the relative increase in demand for labor would be completed offset by new immigrants. The reality in China was that labor were somewhat mobile but with strong restrictions. In that sense, the estimated effect of highways below should serve as a lower bound for the actual effect.

[23]The data mainly come from the Australian Consortium for the Asian Spatial Information and Analysis Network (ACASIAN). For data inquiries, please contact `lcrissman@optusnet.com.au`. I would like to thank Prof. Dai at the Geology Department of Beijing Normal University for helping me get the newest transport route map.

are explained below. Figure 5 present highways construction in China for the period of this study.

Second, firm-level data and industry-level data are constructed from the Annual Survey of Industrial Enterprises (ASIE) from 1998 to 2007. The survey conducted by the National Bureau of Statistics of China span the period from 1998 to 2009. The survey contains all State-Owned Enterprises (SOEs) and all private enterprises with annual sales of $650,000 and above. The data set contains very detailed information on firm's balance sheet and income statement, as well as information on ownership, export status, employment among others. For the year 2004, we have information on average education and skill level of labor. There are over 100 variables in the data. The most useful variables for this project are firms identifier, industry identifier, gross output, total sales, wage bill, employment, stock of fixed capital, value of intermediate inputs, export status, year of establishment, ownership, skill/education level and location identifier. I also use Economic Census for 1995, 2004, 2008 and Basic Units Census for 1996 and 2001 to construct aggregate output and employment for each industry at the prefecture city level.

Finally, to estimate bilateral transportation costs between any two cities, I collect prices quoted by logistics companies for freight transportation between any two cities in China. More specifically, I gather all price quotes from two of the largest logistics companies in China from their websites[24] to estimate transportation costs between any two cities. These two companies cover freight transportation over most of the cities in China. I will argue below that price quotes they publish are good estimates of the actual costs that firms or their clients have to pay to deliver the goods. By using actual transportation costs data, I contribute to the literature by estimating the bilateral transportation costs between Chinese cities.

---

[24]I use quotes from Arima World Group (website: www.hoau.com) and Deppon (website: www.deppon.com). Please see Appendix A.2. for details of the price information on their websites. They are two of the largest logistics companies in China. They both have a very extensive network in the whole nation.

4.1 Estimation of Bilateral Transportation Costs

To construct a measure of market access for each city, it is necessary to first gauge the cost of transporting goods between cities. A challenge in the literature is the lack of reliable estimates of transportation costs. Some of the earlier papers use the Euclidean distance between locations to proxy for transportation costs. There are two problems with this approach. First, Euclidean distance is a poor proxy for transportation costs. Euclidean distance is not able to take into account road availability and conditions, topography and many other factors that will affect transportation costs. Second, Euclidean distance is time-invariant and thus prevents researchers from using changes in transportation costs over time to evaluate the impact of a transport infrastructure project.

Some of the recent work in this literature uses digitized transport route data to estimate transportation costs. With digitized transport route data, the most widely used approach is to estimate travel distance or time first and then to estimate transportation costs between locations with parameter assumptions, such as dollars per ton-mile (Donaldson and Hornbeck 2016) or dollars per hour (Baum-Snow et al. 2015). This approach is arguably a substantial advance since transport infrastructure is a key determinant of trade costs between locations and has attracted considerable interest in the literature. However, there are still a number of pitfalls associated this approach. First, the assumed per unit cost parameters are usually taken from previous studies that investigate a different question in a different country for a different time period. There are many potential issues in using, for example, overland shipping costs in the US in the 1980s to investigate the current Chinese market, which could undermine the relevance of the estimated transportation costs. Second, this simplistic approach ignores many other factors that we are known to affect transportation costs, such as congestion, economies of scale and local labor costs. It is the acknowledgment of these pitfalls that lead me to adopt a different strategy for estimating transportation costs in this paper. One contribution of this paper is to estimate the transportation costs between any two cities in China with data on actual transportation costs.

As mentioned at the beginning of this section, I collect all price quotes from two of the largest logistics companies in China. The data indicate how much the seller or the buyer

has to pay to have the product shipped from the factory to the user. The price quotes incorporate all information, observable or unobservable, such as travel distance, road availability, congestion, costs of labor, gasoline and other inputs. Thus, in contrast to previous papers, I measure the actual transportation costs between any two cities in China. A potential drawback of using the price quotes is that companies may not strictly follow the quotes when they finalize the contracts with their clients. I checked with representatives from these two companies and with some of their clients to determine whether the price quotes they publish on their websites truly reflect the actual transaction prices. To the best of my knowledge, the two companies very closely follow the price quotes in practice. An established client will be able to secure a 10%-20% discount, while the logistics companies occasionally award discounts as high as 50% to very large orders. Although they do exercise discretion on an order-by-order basis, the price quotes are reasonably good estimates of the actual shipping costs that a firm needs to pay if it chooses to use one of the two logistics companies. There are many other smaller logistics companies in China, and many of the smaller companies only maintain a local presence. It is impossible to collect price quotes from all logistics companies in China, but to the extent that there are many companies and the market is competitive, I will assume that the price quotes from these two largest companies reflect the "market price".[25] Since I need to estimate iceberg trade costs, I also need to obtain value-to-weight ratio of manufacturing goods. I use revenue and quantity data of Chinese manufacturing firms to estimate the average value-to-weight ratio for each manufacturing industry.

Table 1 shows some of the estimated shipping costs between cities. The first lesson we learn from the table is that shipping costs vary widely within China. We see that for some cities that are close to each other and/or have easy highway access, shipping costs between them are only a very small fraction of total value of manufacturing goods. However, for many cities that are far away from each other or do not have highway access, then shipping costs can be as high as 30% of total goods value. An example would be that the estimated shipping costs between the two largest cities in China, Beijing and Shanghai, is 5.5% of total goods value. Second, these estimated costs are very different from the parameters assumed

---

[25]For any city pair, I use the lower price of those offered by the two companies as the actual shipping costs that a firm has to pay to transport goods between the two cities, conditional on firms using the same class of delivery service

in the literature. For example, Limao and Venables (2001) find that the cost of shipping 1 ton of freight overland for 1000 miles is about \$2,100, or about 2% of its value. Some of the papers that use these estimates will end up estimating a much smaller transportation cost[26].

While I can collect the current price quotes from the two companies' websites, they do not publish any information on historical prices. To estimate transportation costs for previous years, I employ a simple linear regression model incorporating both macro data and origin- and destination-specific data. To obtain transportation costs, I run a simple regression model to estimate transportation costs for previous years.

The first step is to estimate the following regression specification:

$$\tau_{o,d,15} = \alpha + \beta_1 hwy_{o,d,15} + \beta_2 local_{o,d,15} + fe_o + fe_d + \epsilon_{o,d,15}$$

where $\tau_{o,d,15}$ is the iceberg transportation cost between city o and d in 2015, expressed as a percentage of goods value. $hwy_{o,d,15}$ is the length of highway on the optimal route, whereas $local_{o,d,15}$ is the length of local roads on the optimal route in 2015. By estimating this regression specification, I obtain $\hat{\alpha}$, $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{fe}_o$, $\hat{fe}_d$, which are estimates for α, β_1, β_2, fe_o, fe_d,respectively. The second step is to estimate transportation costs between origin and destination.

$$\hat{\tau}_{o,d,t} = \hat{\alpha} + \hat{\beta}_1 hwy_{o,d,t} + \hat{\beta}_2 local_{o,d,t} + \hat{fe}_o + \hat{fe}_d$$

One way to assess the quality of estimated historical transportation costs is to examine the goodness-of-fit of the regression. In Column (2) of Table 5, I find that the model can explain 85% of the variation in bilateral transportation costs. If I exclude all of the origin and destination fixed effects and only use the road data, the model with road data alone explains 69% of the variation in transportation costs (see Column (1) of Table 5). I also add nonlinear terms of highway and non-highway length, as well as total travel time and its nonlinear terms to control for factors that are not collinear with highway and non-highway distance. In Columns (3) and (4) of Table 5, I find that the R-squared does increase, but only

---

[26]For example, an earlier version of Baum-Snow et al. (2015) estimates that travel along the 1990 road network from Beijing to Shanghai would cost 2% of value

marginally. These results suggest that road data can effectively predict transportation costs, and city-specific factors such as labor costs and congestion also matter for transportation costs. The results offer some confidence that this simple linear regression model with road data can estimate historical transportation costs reasonably well. Please also note that it is the *relative*, not the *absolute* transportation costs that matter for my purpose. Any common factors such as oil price and wage growth that affected the entire country symmetrically will not invalidate my analysis. One may be temped to estimate the above regression at the industry level if one believes that the cost structure is heterogeneous across industries. However, the price quotes from the logistics companies do not distinguish between product groups. Thus, there is no need to separately estimate transportation costs for each industry.

Table 1: Transportation cost (% of Good Value)

Panel A: Low Iceberg Transportation cost		
origin	destination	Iceberg Cost
Yangquan	Taiyuan	1.04%
Liangshan	Ya'an	1.30%
Chengdu	Zigong	1.30%
Kaifeng	Zhengzhou	1.51%
Panel B: High Iceberg Transportation cost		
origin	destination	Iceberg Cost
Huainan	Lasa	20.32%
Hetian	Wuzhou	20.97%
Lasa	Kelamayi	23.70%
Shannan	Bayingguole	28.11%
Panel C: Between Beijing and Shanghai		
Beijing	Shanghai	5.54%

Notes: This table shows iceberg trade costs between city pairs. Panel A shows the top-5 cities pairs that have the lowest transportation costs. Panel B shows the bottom-5 city pairs that have the highest transportation costs. Panel C shows the transportation costs between Beijing and Shanghai, the two largest cities in China.

4.2 Construction of Market Access

The idea of "market access" dates back to Harris (1954). Harris argued that the potential demand for goods produced in a location depends on the sum of distance-weighted GDP from all locations. Mathematically, Harris's "market potential" term equals $\sum_d (d_{od})^{-1} N_d$, where d_{od} is the distance and N_d is location's population or GDP. Redding and Venables (2004) and Donaldson and Hornbeck (2016) derive a similar term–"market access"–that measures each location's proximity to markets. To construct market access for all 340 prefectures in China, I assemble a prefecture-level dataset of employment and the number of firms in all manufacturing industries. To assess the effects of highway access on productivity, I use data for all manufacturing firms in China. Over the period that I study, the administrative boundaries of many prefectures changed because counties were occasionally reassigned to a different prefecture. I establish a county-level correspondence from 1995 to 2008 and construct time-consistent prefecture boundaries.

After the massive and rapid expansion of highway network during the period 1998-2005, China established the National Trunk Highway System (NTHS). Thanks to this massive infrastructure project, highway transportation gradually became the main transportation method for inter-city trade after 2000, whereas railroads were little changed and have been used primarily for transporting commodities such as coal and metal since then. In fact, estimates from different sources suggest that road transportation accounted for nearly 70% of freight value and logistics costs in 2007. Since I only consider manufacturing industries from 1998 to 2007, it makes sense for me to focus on the highway "shock" to Chinese cities. Therefore, I use digitized highway network data and local road data to construct each city's market access after China's highway network expansion. As mentioned above, I control for access to railways and waterways in my empirical specification.

I use firm census data for the years 1995, 1996, 2001 and 2008 to construct employment, the number of firms and sales for 42 2-digit manufacturing industries at the prefecture level. The reason for only using manufacturing firms is that it is generally accepted that manufactured goods are much more tradable than services. For example, a KPMG report estimates that the movement of industrial products accounted for 87% of the value of goods

moved in 2006. I use the Annual Survey for Industrial Enterprises (ASIE) from 1998 to 2007 to estimate firm-level productivity. I use the standard Olley-Pakes estimation method to estimate firm TFP.[27] I also use firm-level census data to construct output and employee data at the industry-prefecture level from 1995 to 2008. The reason that I use the average growth rates from the period 2001-2008 is twofold. First, China did not have a national highway network until the early 2000s. Thus, by using the period from 2001 to 2008, I ensure that my measures of market access using highway data truly capture proximity to markets and supplies. Second, the Chinese government announced a 4-trillion-yuan stimulus package after the global financial crisis began in 2008. In the following years, they committed considerable resources to building roads as part of the stimulus package in an effort to boost domestic demand. Thus, the period after 2008 might be affected by the crisis and the stimulus package, which might need additional attention.

