Market Access and Firm Performance: Evidence from China*

Guojun HE† Xing HU‡ Shengyu LI§ Hongsong ZHANG¶

December 17, 2024

Abstract

Distortions within and across regions, arising from natural conditions or institutional frictions, can hinder economic growth. This paper examines how improved market access – induced by transportation network development – alleviates these distortions and impacts firm performance. We emphasize the importance of input-output networks, as expanding transportation infrastructure into regions with richer trade links brings larger shifts in market accessibility. Our findings reveal that improved market access to upstream suppliers and downstream buyers significantly boosts firm productivity and markups while reducing input prices. However, these gains are partly offset by intensified within-industry competition resulting from enhanced transportation networks. Overall, the expansion of China's expressways between 1998 and 2007 increased firm productivity by approximately 40%, raised markups by 1.7%, and reduced input prices by 17%. Moreover, firms initially at a disadvantage benefited disproportionately more from the expressway improvements, fostering convergence in firm performance within and across regions. This study provides robust evidence that reduced regional distortions through enhanced market access is a critical mechanism for firm-level gains.

Keywords: Distortions, market access, firm performance, productivity, input price, markup

*The authors thank Xiaodong Zhu, Yifan Zhang, Heiwai Tang, Bingjing Li, Chang Sun, and Jiawei Mo for the insightful comments. The authors also benefited from the discussions in the 2024 Firm and Industry Dynamics Workshop at the University of Hong Kong, the International Trade and Empirical Industrial Organization Workshop at Peking University, and the 2024 Glasgow-Nankai International Workshop on Trade, Productivity and Development, as well as the seminar participants at the University of Hong Kong, Peking University, Renmin University of China, Zhejiang University, and Wuhan University. All errors are the authors' responsibility.

†HKU Business School, The University of Hong Kong, gjhe@hku.hk

‡HKU Business School, The University of Hong Kong, huxing@connect.hku.hk

§School of Economics & Centre for Applied Economic Research, The University of New South Wales, shengyu.li@unsw.edu.au

¶HKU Business School, The University of Hong Kong, hszhang@hku.hk

1 Introduction

Distortions within and across regions, caused by natural conditions or institutional frictions, may hinder economic growth. Institutional frictions, in particular, may favor a particular group of firms while discriminating against others. The improved market access due to the improvement of transportation may reduce distortions, such as the misallocation in inputs (e.g., Wu et al., 2023; Hornbeck and Rotemberg, 2024) and government assistance (e.g., Restuccia and Rogerson, 2008; Aghion et al., 2015; Harris and Li, 2019). This paper investigates how market access induced by the development of transportation networks influences firm performance, emphasizing the role of market access in reducing distortions within and across regions.

The input-output linkages and the distribution of industrial clusters plays an important role in driving the impact of transportation networks. This is because extending transportation networks to a region with more and better trade partners generates larger