As pointed out by Baum-Snow et al. (2015), there were very few expressways before 2000 and almost all long-distance shipping was through railways or waterways. With the rapid and massive expansion of national highway network from 1998-2003, road transportation emerged as the prominent transportation method except for a few commodities. In fact, coal alone consists of roughly 50% of inter-provincial trade by railway in 2007. Coal and a few other commodities such as iron ore, lumber and other metals account for over 90% of freight transportation by railway [28]. Moreover, China's railway network has been little changed since 1990. In this paper, I only consider manufacturing industries, which should further alleviate the concern that constructing market access solely based on roads data may distort results. In addition, I include access to railways and waterways in the regressions to address any potential bias. The effects of the highway network expansion on regional trade can be seen in Figure 4 inter-provincial trade exploded since early 2000, reaching 106% of GDP from 85%. Even though there is no data on inter-prefecture trade, it is very likely that the increase in trade was even more drastic. Since then, China continued to expand the reach of its highway network. The average transport distance of highways nearly tripled from 60 kilometers in 2000 to 172 kilometers in 2008, whereas the average transport

---

[27]There is now a large literature on the estimation of firm productivity. Other popular methods include Levinsohn and Petrin (2003), Ackerberg et al. (2015) and Wooldridge (2009).

[28]Data used for calculating these statistics are from the National Bureau of Statistics

distance of railways and waterways did not change at all and stayed at 760 kilometers and 1800 kilometers respectively during the same period. Figure 5 shows how China's highway network evolved over the years.

The market access term derived in Section 3 resembles the "market potential" approach of Harris (1954) and the "market access" approach developed by Redding and Venables (2004) and Donaldson and Hornbeck (2016). The Harris approach is simple, whereas the Donaldson and Hornbeck approach requires numerically solving 318 nonlinear equations simultaneously. Donaldson and Hornbeck (2016) verify that the numerically solved market access term is highly correlated with the simpler market potential term, and the results do not depend on what term they use. In the model, I show that productivity is determined by a "market potential" term, whereas industry specialization patterns are more closely related to the "market access" term. Henceforth, I use the term "market access" to refer to a location's proximity to suppliers and consumers, and use the market potential term as my baseline. I also numerically solve for the market access term and verify that these two approaches generate similar results. I use the following equation to estimate each location's market access:

$$MA_o \approx \sum_d \tau_{od}^{-\theta} Y_d,$$

where $\tau_{od,t}$ represents the iceberg trade costs and Y_d is the size of the economy of destination d. Redding and Turner (2015) note that most shipments cover over very short distances and that the time cost of freight seems to be important. This is why I choose the least-time-cost paths as the "optimal" route for ground transportation. The choice of routes is also confirmed by representatives from logistics companies in China.

What I have constructed is *domestic* market access. Cities also trade with the rest of the world, and this is especially true for the cities in coastal areas. Two forces increased Chinese cities' access to the international market during the period of study. First, tariffs declined rapidly following China's accession to the WTO at the end of 2001. Second, highway connections reduced the domestic transportation costs from the origin cities to the ports. Since I am interested in transportation infrastructure in this paper, I control for industry-

year fixed effects in the empirical specification to abstract from tariff effects.[29] I use the transportation cost from the origin city to the nearest port[30] to measure the city's access to the international market.

A concern associated with using only highway data is that railways also played an important role in freight transportation in China, especially before the highway era. Even at present, a very large share of coal, coke and metals are transported via rail, not highways. While short-distance travel is dominated by highway transportation, rail remains important for long-distance travel. However, rail only constitutes a small share of transportation except for those commodities mentioned above. A first argument against this concern is that my empirical strategy identifies the effect of market access over the time dimension. The railway system for freight transportation only changed very slightly over the period of study (Baum-Snow et al. 2015). Therefore, any time-invariant effect of railways on productivity and industry specialization should be absorbed by the city fixed effects. However, one may still be worried any time-varying effect of the almost static railway system. To address this issue, I multiply cities' access to railways by a year dummy and include this interaction term in my empirical specifications to allow for potentially time-varying effects of railways.

One may also be concerned about industry heterogeneity in transportation costs. Indeed, the iceberg transportation costs depend on the weight-to-value ratio of the industry and can differ markedly across different industries. I use the firm production quantity data set along with ASIE to estimate the weight-to-value ratio for each industry. Then, I construct city- and industry-specific market access. I verify that the estimated effects using industry- and city-specific market access are very similar to the city-level market access measure I construct. Figure 1 shows the change in market access for all Chinese cities from 1995 to 2005. Figure 6 presents the market access measure I construct for each city in 1995, 2001 and 2005.

---

[29]The magnitude of the tariff reduction was industry specific. A caveat is that the HS classification used for exports and imports differs from the Chinese Industry Classification.

[30]I calculate the transportation costs from every city to 9 of the largest ports in China. Then, I select the lowest cost from the 9 for each city.

Figure 1: Estimated Changes in Market Access
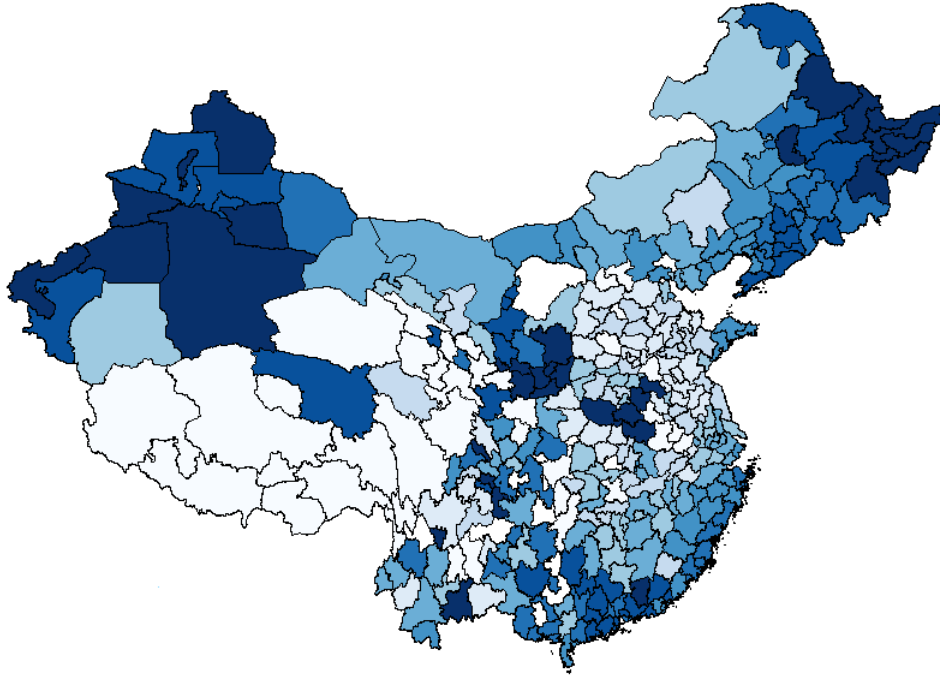
Notes: this figure shows the changes in market access from 1995 to 2005 for 339 Chinese cities. Darker color means larger increase in market access. We can see that areas in the west and the northeast gained large increases in market access. A few cities in the central area also experienced large increases in market access.

4.3   Firm Productivity and Markups

Similar to Feenstra et al. (2014), I use the augmented Olley and Pakes (1996) approach to estimate and calculate the TFP.[31] There are two approaches to estimating firm TFP. One is the Olley-Pakes approach that uses value added, and the other is the Levinsohn and Petrins approach that uses total output. Feenstra et al. (2014) argues that the Olley-Pakes approach is more appropriate in the Chinese context because processing trade in China accounts for more than a half of the country's total trade since 1995. The prices of imported intermediate inputs are different from those of domestic intermediate inputs. Using the domestic deflator to deflate imported intermediate input would create another unnecessary source of estimation bias. A potential issue with the Olley-Pakes approach is that a large number of firms that have zero investment will be dropped from the estimation exercise. However, As shown in

---

[31]All nominal variables are deflated by input and output deflators. Deflators are taken from Brandt et al. (2016). I use the perpetual inventory method to construct real capital stock, similar to Brandt et al. (2016).

Brandt et al. (2012), in the Chinese data there is only negative real investment for 1% of continuing firms. Moreover, I do not observe investment decisions directly, but estimate investment from the capital stock series, which will smooth out most of the zero investment decisions. I follow De Loecker and Warzynski (2012) to estimate firm markups. Since I use the value-added approach to estimate TFP, I choose labor as the "flexibly adjustable" input.[32] I also back out prices, marginal costs and physical productivity for a subset of firms that report quantity data. Prices and marginal costs will be used in Section 6 to disentangle gains retained by firms from gains passed on to consumers. Table 6 presents the estimated industry-level TFP by aggregating estimated firm TFP.

5 Instrumental Variable Approach

One of the greatest challenges in the literature is the endogenous placement of transport infrastructure, which may bias the estimated effects of transportation infrastructure. The official documents from the Chinese government state that the objective of the National Trunk Highway System was to connect all cities with populations over 500,000. It is highly probable that when choosing a route to connect two large cities, policymakers and urban planners would choose to build highways near the cities that they expected to have higher economic growth for reasons unobserved by the researcher. Our OLS estimates would have an upward bias in the existence of such correlation. Another possibility is that the location and funding decisions for highways were the product of a bargaining process between the central and local governments. Cities that had better political connection could get preferential policies from the central government. It is possible that highway was just one of the placed-based preferential policies during my period of study that shaped the subsequent development of the preferred cities. If that was the case, then the effect of market access on TFP may also be biased upward.

---

[32] Although it is more appropriate to use the value-added approach to estimate firm TFP, it is somewhat difficult to justify the assumption that labor input is fully flexible in China. Therefore, I also employ the output approach to estimate firm TFP and use material inputs as the "adjustable" input. These results are consistent with my baseline results.

Researchers have proposed historical roads (Donaldson 2016; Baum-Snow et al. 2015), planned networks (Baum-Snow 2007; Michaels 2008; Duranton and Turner 2011; Duranton et al. 2014), and algorithm-generated networks (Faber 2014) as instruments for actual highway or railway networks. The validity of these instruments hinges on whether these networks only influence productivity growth and industry specialization patterns through their predictive power for the actual transportation network conditional on control variables. In other words, these instruments may fail the exclusion restriction if they are correlated with unobserved economic fundamental or policy variables which also affect productivity growth and spatial reallocation of industries.

To address the identification issues discussed above, I construct a time-varying least-cost path spanning-tree network as an instrument for actual highway connections to evaluate China's highway expansion.[33] Note that I need to construct a time-varying instrument since the goal is to evaluate the effects of the NTHS on productivity and industry specialization. Thus, I also instrument for the *timing* of highway placement, which could also be endogenous as planners might choose to build roads near their most preferred cities first. The construction of the time-varying instrument is executed in two steps. In the first step, I use the Kruskal minimum spanning tree algorithm to calculate the least-cost path connections between any two city centers.[34] I use remote sensing data on terrain ruggedness collected by satellite images to estimate the construction cost of each small piece of land in China. Please refer to Appendix A.4. for all the estimated construction costs. I then construct least cost path connections for any two cities in China. Given all the least cost path connections, a least-cost path spanning-tree network is constructed that connects all the node cities and minimizes construction costs at the same time.

In the second step, I use the Girvan-Newman algorithm from network theory to predict the optimal timing of the construction of each connection.[35] The Girvan-Newman Algorithm

---

[33]In this paper, I am interested in the effects of market access on TFP growth and industry reallocation. A city's market access is affected by the expansion of the highway network even far away from that city. Thus, I am not testing a simple highway connection effect, which is likely to be correlated with city TFP growth. Instead,

[34]Faber (2014) constructs a similar least-cost path network. However, my instrument varies over time and predicts the timing of highway construction.

[35]Frye (2014) uses the Girvan-Norman algorithm to construct an instrument for the US interstate highway system.

ranks each edge by counting the number of shortest paths that move along that edge. The top-ranked edges are the most "important" ones to the network and should be built first. After the sequence of construction has been solved by the algorithm, I determine highway construction on the least-cost path network for each year. To do so, I first calculate the *actual* length of highway construction in the entire country for a given year and take that as an exogenous variable–one can regard this as a pre-approved government budget devoted to highway construction. Given the constraint that the planner faces, he or she constructs the least-cost path network by performing constrained optimization for each year.

The final step is to estimate city market access with the algorithm-generated highway networks. The way to estimate the algorithm-based market access is the same as described in 4.2. The only difference is that now the least-cost path spanning-tree network is used for estimation instead of the actual highway network. We use it to instrument for the market access measure based on actual transport routes.

My time-varying instrument addresses the identification concerns on both the *location* and *timing* of highway connection. The location dimension of the instrument is constructed based on cost minimization. The timing dimension of the instrument is constructed by on a "centrality" measure from network theory. Figure 2 plots the hypothetical construction of the least-cost path network. The highway routes in red represent hypothetical construction before 1995, the routes in green represent hypothetical construction during the period 1995-2000, and the routes in blue represent hypothetical construction during the period 2000-2005. The highway routes in black are the actual National Trunk Highway Network. It is evident from the figure that the least-cost-path spanning tree network resembles the actual highway network, but the two also differ in both the location and timing of highway assignment.[36]

The identifying assumption is that the hypothetical highway network should affect city productivity and the spatial allocation of industries only through the actual highway network, conditional on time-invariant city characteristics, city population and pre-existing railway and waterway access. I will discuss two threats to my exclusion restriction. First, local construction costs may be correlated with the potential economic returns of highway connection.

---

[36]For the timing of actual highway assignment, please refer to Figure 5. Details of construction for each year is available upon request.

For instance, it is costly to build highways in a hilly region, and the economic benefits of highway connection for the region is very high because the region is not well connected by any other transport mode. The inclusion of pre-existing access to railway and waterway in the empirical specifications should mitigate this issue. I nevertheless construct an alternative Euclidean distance spanning tree network to replace the cost-based network. The Euclidean distance spanning tree network minimizes the total distance, not total construction costs of the network. None of the results are drastically different with this alternative instrument. We may also be concerned that placed-policies is correlated with the "centrality" of an edge, which is related to the "centrality" of its connecting cities. It is reassuring that I do not find evidence that city centrality is strongly correlated with pre-trends in the data.

6 Market Access and City Productivity

In this section, I first examine whether the large increase in Chinese cities' market access raised the production efficiency with firm-level data. In Section 6.2, I aggregate TFP growth at the industry level and decompose the aggregate TFP growth into four channels and quantify the importance of each channel. In Section 6.3, I then estimate the extent of the revenue-based TFP that are attributable to changes in markups and changes in physical productivity. Section 6.4 presents robustness checks to ensure the validity of my findings.

6.1 Firm-Level Regressions

First, I use the following baseline specification to examine the effects of market access on firm TFP:

$$logTFP_{i,k,t}^s = \alpha + \delta logMA_{k,t}^s + \left(X_{i,k,t}^s\right)' \beta + \lambda_s + \mu_{s,t} + \epsilon_{i,k,t}^s$$

where $TFP_{i,j,k,t}^s$ is measured productivity for firm i in industry s and city k in year t. $MA_{k,t}^s$ is the market access of city k in industry s in year t.

I am interested in δ, which measures the effect of market access on firm productivity. I include industry × year fixed effects, $\mu_{s,t}$, and city fixed effects, λ_k. I also include additional

Figure 2: Instrument: A Least-Cost Path-Spanning Tree Network
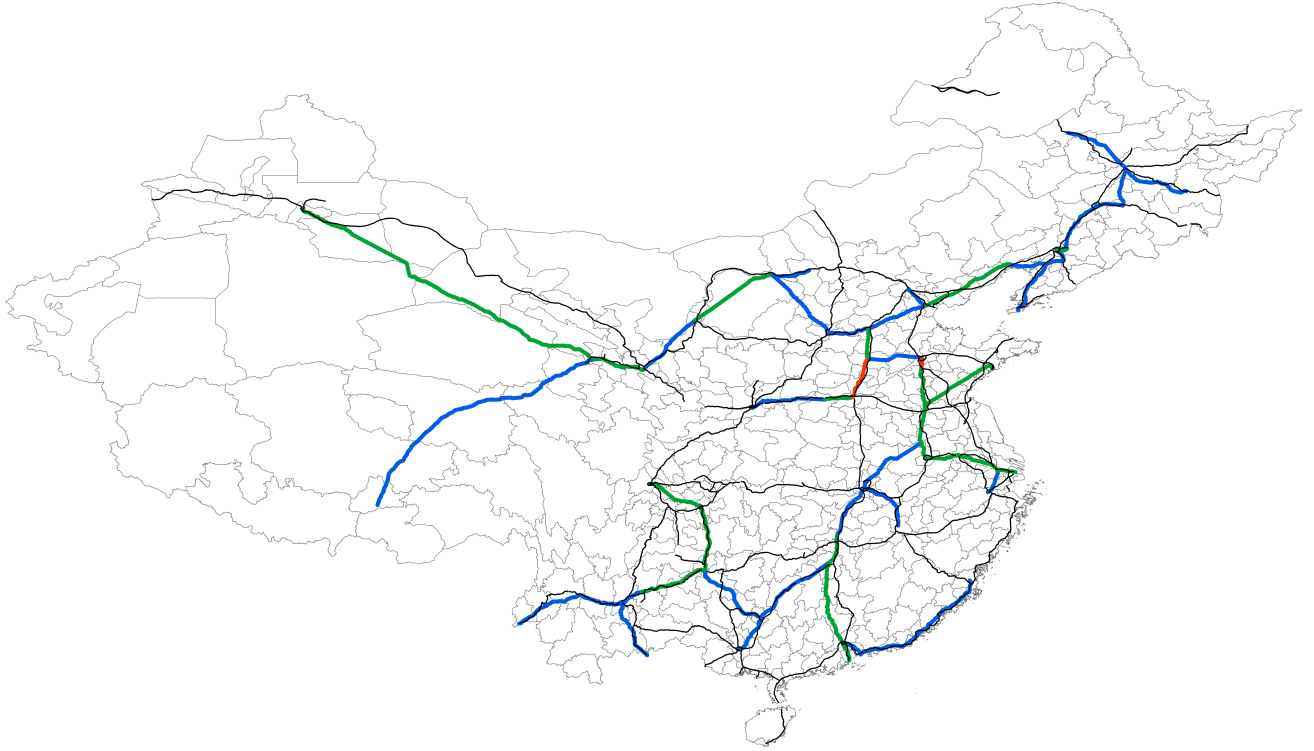
Notes: this figure shows the time-varying instrument. The black lines are the actual National Trunk Highway System (NTHS). The red lines are hypothetical highway construction in 1995; the green lines are hypothetical highway construction between 1996 and 2001; the blue lines are hypothetical highway construction between 2002 and 2005. The complete network is the least-cost path spanning-tree network.

control variables, $X_{i,k,t}^s$. The industry $\times$ year fixed effects account for any time-varying industry characteristics that may be correlated with the location of highway construction. As the tariff reductions were industry-specific after China's accession to the WTO in 2001, and the government's positive attitude towards active industrial policies, it is crucial to control for these potential confounding factors. By including city fixed effects, I control for any city characteristics that are not time-varying and that may be correlated with highway construction. Thus, the treatment effect of highway construction is only identified from variation within a city over time. I include city population, the interaction between distance to a railway and a year dummy, and the interaction of distance to a waterway and a year dummy as additional controls. The reason for including the interaction is because railways

and waterways were little changed over the period I study. Thus, I interact them with year dummies to account for the potentially time-varying effects of these variables. I also include firm-level control variables, i.e., firm markup, firm size, ownership structure, and an exporter status dummy.

I choose market access in three periods, 1995, 2001 and 2005, to perform my empirical analysis.[37] The reason for choosing these three years is twofold. The first reason is mainly data driven. It is not easy to identify what exactly was on the ground, even with the officially published highway network maps. 2001 and 2005 are the years in which there are more sources, and thus, I can cross check to ensure that the truly functioning highways are in my dataset. The second reason is that China had a "structural break" within each of the two time intervals. The project accelerated after the 1997 Asian Financial Crisis, as the central government decided to invest in the NTHS as part of its fiscal stimulus plan, whereas China joined the WTO at the end of 2001. To allow for the possibility that the effect of highway connection was not instantaneous, I use firm TFP data in 1998, 2003 and 2007 to allow for a possible lagged effect of highway connection.

OLS Regressions

Table 2 shows that firms in a city became more productive on average as the city's market access expanded. To check the stability of coefficients, I include a slew of controls in the regression. The estimated coefficient remains stable as I include aggregate and firm-level covariates.[38] More specifically, a 1% increase in market access increases firm productivity by 0.05%. Please note that market access increased due to either GDP growth or transportation cost reduction. Combining these two sources implies that there are large variations in market access across cities and cities in general experienced large changes in market access over the years. A one-standard-deviation increase in market access would boost firm TFP by 5%. For the purpose of this paper, we are more interested in the productivity gain due to reduction in

---

[37]For firm-level data, I use 1998,2003 and 2007 to account for any lagged effects of highway construction. It is plausible that the effect is not immediate after the completion of the construction project for several reasons. First, firms need to learn about the faster routes and build connections in other cities first. Second, it also takes time for logistics companies to update their routes and allocate labor and capital to the new routes.

[38]This is not because these control variables are not relevant. In fact, the R-squared increases as more covariates are added.

transportation cost only. After all, in a country like China growing at ten percent per year, the increase in market access is mainly driven by economic growth, and therefore, if growth is correlated across cities, this will mechanically induce a positive correlation between firm growth and market access. In one of the robustness checks below, I fix all cities' GDP to their 1995 level and only look at the effect coming from reductions in transportation costs. The results are similar to the ones in Table 2. In Column (2), I look at how firm TFP responds to changes in a city's access to international market. I find that a reduction in transportation costs from origin to the nearest port–an effective increase in access to the international market–raise average firm productivity. A 1% reduction in domestic transportation costs to ports boosted firm TFP in the city by 0.6 %. When I include both domestic market access and international market access, I find that both promoted firm TFP growth.

I then include firm fixed effects into the regression specification and test if an increase in market access raised TFP within firms. This channel has been emphasized in a few recent papers (Lileeva and Trefler 2010; Bustos 2011; Garcia and Voigtländer 2013). I do not find strong evidence for that channel. As we can see in Table 8, the coefficient on market access is insignificant once I include firm fixed effects. I do find that a reduction in transportation costs to the nearest port led to within-firm productivity growth. However, we need to construct measures of output market access and measures of input market to investigate whether the effect comes from easier access to cheaper and higher quality foreign inputs or from larger access to demand.[39]

Table 9 presents results on markups. I do not find strong evidence that firm markups respond to market access. Since higher productivity is another way of saying a firm has a cost advantage over other firms, the fact that markups do not respond to market access implies that the higher production efficiency is translated into lower product prices. This finding potentially has important welfare implications–both the more productive firms and consumers seem to reap the benefits from highway infrastructure. I will decompose the revenue-based TFP into markups, marginal costs and physical productivity.

---

[39]Currently I am using China's input-output tables to construct output market access and input market access at the city- and the industry-level. This exercise will allow us to disentangle the effects of proximity to suppliers from the effects of proximity to consumers.

As explained in Section 5, the instrument I construct is the hypothetical city market access constructed from a least-cost path network. I present results from the first stage and followed by the second stage. In the first stage, I regress the actual market access on the hypothetical market access, along with all the control variables. Results from the first stage of the two-stage-least-squares regressions are presented in Table 7. The hypothetically constructed market access is highly correlated with the market access estimated with the actual highway networks.[40] In Table 2, we see that the estimated effect is slightly smaller than the results from OLS regressions, which suggests that there were indeed unobserved forces that were correlated with cities' expansion of market access and affected TFP growth.

6.2 Counterfactual Impact of Removing Highways

Based on the estimated effect of market access on city productivity,[41] I evaluate the economic impact of China's national highway system. I consider a baseline counterfactual scenario of removing all highways in China.[42] Other counterfactual scenarios, such as replacing the national highway network with a more extensive railway system or a different highway network, can be analyzed in a similar fashion. I first calculate the decline in market access for each city if we remove all highways in China.[43] The median loss in market access by removing all highways is 62%. I then estimate the decline in productivity in each city under the counterfactual scenario based on the estimated effect of market access on productivity.

---

[40]The F statistic in the first stage is very large, which suggests a reassuring strong first stage. However, it also rings some alarm bells about the validity of the instrument. Spatial correlation in the error term could potentially lead to the very high F-statistic for weak instrument test. A solution to spatial correlation in panel IV regressions has been proposed in König et al. (2015). As a next step, I will follow their procedure and correct for spatial correlation in the error term.

[41]We need to assume that the effect of market access on aggregate productivity is log linear.

[42]Donaldson and Hornbeck (2016) conduct a similar counterfactual analysis to evaluate the loss of total land value in the US if all the railways were eliminated.

[43]A more realistic quantification of the effects of a reduction in transport costs on economic performance should include intermediate goods, which accounts for more than 70 percent of total trade. However in a more general model with trade in intermediate goods, a change in trade costs may make the current equilibrium unstable and lead to a bifurcation. In such case, the counterfactual exercise of studying the impact of removing all highways on economic performance may become very complicated due to the issue of multiple equilibria.

Table 2: City Market Access and Firm TFP

Dependent variable:			OLS					IV		
Firm TFP	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Market access	0.134***	0.065***		0.064***	0.069***	0.064***	0.042**		0.040**	0.046**
	(0.023)	(0.019)		(0.018)	(0.019)	(0.022)	(0.020)		(0.020)	(0.020)
Access to port			0.087**	0.087**	0.088**			0.087**	0.087**	0.088**
			(0.034)	(0.035)	(0.034)			(0.034)	(0.035)	(0.034)
Exporter dummy		-0.008	-0.008	-0.008	-0.008		-0.008	-0.008	-0.008	-0.008
		(0.009)	(0.009)	(0.009)	(0.009)		(0.009)	(0.009)	(0.009)	(0.009)
City population		-0.031	0.027	0.027	0.027		-0.031	0.027	0.027	0.027
		(0.080)	(0.070)	(0.073)	(0.074)		(0.080)	(0.070)	(0.071)	(0.073)
Node city × year					Yes					Yes
Railway Dist × year					Yes					Yes
Waterway Dist × year					Yes					Yes
Ownership structure		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	612484	598381	597085	597085	594476	612484	598381	597085	597085	594476
Adjusted. R squared	0.380	0.655	0.656	0.656	0.656	-	-	-	-	-
No. of clusters	339	339	335	335	335	339	339	335	335	335

Notes: This table includes results from OLS and IV regressions. The dependent variable is firm TFP and the main explanatory variable is market access. All regressions include industry-year and city fixed effects. Standard errors are clustered at the city level. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Finally, the loss in productivity for each city is weighed by city size to estimate the total national decline in productivity in absence of the national highway system.

The counterfactual analysis suggests that eliminating all highways in China would lower aggregate TFP by 3.2%.[44] In the baseline counterfactual, population is held fixed but it is likely that removing highways would change the distribution of population and production across cities. In fact, the effect of transport infrastructure on the distribution of population is the focus for a few studies mentioned above (Faber 2014; Baum-Snow et al. 2015).[45] I relax this assumption to allow highways to influence the distribution of population across cities. I use the observed population distribution in 1995 as a proxy for the counterfactual population distribution without highways.[46] I find that removing all highways would lower aggregate TFP by 3.8%. It is interesting that we get a larger increase in aggregate TFP by allowing population redistribution. The larger economic impact of highways suggests that workers moved to cities which experienced larger gain in productivity. Table 3 presents the estimated impact of removing all highways.

6.3   Decomposition of Productivity Gains from Highway Access

From the firm-level regressions, I have established that highways increased Chinese cities' productivity. The firm-level regressions only indicate that highway connections increased the production efficiency of a city. However, the channels through which highway connections enhanced aggregate production efficiency are unclear. In this section, I explore the mechanisms underlying the aggregate TFP gains resulting from highway connections. There are potentially four channels through which market access could affect a city's production efficiency: a within-firm productivity enhancement among continuing firms, the reallocation of market shares across continuing firms, the entry of productive new firms and the exit of inefficient firms. All channels except for the first are present in the model and the central

---

[44]Alternatively, I fix population to 1995 level, and GDP to 1995 and 2005 level. Results are not sensitive to different weights.

[45]With perfect labor mobility, city productivity is not determined by its market access according to the model I present in this paper. In that sense, the estimated impact from the empirical part should serve as a lower bound of the actual effect of highways on aggregate TFP.

[46]It is possible that the counterfactual population distribution differs from the population distribution in 1995 in important ways. The model I present does not predict migration. Donaldson and Hornbeck (2016) and Baum-Snow et al. (2015) do model population change due to transport infrastructure projects.

Table 3: Counterfactual Impacts on Aggregate TFP

	Estimated Decrease in Aggregate TFP without Highways
Baseline counterfactual scenario in absence of highways (holding city population constant from 2005)	3.16%
1. Holding the population distribution from 2000	3.12%
2. Holding the population distribution from 1995	3.09%
Allowing for changes in the distribution of population (holding total population constant)	3.81%

Notes: This table shows the estimated impact of removing all highways. Row 1-3 reports the counterfactual impact on aggregate TFP from eliminating all highways in China. In the baseline scenario, population distribution is held fixed from 2005. In Row 2, population distribution is assumed to be from 2000. In Row 3, population distribution is held fixed from 1995. In Row 4, I allow for changes in population distribution over time. The population distribution from 1995 is used as a proxy for the counterfactual population distribution without highways. All results assume that total population is held unchanged.

argument in the recent international trade literature (Pavcnik 2002; Bernard et al. 2003; Melitz 2003), but the first is not. However, many more recent papers (Lileeva and Trefler 2010; Bustos 2011; Garcia and Voigtländer 2013) document sizable within-plant or within-firm productivity growth after trade liberalization. To quantify the effects of each channel, I follow Haltiwanger (1997) to decompose changes in city productivity at the industry level into four terms. First, a productivity index for industry s is defined as follows:

$$lnTFP_{k,t}^s = \sum_i \theta_{i,k,t}^s lnTFP_{i,k,t}^s, \tag{22}$$

where $\theta_{i,t}^s$ is the share of output for firm i in industry s in city k at time t.

$$\Delta lnTFP_{k,t}^s = \sum_{continuers} \theta_{i,k,t-1}^s \Delta lnTFP_{i,k,t}^s +$$

$$\sum_{continuers} \Delta\theta_{i,k,t}^s (lnTFP_{i,k,t-1}^s - lnTFP_{k,t-1}^s) +$$

$$\sum_{entering firms} \theta_{i,k,t}^s (lnTFP_{i,k,t}^s - lnTFP_{i,k,t-1}^s) -$$

$$\sum_{exiting firms} \theta_{i,k,t-1}^s (lnTFP_{i,k,t-1}^s - lnTFP_{k,t-1}^s)$$

I regress each of the four components on changes in city market access, transportation cost to the nearest port, and other city-level characteristics and industry fixed effects. Table 4 reports the results from the decomposition exercise. Overall, the effect is very similar to that observed in the firm-level regressions – a 1% increase in market access leads to a 0.04% increase in the productivity index. I find that the entry of productive firms contributed the most to the aggregate gains in TFP. Reallocation among large incumbent firms and exit of inefficient firms also contributed to the productivity gain. These findings are consistent with evidence from other strands of the literature.[47] Consistent with the firm-level regressions, I do not find strong evidence of within-firm productivity improvements from an increase in domestic market access. Although the OLS result is significant and large, the IV result is much smaller and insignificant.

The decomposition exercise reveals the sources of TFP gains from highway infrastructure investments in China: transportation infrastructure affects city productivity mainly though the entry and expansion of relatively new and small firms and the contraction and exit of inefficient firms. Table 4 also presents the sources of TFP gains from easier access to the international market. Overall, a 1 % reduction in transportation costs to the nearest port increases aggregate TFP by 0.14%. Similar to domestic market access, the entry and expansion of productive young firms contributed to the TFP gains. In contrast to domestic market access, I find that within-firm productivity growth was an important source of ag-

---

[47]However, the exit effect is small. This may be because large firms, especially state-owned enterprises, were given many preferential policies to keep them alive, to the extent that the shut-down decision itself was subject to local politics.

gregate TFP growth. I find little evidence on reallocation across large incumbents. This is consistent with findings from Lileeva and Trefler (2010) on the liberalization of trade between the US and Canada.

Table 4: Decomposition of TFP gains

(a) OLS	Total effect	Within	Between	Entry	Exit
	(1)	(2)	(3)	(4)	(5)
Change in domestic market access	0.061**	0.039***	0.014***	0.045**	0.009***
	(0.031)	(0.012)	(0.005)	(0.020)	(0.002)
Change in access to ports	0.139***	0.065***	0.006	0.097***	-0.000
	(0.027)	(0.018)	(0.008)	(0.020)	(0.004)
No. of Observations	6413	6413	6413	6413	6413

(b) IV	Total effect	Within	Between	Entry	Exit
	(1)	(2)	(3)	(4)	(5)
Change in domestic market access	0.080**	0.007	0.021***	0.055**	0.011***
	(0.032)	(0.012)	(0.006)	(0.023)	(0.003)
Change in access to ports	0.134***	0.062***	0.006	0.095***	-0.001
	(0.027)	(0.017)	(0.008)	(0.020)	(0.004)
No. of Observations	6413	6413	6413	6413	6413

Notes: This table includes results from OLS and IV regressions. I decompose aggregate TFP into four components (the within-firm productivity growth, the reallocation between firms, the entry of new firms and the exit of incumbents). I then regress the change in each component on market access. Control variables include city population, provincial capital dummy, average years of schooling, access to railways, access to waterways and industry fixed effects. All regressions include industry fixed effects. Robust standard errors in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

6.4   Robustness Checks

Revenue Productivity and Prices

The distinction between revenue productivity and physical productivity has become increasingly important in the literature. Up to now, we have measured revenue productivity, not physical productivity (the connection and difference between the two will become clear below). As a result, we can not rule out a somewhat counter-intuitive case: the average firm's physical productivity decreases as a city's market access expands, which implies an increase in marginal cost. Since both terms enter revenue productivity and offset each other, it is still possible that revenue productivity goes up. Next, I investigate how firms' marginal costs

and prices respond to changes in market access.[48]

Results on marginal cost and price are presented in Table 10. Similar to earlier results, I do not find a strong effect of market access on markups.[49] However, we find that a 1% increase in market access leads to a 0.12% reduction in firms' marginal costs, which suggests that the effect of market access on firm productivity is more than twice as large as suggested by the regressions on TFPR if we consider the price effect. An interesting and important take-away from these regressions is that firms pass most of their cost advantages to consumers, implying economic integration through transportation costs reductions could potentially sizable welfare gains for consumers.

Isolating transportation costs reduction from GDP growth

In my market access measures, two factors affect a city's market access: 1) transportation costs 2) the size of its neighbors. One might be concerned that changes in a city's market access due to changes in its neighbors' size might be correlated with some unobserved factors that also affect firm productivity in the city. Moreover, since I am mainly interested in the effects of transportation costs in this paper, I would like to isolate the changes in market access due to a reduction in transportation costs from changes in sizes of neighbors. In Table 11, I show results when I use population of all cities in 1995 and the only time variation in a city's market access comes from changes in bilateral transportation costs. It is clear that shutting down the size channel does not substantially alter my results. If anything, the effects from the transportation costs channel are slightly larger than the corresponding baseline results.

Excluding the "Node" Cities

Another threat to my identification strategy is the potential correlation between highway placement and some unobserved placed-based policies. In the baseline regressions, I remove city GDP to alleviate the this concern, but one may still be concerned that those "node"

---

[48]Please refer to Appendix A.3 for an illustration of the idea in Garcia and Voigtländer (2013).

[49]This result at first seems to be in contrast with one of the main predictions in Melitz and Ottaviano (2008). However, De Loecker et al. (2016) also find markups did not decrease after trade liberalization India. In fact, they find markups went up because firms' marginal costs decreased and firms could cut prices and raise markup simultaneously.

cities targeted by the government may have other characteristics that affected both highway placement and TFP growth. I re-run the same regressions as in the baseline but exclude the 52 cities that were targeted as "nodes" by the government when it planned the national highway network. The results are presented in Table 12. Note that the exclusion of node cities does not change the results. Since my results are not driven by those target cities, I can be confident that unobserved city characteristics do not threaten the validity of my results.

An Alternative Way for Estimating Market Access

Donaldson and Hornbeck (2016) show that a location's market access can be expressed as the sum over the transportation costs with each other county, that other location's population, and that other location's access to other markets. The derivation is relegated to the the Appendix. They numerically solve for market access for all US counties. Donaldson and Hornbeck (2016) also verify that the the results they get from using their market access measure is consistent with their baseline results from using the "market potential" term by Harris (1954). I follow their strategy of estimating city market access by solving a system of nonlinear equations. I then run regressions with the numerically solved market access measure, and confirm that the two measures generate very similar results.[50]

7 Spatial Reallocation of Industries

Now I turn to the effects of highways on the spatial reallocation of industries. As explained in the introduction, classic trade theories predict industry specialization patterns after trade liberalization based on either technological differences, endowment differences or returns to scale. Following the literature, I explore three important dimensions of industry characteristics that will lead to heterogeneous responses to highway connection. I use weight to value ratio to measure the iceberg transportation cost of an industry. I use the median firm's capital to labor ratio to measure industry's capital intensity. I also take the trade elasticities

---

[50]The estimation of market access is implemented in MATLAB. MATLAB code is available upon request.

estimated in the trade literature as a measure of the degree of product differentiation of an industry.[51] I use the following baseline specification[52] for estimating the effects of market access on employment growth and firm growth:

$$\Delta log Y_{j,k} = \alpha + \delta \Delta log M A_k \times Ind_j + \beta X_{k,j} + \mu_k + \nu_j + \epsilon_{j,k}$$

where $logMA_k$ is city k's change in market access, Ind_j is a ranking of industries depending on what aspects we are interested in. I am most interested in three sources of industry heterogeneity: industry's degree of product differentiation, as measured by the elasticity of substitution; industry's capital intensity, as measured by the median capital-labor ratio of firms; industry's transportation cost, as measured by the median weight-to-value ratio. I include initial share of the industry in total output as a control variable to account for any convergence effect.

In Table 12-14, I present the effects of highway network on employment growth across industries and across cities. In Table 13, I find that the coefficient on the interaction term between market access and capital to labor ratio is positive, meaning that industries with higher capital intensity grew disproportionately faster in cities with large market access. Results do not change very much if I include the interaction between market access and city population, as well as the inclusion of interaction between market access and driving time to nearest port. The instrument variable approach generates similar results to OLS regressions. One way to interpret the coefficients of the interaction term is the following: a one-standard deviation of a city's market access would increase annual employment growth in the sector at the 75th percentile of the distribution by capital intensity by 1.6 percentage points more than annual employment growth in the sector at the 25th percentile for the period from 2001 to 2008. Cities with large market access will find themselves have a comparative advantage in capital-intensive industries.[53] In Table 14, the coefficient on the interaction term

---

[51]Please note that here the market access term is very similar to Donaldson and Hornbeck (2016) The model in Section 3 does not generate precise predictions for industry reallocation. So, I follow the literature and use the market access term used in recent papers.

[52]This type of specification has been use extensively in the literature to examine the effects of various frictions on growth and trade, such as financial constraints (Rajan and L. 1998; Manova 2013), labor market protection (Tang 2012).

[53]More analysis will need to be done to identify the underlying mechanisms. Gaubert (2014) argues that larger cities tend to specialize in capital intensive industries because wage is higher in big cities. Hanson

between market access and weight to value ratio is again positive, suggesting that industries with weight-to-value ratio (high transportation costs) grew disproportionately faster in cities with large market access. Similarly, Table 15 presents results on product differentiation. The results suggest that industries with a low degree of product differentiation grew disproportionately faster in cities with large market access. This result is in contrast to Hanson and X. (2004), in which they find more industries with differentiated products tend to locate in large *countries*.[54] Overall, my results show that transport infrastructure redistributes industries spatially, and the findings are in general consistent with theories of trade and economic geography.

8   Discussion and Conclusion

Transport infrastructure is one of the most expensive public goods in the world. Governments across the world spend billions of dollars to build highways and railways to facilitate the movement of goods and people. Existing empirical studies focus on the effects of railways or highways on GDP or population growth and have found mixed results. In this paper, I examine the channels through which transport infrastructure affects economic outcomes. I find that the national highway network in China promoted production efficiency, delivered sizable welfare gains to consumers, and led to a sectoral reallocation between cities. Aggregate TFP growth resulting from reduced transportation costs is attributable primarily to firm entry and resource reallocation. I also find that cities with large market access specialized in industries that have low unit cost, higher capital intensity, and low product differentiation.

The findings presented in this paper have important policy implications. Facing the threat of secular stagnation, policymakers are searching for tools to boost aggregate demand in the short run and promote economic growth in the long run. After the global recession, there has been a growing interest among policymakers worldwide in using infrastructure investments

---

(2005) also finds that wages tend to be higher in locations with larger market access.

[54]As mentioned in Hanson and X. (2004), the robustness of the home-market effect in Krugman (1980) is still an open question. The prediction hinges critically on model assumptions.

both as a short-term fiscal policy instrument and as a long-term growth generator.[55] For example, the World Bank has consistently dedicated itself to investing in infrastructure in low-income countries to fight poverty. The International Monetary Fund is also actively advocating for more infrastructure investments in Latin America and Africa to meet the infrastructure needs and boost economic growth in these regions. The two recently-founded development banks–the Asian Infrastructure Investment Bank and the New Development Bank, were established under the leadership of China to address the increasing infrastructure needs in Asia. The US president-elect Donald Trump envisions a trillion-dollar infrastructure plan. The increasing use of infrastructure projects by policymakers begs the question of whether the huge amount of tax dollars spent on infrastructure is well justified by their potential benefits.

This paper provides novel evidence that transport infrastructure promotes economic growth and sheds light on the mechanisms underlying the gains from infrastructure investment. I examine the channels of productivity gains resulting from infrastructure investments and quantify the relative importance of each channel. I highlight the role of highway infrastructure in raising aggregate TFP by facilitating resource reallocation between heterogeneous firms. This paper also evaluates the aggregate economic impact of China's national highway system. Findings in this paper suggest a sizable economic impact of infrastructure investments and market integration when domestic transportation costs are large and misallocation is pervasive.[56] To make investment decisions, however, we will also need to evaluate potential dynamic gains from such as a large-scale transport infrastructure project and its economic costs, which is beyond the scope of this paper.

There are a few avenues for future research. As mentioned in the introduction, a natural byproduct of this paper is a quantitative exercise that compares aggregate productivity growth under the actual highway construction with counterfactual scenarios. Given the enormous cost of infrastructure projects, it is important to compare the potential economic growth resulting from infrastructure to both its direct and indirect cost to aid policy rec-

---

[55]During the crisis, China, India and Korea led the way to spend on infrastructure as part of their stimulus packages.

[56]This paper focuses on productivity gains from highways in the manufacturing sector but neglects potential gains in other sectors.

ommendations. Another interesting area for future research is inter-sectoral linkages. So far, I have ignored the inter-connectedness of sectors when estimating market access. One interesting extension would be to incorporate inter-sectoral linkages and construct output market access and input market access measures using input-output tables. Such output and input market access measures would not only allow for the use of cross-industry and within-city variation to identify the impact of transportation infrastructure but also make it possible to differentiate the impact of output access from input access. Finally, it would also be interesting to conduct an in-depth focus study on some of the most remote and poorest areas. To do so, one would need to examine the county or even village level data. Such studies at the more micro-level would help to shed light on the role of infrastructure in reducing poverty and affecting income inequality.

References

Ackerberg, D. A., K. Caves, and G. Frazer, "Identification properties of recent production function estimators," *Econometrica*, 2015, *83(6)*, 2411–2451.

Alder, S., "Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development," in "2015 Meeting Papers" number 1447 Society for Economic Dynamics 2015.

Anderson, J. E. and E. Van Wincoop, "Gravity with gravitas: a solution to the border puzzle," *the american economic review*, 2003, *93* (1), 170–192.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare, "New trade models, same old gains?," *The American Economic Review*, 2012, *102* (1), 94–130.

Asturias, J., M. García-Santana, and R. Ramos Magdaleno, "Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India," 2015.

Autor, D, David Dorn, and Gordon H Hanson, "The China syndrome: Local labor market effects of import competition in the United States," *The American Economic Review*, 2013, *103* (6), 2121–2168.

Banerjee, A., E. Duflo, and N. Qian, "On the road: Access to transportation infrastructure and economic growth in China," *National Bureau of Economic Research*, 2012, (No. w17897.).

Baum-Snow, N., "Did highways cause suburbanization?," *The Quarterly Journal of Economics*, 2007, pp. 775–805.

_ , **L. Brandt, V. Henderson, M. Turner, and Q. Zhang**, "Transport Infrastructure, Urban Growth and Market Access in China," *ERSA conference papers. European Regional Science Association*, October 2015.

Behrens, K., G. Duranton, and F. Robert-Nicoud, "Productive cities: Sorting, selection, and agglomeration," *Journal of Political Economy*, 2014, *122(3), 507-553.*

Bernard, Andrew B, Jonathan Eaton, J Bradford Jensen, and Samuel Kortum, "Plants and productivity in international trade," *The American Economic Review*, 2003, *93* (4), 1268–1290.

Brandt, L., J. Van Biesebroeck, and Y. Zhang, "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing," *Journal of Development Economics*, 2012, *97* (2), 339–351.

_ , _ , **L. Wang, and Y. Zhang**, "WTO accession and performance of Chinese manufacturing firms," *Center for Economic Policy Research*, 2016.

Bustos, Paula, "Trade liberalization, exports, and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms," *The American economic review*, 2011, *101* (1), 304–340.

Chaney, T., "Distorted gravity: the intensive and extensive margins of international trade," *The American Economic Review*, 2008, *98.4*, 1707–1721.

Costinot, A. and A. Rodriguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," *The Handbook of International Economics*, 2014, pp. 197–261.

Donaldson, D., "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *forthcoming, American Economic Review*, 2016.

Donaldson, Dave and Richard Hornbeck, "Railroads and American Economic Growth: A "Market Access" Approach," *The Quarterly Journal of Economics*, 2016, p. qjw002.

Duranton, G. and M. A. Turner, "The fundamental law of road congestion: Evidence from US cities," *The American Economic Review*, 2011, *101* (6), 2616–2652.

_ , **P. M. Morrow, and A. Turner**, "Roads and Trade: Evidence from the US," *The Review of Economic Studies*, 2014, *81* (2), 681–724.

Eaton, J. and S. Kortum, "Technology, geography, and trade," *Econometrica*, 2002, *70* (5), 1741–1779.

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

Feenstra, R. C., Z. Li, and M. Yu, "Exports and credit constraints under incomplete information: Theory and evidence from China," *Review of Economics and Statistics*, 2014, *96* (4), 729–744.

Frye, D., "Transportation Networks and the Geographic Concentration of Industry," 2014.

Garcia, A. and N. Voigtländer, "Exporting and Plant-Level Efficiency Gains: It's in the Measure," Technical Report, National Bureau of Economic Research 2013.

Gaubert, C., "Firm sorting and agglomeration," *Unpublished manuscript, UC Berkeley*, 2014.

Ghani, E., A. G. Goswami, and W. R. Kerr, "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing," *Econ Journal*, 2016, *126*, 317–357 doi:10.1111/ecoj.12207.

Haltiwanger, J. C., "Measuring and analyzing aggregate fluctuations: the importance of building from microeconomic evidence," *St. Louis Fed Review*, 1997, *79.*

Hanson, G. H., "Market potential, increasing returns and geographic concentration," *Journal of international economics*, 2005, *67(1)*, 1–24.

_ **and Chong X.**, "The Home-Market Effect and Bilateral Trade Patterns," *American Economic Review,*, 2004, *94(4)*, 1108–1129.

Harris, C. D., "The, Market as a Factor in the Localization of Industry in the United States," *Annals of the association of American geographers*, 1954, *44* (4), 315–348.

König, M, D Rohner, M Thoenig, and F Zilibotti, "Networks in conflict: Theory and evidence from the great war of africa," *University of Zurich, UBS International Center of Economics in Society, Working Paper*, 2015, (14).

Krugman, P., "Scale economies, product differentiation, and the pattern of trade," *The American Economic Review*, 1980, *70* (5), 950–959.

_ , "Increasing Returns and Economic Geography," *The Journal of Political Economy*, 1991, *99(3)*, 483–499.

Levinsohn, J. and A. Petrin, "Estimating production functions using inputs to control for unobservables," *The Review of Economic Studies*, 2003, *70(2)*, 317–341.

Lileeva, A. and D. Trefler, "Improved Access to Foreign Markets Raises Plant-level Productivity...For Some Plants," *The Quarterly Journal of Economics*, 2010, *125 (3)*, 1051–1099.

Limao, N. and A. J. Venables, "Infrastructure, geographical disadvantage, transport costs, and trade," *The World Bank Economic Review*, 2001, *15* (3), 451–479.

Loecker, J. De and F. Warzynski, "Markups and firm-level export status," *The American Economic Review.*, 2012, *102(6), 2437-2471.*

_ **, K. Goldberg Pinelopi, Amit K. Khandelwal, and Nina Pavcnik**, "Prices, Markups, and Trade Reform," *Econometrica*, 2016, *84* (no. 2), 445–510.

Manova, K., "Credit constraints, heterogeneous firms, and international trade," *The Review of Economic Studies*, 2013, *80* (2), 711–744.

Melitz, M. J. and D. Trefler, "Gains from trade when firms matter," *The Journal of Economic Perspectives*, 2012, *26(2)*, 91–118.

Melitz, M.J., "The impact of trade on intraindustry reallocations and aggregate industry productivity," *Econometrica*, 2003, *71(6)*, pp.1695–1725.

_ **and G.I. Ottaviano**, "Market size, trade, and productivity," *The review of economic studies*, 2008, *75(1)*, 295–316.

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

Nagy, D., "City location and economic development," *mimeo*, 2015.

Olley, S. and A. Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64 (6)*, 1263–1298.

Pavcnik, N., "Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants," *The Review of Economic Studies*, 2002, *69(1)* (245-276).

Rajan, R.G. and Zingales L., "Financial Dependence and Growth," *The American Economic Review 88 (3) American Economic Association: 559–86*, 1998.

Redding, S., "Goods trade, factor mobility and welfare," *Journal of International Economics*, 2016, *101*, 148–167.

_ **and A.J. Venables**, "Economic geography and international inequality," *Journal of international Economics*, 2004, *62(1)*, 53–82.

Smeets, V. and F. Warzynski, "Estimating productivity with multi-product firms, pricing heterogeneity and the role of international trade," *Journal of International Economics*, 2013, *90(2)*, 237–244.

Tang, H., "Labor market institutions, firm-specific skills, and trade patterns," *Journal of International Economics*, 2012, *87* (2), 337–351.

Wooldridge, J. M., "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics Letters*, 2009, *104(3)*, 112–114.
World Development Report 2009: Reshaping Economic Geography

World Development Report 2009: Reshaping Economic Geography, *The World Bank, Washington DC, 2009.*

Figures

Figure 3: The National Trunk Highway System
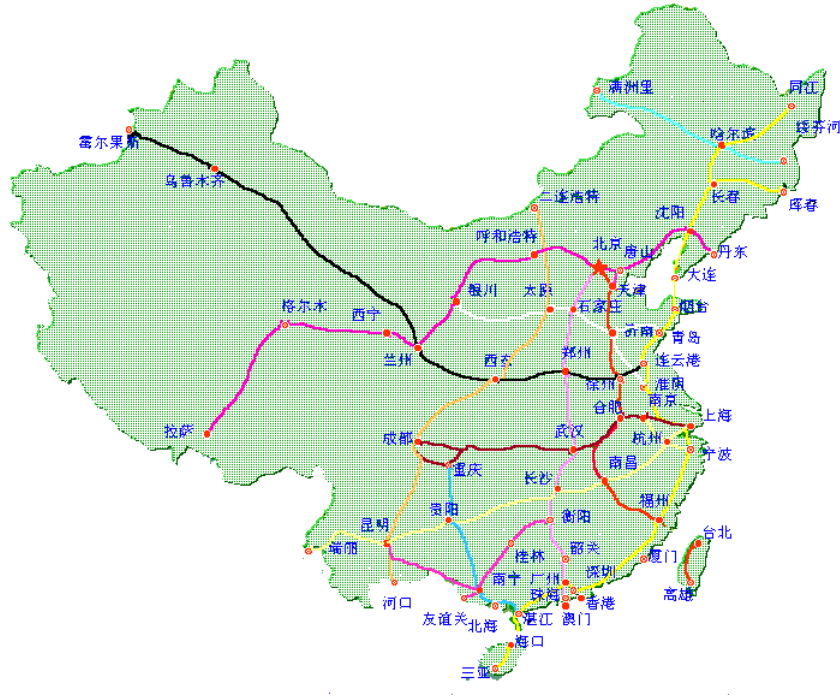
Figure 4: Inter-Provincial Trade
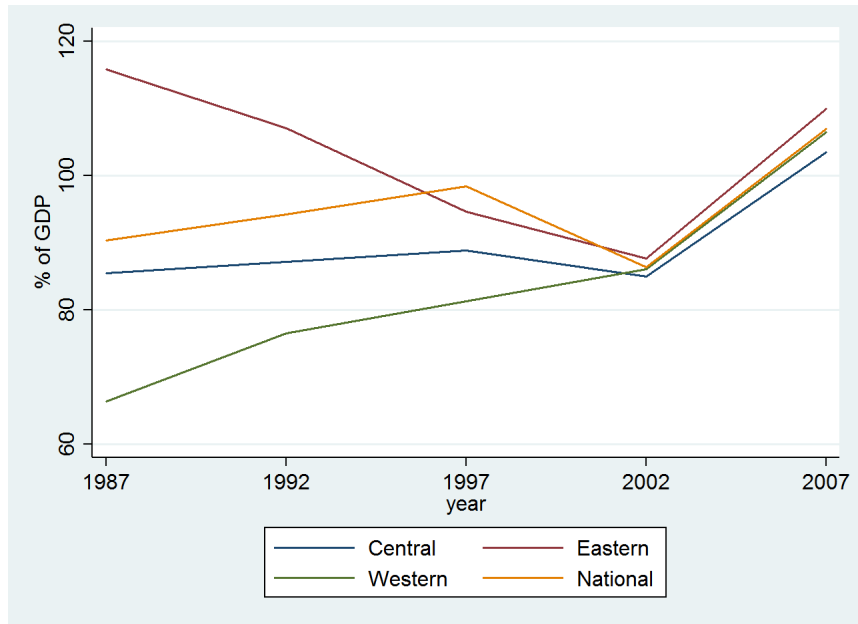
Figure 5: Highway network in 1995, 2001, 2005
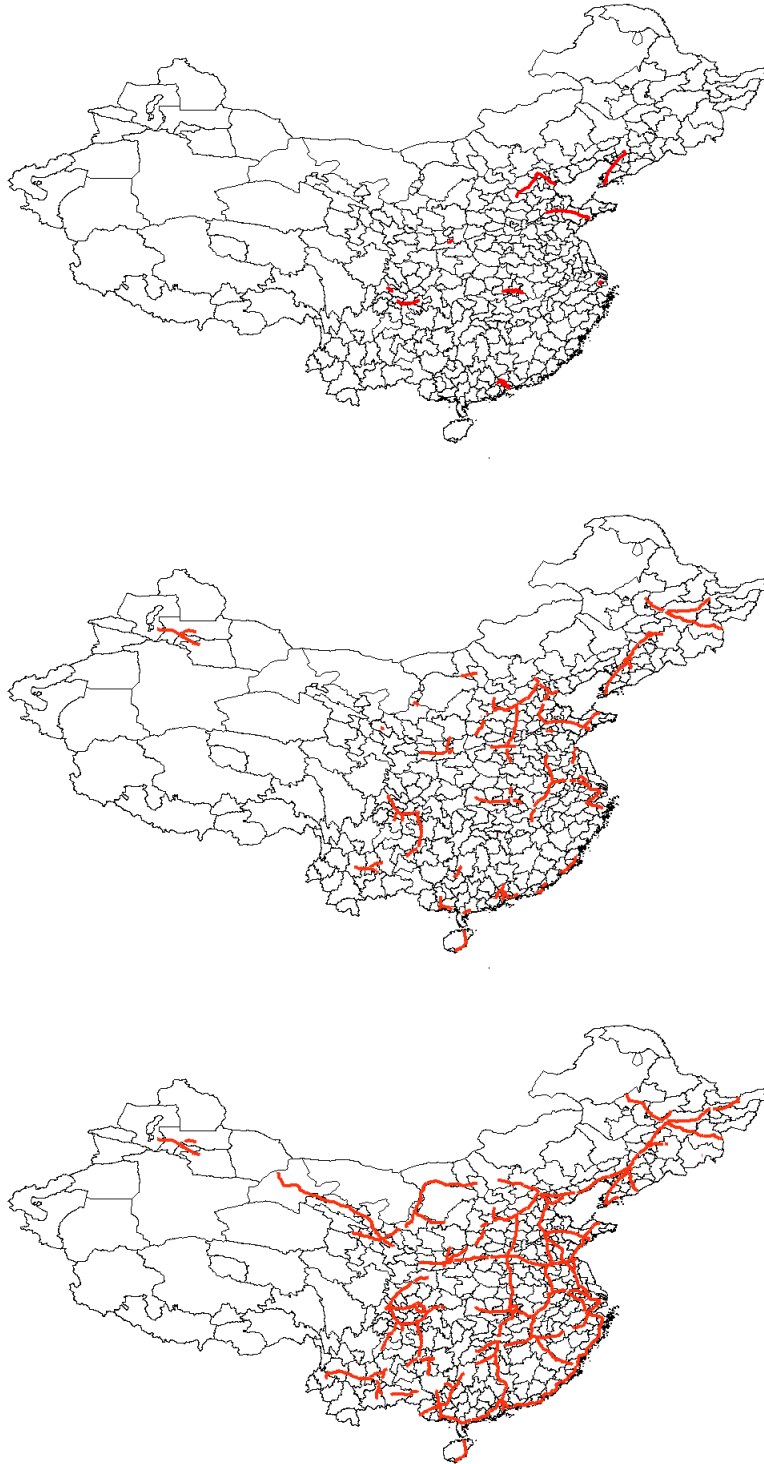
Notes: this figure shows how China's National Trunk Highway System evolved over time. Please note that there were many local construction that were not part of the NTHS. These highways were built for intra-city connections or connecions to nearby cities.

Figure 6: Estimated changes in market access 1995-2005

1995

2001

2005

(17.881,18.71453]
(17.04747,17.881]
(16.21394,17.04747]
(15.38041,16.21394]
(14.54688,15.38041]
(13.71335,14.54688]
(12.87982,13.71335]
(12.04629,12.87982]
(11.21277,12.04629]
No data
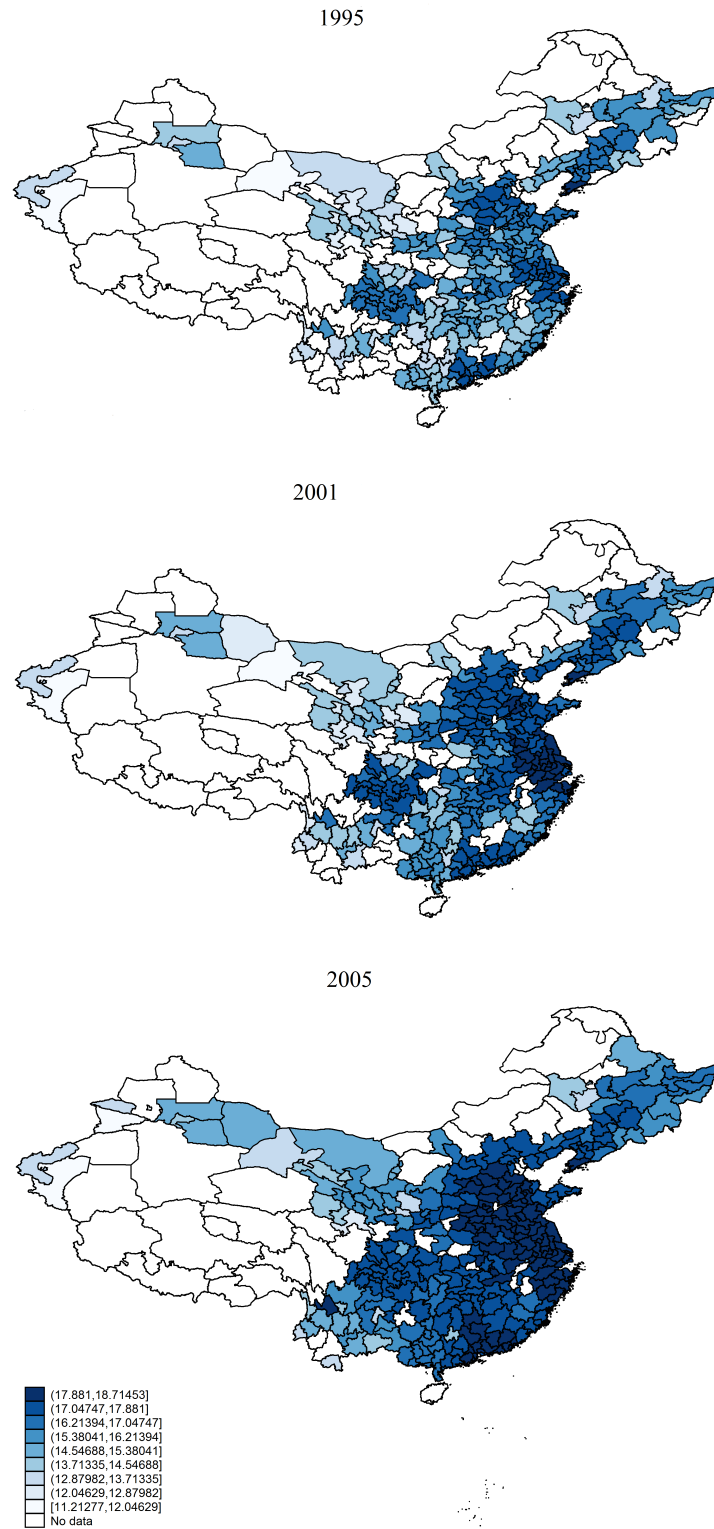
Notes: this figure shows how market access evolved over time for 339 Chinese cities. Darker color means larger market access.

Tables

Table 5: Estimation of Transportation Costs

Dependent Variable: Iceberg Transport cost	(1)	(2)	(3)	(4)
Length of highway	0.454***	0.539***	1.331***	1.105***
	(0.001)	(0.001)	(0.007)	(0.008)
Length of local road	1.092***	0.712***	1.215***	1.199***
	(0.005)	(0.026)	(0.045)	(0.045)
Length of highway (squared)			-0.381***	-0.315***
			(0.004)	(0.004)
Length of local road (squared)			-0.938***	-0.780***
			(0.056)	(0.055)
Length of highway (3 squared)			0.048***	0.041***
			(0.001)	(0.001)
Length of local road (3 squared)			0.327***	0.258***
			(0.017)	(0.016)
Travel time				0.264***
				(0.005)
Travel time (squared)				-0.022***
				(0.001)
Travel time (3 squared)				0.001***
				(0.000)
Origin fixed effects	No	Yes	Yes	Yes
Destination fixed effects	No	Yes	Yes	Yes
No. of observations	97253	97253	97253	97253
Adjusted. R squared	0.689	0.850	0.869	0.875

Notes: This table shows estimation results from a simple linear regression model designed to estimate the relationship between transport costs and the road network structure. All regressions include origin and destination fixed effects. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 6: Productivity of Chinese Manufacturing Firms

Chinese Industrial Classfication (2-digit)	Industry Aggregate TFP (log)									
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Processing of Foods (13)	5.05	4.96	5.09	5.18	5.35	5.46	5.57	5.62	5.67	5.80
Manufacturing of Foods (14)	5.49	5.16	5.53	5.54	5.62	5.79	5.89	6.00	6.09	6.14
Manufacture of Beverages (15)	3.88	3.77	3.89	3.99	4.05	4.18	4.30	4.55	4.56	4.60
Manufacture of Tobacco (16)	2.49	2.47	2.49	2.64	3.05	3.21	3.13	3.15	3.34	3.42
Manufacture of Textile (17)	5.97	6.01	6.07	6.21	6.33	6.45	6.46	6.70	6.78	6.96
Manufacture of Apparel, Footwear & Caps (18)	5.46	5.16	5.18	5.33	5.32	5.45	5.54	5.74	5.95	5.89
Manufacture of Leather, Fur, & Feather (19)	5.04	5.04	5.00	5.08	5.13	5.18	5.29	5.28	5.44	5.54
Processing of Timber, Manufacture of Wood and Bamboo Products (20)	4.68	3.85	3.85	3.93	4.07	4.09	4.33	4.27	4.35	4.57
Manufacture of Furniture (21)	5.05	4.66	4.77	4.77	4.96	4.92	5.14	5.07	5.26	5.41
Manufacture of Paper & Paper Products (22)	5.08	5.08	5.18	5.31	5.48	5.64	6.00	5.88	6.00	6.17
Printing, Reproduction of Recording Media (23)	5.52	5.45	5.57	5.80	5.87	5.97	5.99	6.25	6.32	6.43
Manufacture of Articles For Culture, Education & Sport Activities (24)	6.27	6.26	6.31	6.31	6.41	6.56	6.54	6.69	6.85	6.97
Processing of Petroleum, Coking, &Fuel (25)	4.95	4.86	4.93	5.05	5.15	5.29	5.36	5.39	5.46	5.48
Manufacture of Raw Chemical Materials (26)	5.64	5.47	5.69	5.81	5.94	6.31	6.69	6.52	6.56	6.69
Manufacture of Medicines (27)	6.16	6.18	6.23	6.30	6.50	6.55	6.58	6.66	6.77	6.83
Manufacture of Chemical Fibers (28)	4.35	4.51	4.79	4.68	4.80	5.12	5.21	5.23	5.38	5.48
Manufacture of Rubber (29)	5.77	5.73	5.82	5.98	6.24	6.42	6.47	6.75	6.70	6.75
Manufacture of Plastics (30)	4.72	4.58	4.65	4.74	4.87	4.96	4.91	4.96	5.14	5.26
Manufacture of Non-metallic Mineral goods (31)	4.81	4.73	4.82	4.97	5.08	5.24	5.49	5.80	5.68	5.87
Smelting & Pressing of Ferrous Metals (32)	4.49	4.32	4.43	4.62	4.83	5.22	5.52	5.52	5.51	5.50
Smelting & Pressing of Non-ferrous Metals (33)	6.30	6.49	6.60	6.69	7.02	7.01	7.24	7.47	7.87	7.96
Manufacture of Metal Products (34)	5.58	5.49	5.57	5.66	5.92	6.00	5.94	5.96	6.07	6.32
Manufacture of General Purpose Machinery (35)	5.34	5.39	5.44	5.65	5.86	6.10	6.40	6.55	6.74	6.90
Manufacture of Special Purpose Machinery (36)	5.01	5.13	5.30	5.39	5.75	6.01	6.06	6.05	6.27	6.42
Manufacture of Transport Equipment (37)	4.85	4.85	5.00	5.33	5.60	6.01	6.06	6.11	6.14	6.22
Electrical Machinery & Equipment (39)	5.41	5.62	5.69	5.74	5.81	5.97	6.05	6.11	6.28	6.34
Computers & Other Electronic Equipment (40)	7.78	7.97	8.00	8.20	8.52	8.40	8.77	8.70	8.61	8.35
Measuring Instruments & Machinery for Cultural Activity & Office Work (41)	5.85	5.41	5.38	5.69	5.91	6.18	6.01	6.23	6.34	6.42
Manufacture of Artwork (42)	5.52	5.32	5.44	5.60	5.68	5.81	5.68	6.08	5.93	6.46

Notes: This table shows the industry-level productivity index in China from 1998 to 2007. The construction of the productivity index follows (22) in 6.2. The index is expressed in log terms. The 2-digit industry classification changed in 2003, and the years after 2003 are converted to the classification before the change to ensure consistency over time.

Table 7: First Stage Regressions

	log MA	log MA (excluding nodes)	log MA (constant GDP)
log MA (IV)	0.967***	0.922***	0.960***
	(0.035)	(0.041)	(0.036)
Access to Port	0.022	0.060	0.028
	(0.033)	(0.041)	(0.034)
Exporter dummy	0.002**	0.004***	0.002**
	(0.001)	(0.001)	(0.001)
City population	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)
Railway Dist $\times$ year	Yes	Yes	Yes
Waterway Dist $\times$ year	Yes	Yes	Yes
Ownership structure	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
No. of observations	608208	370615	608208
Adjusted. R squared	0.991	0.992	0.982
First-stage F statistic	723.433	489.353	656.805

Notes: This table presents results from the first-stage regressions. The dependent variable is the market access estimated with actual highway network data. The independent variable is the hypothetical market access constructed from the least-cost path spanning tree network. Standard errors are clustered at the city level. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 8: Market Access and Within-Firm TFP Growth

Dependent variable:			OLS					IV		
Firm TFP	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Market access	0.038	0.019		0.016	0.022	0.023	0.017		0.013	0.021
	(0.024)	(0.014)		(0.014)	(0.015)	(0.026)	(0.015)		(0.016)	(0.016)
Access to port			0.081**	0.080**	0.072**			0.081**	0.080**	0.072**
			(0.033)	(0.033)	(0.033)			(0.033)	(0.033)	(0.034)
Exporter dummy		-0.006	-0.006	-0.006	-0.002		-0.006	-0.006	-0.006	-0.002
		(0.009)	(0.009)	(0.009)	(0.009)		(0.009)	(0.009)	(0.009)	(0.009)
City population					Yes					Yes
					Yes					Yes
Node city × year					Yes					Yes
Railway Dist × year					Yes					Yes
Waterway Dist × year					Yes					Yes
Ownership structure		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	322546	313209	312381	312381	312381	322546	313209	312381	312381	312381
Adjusted. R squared	0.691	0.822	0.823	0.823	0.823	-	-	-	-	-
No. of clusters	339	339	335	335	335	339	339	335	335	335

Notes: This table includes results from OLS and IV regressions. The dependent variable is firm TFP and the main explanatory variable is market access. All regressions include industry-year and city fixed effects. Standard errors are clustered at the city level. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 9: Market Access and Markups

Dependent variable:	OLS			IV		
Firm Markup	(1)	(2)	(3)	(4)	(5)	(6)
Market Access	-0.007		-0.002	-0.003		-0.001
	(0.020)		(0.018)	(0.022)		(0.019)
Access to port		0.072	0.072		0.073	0.075
		(0.046)	(0.046)		(0.045)	(0.045)
Exporter dummy	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
City population	-0.173*	-0.140	-0.140	-0.171	-0.140	-0.140
	(0.097)	(0.086)	(0.086)	(0.098)	(0.086)	(0.086)
Node city × year	Yes	Yes	Yes	Yes	Yes	Yes
Railway Dist × year	Yes	Yes	Yes	Yes	Yes	Yes
Waterway Dist × year	Yes	Yes	Yes	Yes	Yes	Yes
Ownership structure	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	595635	594476	594476	595635	594476	594476
Adjusted. R squared	0.84	0.841	0.841	-	-	-
No. of clusters	333	331	331	333	331	331

Notes: This table includes results from OLS and IV regressions. The dependent variable is firm markups and the main explanatory variable is market access. All regressions include industry-year and city fixed effects. Standard errors are clustered at the city level. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 10: Market Access, Marginal Cost and Price

Dependent variable:	Marginal Cost			Price		
	(1)	(2)	(3)	(4)	(5)	(6)
Market Access	-0.155***	-0.170***	-0.120**	-0.084*	-0.086*	-0.095**
	(0.048)	(0.049)	(0.046)	(0.046)	(0.047)	(0.047)
Access to port		-0.036	-0.022		-0.036	-0.022
		(0.100)	(0.094)		(0.100)	(0.094)
Exporter dummy			0.323***			0.323***
			(0.046)			(0.043)
City population		0.366*	0.210		0.363*	0.242
		(0.209)	(0.190)		(0.208)	(0.192)
Node city × year		Yes	Yes		Yes	Yes
Railway Dist × year		Yes	Yes		Yes	Yes
Waterway Dist × year		Yes	Yes		Yes	Yes
Ownership structure	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	197656	196145	196143	197656	196145	196143
Adjusted. R squared	0.389	0.389	0.406	-	-	-
No. of clusters	338	331	331	338	331	331

Notes: The table presents results from a regression of marginal cost and price on market access. All regressions include industry-year and city fixed effects. All standard errors are clustered at the city level. Standard errors in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 11: The NTHS and Firm TFP: Fixing City GDP

Dependent variable:	OLS			IV		
Firm TFP	(1)	(2)	(3)	(1)	(2)	(3)
Market Access	0.052***		0.053***	0.051***		0.053***
	(0.015)		(0.014)	(0.015)		(0.014)
Access to port		0.075**	0.076**		0.075**	0.076**
		(0.031)	(0.031)		(0.031)	(0.033)
City population	-0.051	-0.004	-0.007	-0.051	-0.004	-0.004
	(0.070)	(0.066)	(0.066)	(0.070)	(0.066)	(0.066)
Exporter dummy	-0.007	-0.008	-0.008	-0.007	-0.008	-0.002
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Ownership structure	Yes	Yes	Yes	Yes	Yes	Yes
Node city × year	Yes	Yes	Yes	Yes	Yes	Yes
Railway Dist × year	Yes	Yes	Yes	Yes	Yes	Yes
Waterway Dist × year	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	595635	594476	594476	595635	594476	594476
Adjusted. R squared	0.654	0.654	0.654	-	-	-
No. of clusters	333	331	331	333	331	331

Notes: This table includes results from OLS and IV regressions. The dependent variable is firm TFP and the main explanatory variable is market access. All regressions include industry-year and city fixed effects. Standard errors are clustered at the city level. Standard errors in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 12: The NTHS and Firm TFP: Excluding the "Node" Cities

Dependent variable: Firm TFP	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Market Access	0.152***	0.053***		0.050***	0.057***	0.096***	0.048***		0.047***	0.053***
	(0.023)	(0.014)		(0.014)	(0.015)	(0.020)	(0.016)		(0.016)	(0.016)
Access to port			0.066**	0.066**	0.072**			0.068**	0.072**	0.072**
			(0.029)	(0.029)	(0.029)			(0.029)	(0.030)	(0.031)
Exporter dummy		-0.008	-0.008	-0.008	-0.008		-0.008	-0.008	-0.008	-0.008
		(0.009)	(0.009)	(0.009)	(0.009)		(0.009)	(0.009)	(0.009)	(0.009)
City population					0.043					0.0064
					(0.069)					(0.069)
Railway Dist × year					Yes					Yes
Waterway Dist × year					Yes					Yes
Ownership structure		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	428739	418867	417960	417960	416133	428739	418867	417960	417960	416133
Adjusted. R squared	0.421	0.682	0.693	0.694	0.694	-	-	-	-	-
No. of clusters	285	285	280	280	280	285	285	280	280	280

Notes: This table includes results from OLS and IV regressions. The dependent variable is firm TFP and the main explanatory variable is market access. All regressions include industry-year and city fixed effects. Standard errors are clustered at the city level. Standard errors in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 13: Employment and Firm Growth: Capital Intensity

Dependent Variable:	Growth of Employment		Growth of No. of firms	
	OLS	IV	OLS	IV
K/L*Δ Market Access	0.642***	0.644***	0.462***	0.463***
	(0.136)	(0.136)	(0.076)	(0.078)
K/L*Δ Population	0.006	0.004	-0.005	-0.006
	(0.010)	(0.009)	(0.006)	(0.006)
K/L*Δ Access to port	0.022*	0.026*	0.012	0.013
	(0.013)	(0.014)	(0.009)	(0.010)
K/L*Rail Access	0.010**	0.010**	0.005**	0.005**
	(0.000)	(0.000)	(0.002)	(0.002)
Initial industry Share	-0.001***	-0.001***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Industry FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
No. of observations	8250	8250	8251	8251
No. of clusters	319	319	319	319

Note: the dependent variable in the two columns on the left is the change in employment. The dependent variable on the right is the change in number of firms. The coefficient of K/L*Δ Market Access is positive and significant, which implies that industries with high capital intensity grew disproportionately faster in cities that gained large market access. All regressions include sector and city fixed effects. Standard errors are clustered at the city level and province-industry level. Standard errors in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 14: Employment and Firm Growth: Weight-to-Value Ratio

Dependent Variable:	Growth of Employment		Growth of No. of firms	
	OLS	IV	OLS	IV
Weight/Value*Δ Market Access	0.004***	0.004***	0.002***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Weight/Value*Δ Population	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Weight/Value*Δ Access to port	0.000	0.000*	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)
Weight/Value*Rail Access	0.001	0.000	0.002	0.001
	(0.000)	(0.000)	(0.003)	(0.000)
Initial industry Share	-0.001***	-0.001***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Industry FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
No. of observations	8250	8250	8251	8251
No. of clusters	319	319	319	319

Note: the dependent variable in the two columns on the left is the change in employment. The dependent variable on the right is the change in number of firms. The coefficient of Weight/Value*Δ Market Access is positive and significant, which implies that industries with large weight-to-value ratio grew disproportionately faster in cities that gained large market access. All regressions include sector and city fixed effects. Standard errors are clustered at the city level and province-industry level. Standard errors in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Table 15: Employment and Firm Growth: Product Differentiation

Dependent Variable:	Growth of Employment		Growth of No. of firms	
	OLS	IV	OLS	IV
ProdDiff*Δ Market Access	-0.008**	-0.009**	-0.007***	-0.009***
	(0.004)	(0.004)	(0.002)	(0.002)
ProdDiff*Δ Population	0.000	0.002	0.000	0.001
	(0.001)	(0.001)	(0.000)	(0.001)
ProdDiff*Δ Access to port	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
ProdDiff*Rail Access	-0.000	-0.000	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Initial industry Share	-0.001***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Industry FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
No. of observations	8250	8250	8251	8251
No. of clusters	319	319	319	319

Note: the dependent variable in the two columns on the left is the change in employment. The dependent variable on the right is the change in number of firms. The coefficient of ProdDiff*Δ Market Access is negative and significant, which implies that industries with low product differentiation grew disproportionately faster in cities that gained large market access. All regressions include sector and city fixed effects. Standard errors are clustered at the city level and province-industry level. Standard errors in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

Appendix

A.1. China's National Trunk Highway System

The Chinese government has long held the belief that a highway system is important for developing regional economies, boosting efficiency of the logistics system and raising living standard. The highway construction plan aimed to promote trade, increase free flow of commodities, and subsequently raise competition in the domestic market. The government also wanted to ease regional inequalities and increase employment through connecting different parts of the country through highway. The rapid economic growth in China during the early years of the economic reform has resulted in an immense demand for an effective transportation system. The Chinese government wanted to solve the problems of heavy traffic and delay due to the backward infrastructure. To fund the construction project, the government established new regulations in 1984, imposing vehicle purchase tax and raising roadway tolls. In addition, it decided to make loans to build highways and repay the loans by charging highway tools.

According to the original plan, the new transportation system is 34422 kilometers long, of which 25765 kilometers are highway, 1479 kilometers are Rate I freeway, and 7178 kilometers are Rate II freeway. The three kinds of express way takes up 74.85%, 4.3% and 20.85% of the total distance respectively. The government planned to finish the construction around 2020. The plan also stated that the new highway system would link the capital Beijing to other provincial cities, connecting 93% of the major cities with other 1 million population and big cities with over 0.5 million population. The number of cities that are linked together will exceed 200, and 0.6 billion people will be affected by the highways. The government planned five vertical expressways and seven horizontal expressways to connect the country into a single network.

The plan was approved by the National People's Congress in 1992 and established by the Department of Transportation in June, 1993. The construction of the highway system can be divided into four time periods: the mid-1980s to 1991, 1992 to 1997, 1998 to 2003, 2003 to 2007. It started in the major cities in the mid-1980s. Since the approval of the construction plan in 1992, the project entered the regular phase. Until the end of 1997,

the total distance of highways in China was 4771 kilometers, of which 70% was the major national highway. In 1998, the government decided to build infrastructure and speed up the highway construction project in response to the financial crisis in Asia. From 1998, the highway construction project entered the rapid development phase, and in 2001, the total distance of highways in China exceeded 19,000 kilometers, the second longest in the world after the United States. From 2003 to 2007, the construction project was further accelerated and was finished in 2007.

A.2. Price Quote Data

Figure 7: Price Quotes from Deppon

Notes: this figure shows the websites of the two logistics companies. I collect all the price quotes from these two companies for any pair of cities in China. There are three transport modes that the two companies provide. The fastest mode comes with a higher price. I always choose the fastest mode to be consistent. I then compare the quotes for every city pair and choose the lower quote as the transportation cost between the city pair.

A.3. Revenue Productivity, Prices and Marginal Cost

If firms with high productivity tend to pass some of the efficiency advantage to consumers in the form of lower prices, then regressing revenue productivity will produce downward bias.[57]. The reallocation effect may also be dampened by changes in prices. I use production quantity data at the firm level from 2000 to 2006 to estimate firm marginal costs and prices for a subset of firms in the Annual Survey of Industrial Enterprises.[58] I follow Garcia and Voigtländer (2013) to decompose revenue productivity into markup, marginal costs and physical productivity:

$$TFPR_{it} = p_{it}A_{it} = \mu_{it} \cdot MC_{it}(A_{it}, w_{it}) \cdot A_{it} \tag{23}$$

where p_{it} is the output price that firm i charges, A_{it} is firm i's physical productivity, μ_{it} is markup, MC_{it} is firm i's marginal cost and w_{it} is the input price of firm i.

As we can see from the above expression, firms with higher physical productivity will have lower marginal cost, and the two effects offset each other. Assuming markups do not change, then the change in revenue productivity is ambiguous. The fact that I find positive significant effects on revenue productivity suggests that physical productivity might have responded even more to highway access. Also note that the input costs of firms are also very likely to respond to changes in market access. If inputs shipped from other cities becomes cheaper in respond to increase in market access, then revenue TFP will cause even larger downward bias. If labor inputs become more expensive after highway expansion and the effect dominates the cheaper prices of other inputs, then revenue TFP will give us smaller biases.[59]

---

[57]Smeets and Warzynski (2013) construct a firm level price index to deflate revenue productivity and show that this correction yields larger international trade premia in a panel of Danish manufacturers. Garcia and Voigtländer (2013) show within-plant physical productivity growth after exporting, in contrast to most previous studies that use revenue productivity

[58]The firm-level quantity data is used by the National Bureau of Statistics in China to estimate the total quantity of production of major manufacturing products. Though the data does not cover all firms in the ASIE, it is meant to capture all median and large firms.

[59]Theoretically, if the effect of labor input is sufficiently large, it is possible that the effect will exactly offset or even dominate physical productivity.

A.4. Bilateral Trade and Market Access

I derive an alternative measure of market access that is very similar to Redding and Venables (2005) and Donaldson and Hornbeck (2015) with the model in Section 3. For now, I will consider a one-sector version of the model to simplify derivations. An extension to the multi-sector version would be an interesting exercise but requires more theoretical and data work.[60] Bilateral trade:

$$X_{ij}^s = \beta_s \times \frac{Y_i \times Y_j}{Y} \times \left(\frac{w_i \tau_{ij}^s}{\Theta_j^s} \right)^{-\gamma_s} \times \left(f_{ij}^s \right)^{-[\gamma/(\sigma_s-1)-1]}. \tag{24}$$

where $\Theta_j^{-\gamma} \equiv \sum_{k=1}^N \left(\frac{Y_k}{Y} \right) \times (w_k \tau_{kj})^{-\gamma} \times f_{kj}^{-[\gamma/(\sigma-1)-1]}$.

Similar to Redding and Venables (2005) and Donaldson and Hornbeck (2015), I define Consumer Market Access (CMA) of city j to be

$$\Theta_j^{-\gamma} \equiv CMA_j. \tag{25}$$

Rewrite (9)

$$X_{ij} = \frac{Y_i}{Y} (w_i \tau_{ij})^{-\gamma} CMA_j^{-1} Y_j. \tag{26}$$

Similarly, Firm Market Access (FMA) of city j is defined as:

$$FMA_i \equiv \sum_j \tau_{ij}^{-\gamma} CMA_j^{-1} Y_j. \tag{27}$$

Note we can also write CMA_j as:

$$CMA_j = \sum_i \tau_{ij}^{-\gamma} FMA_i^{-1} Y_i. \tag{28}$$

Donaldson and Hornbeck (2015) shows that $FMA_i = \rho CMA_i$ for some scalar $\rho > 0$. If

---

[60]Using China's Input-Output tables for 1997, 2002 and 2007, I am constructing industry-specific input supply access and output demand access.
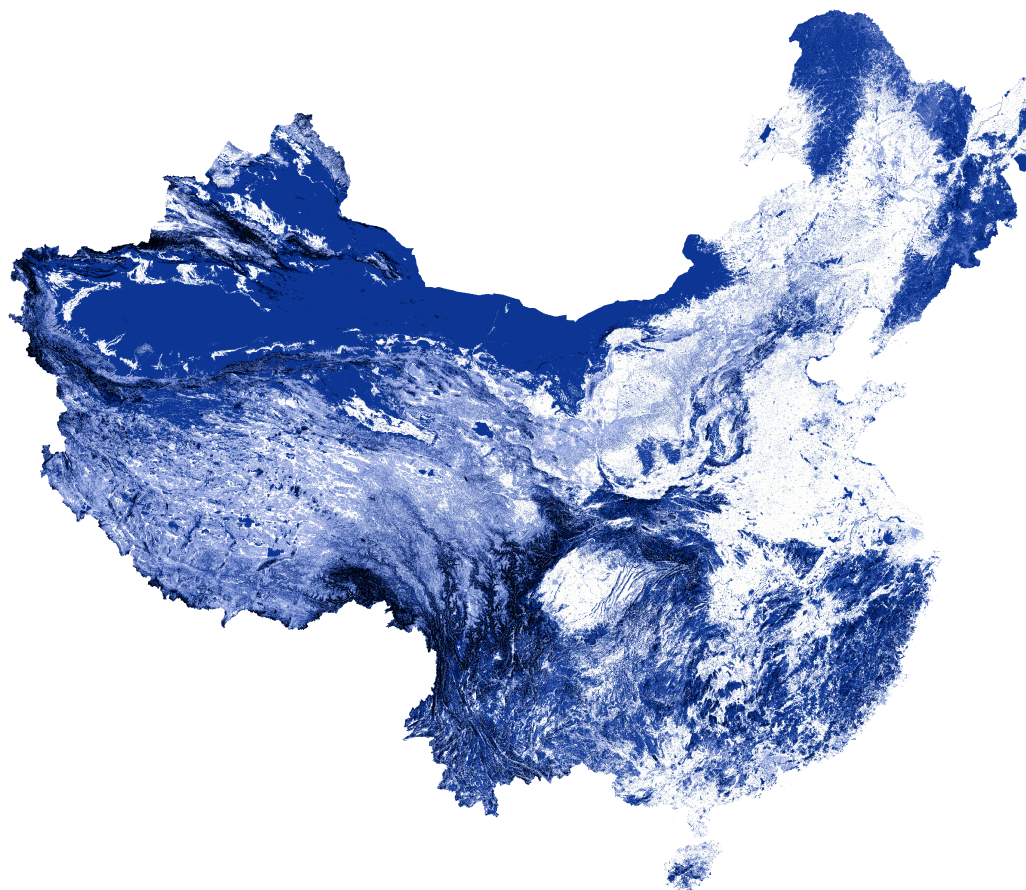
we let $MA_i \equiv FMA_i = \rho CMA_i$ equation (12) implies that:

$$MA_i = \kappa \sum_j \tau_{ij}^{-\gamma} MA_j^{\frac{-(1+\gamma)}{\gamma}} Y_j.$$ (29)

From the expression above, we see that a city's market access is the weighted sum of all its neighboring cities' market access. To solve the market access measure, I solve a system of 339 equations with MatLab.

A.5. Highway Construction Costs in China

Figure 8: Estimated Construction Costs

Notes: this figure presents the estimated construction costs for each pixel. I follow Faber (2014) to construct the construction costs using remote sensing data on terrain ruggedness. The darker the pixel is, the higher the construction costs.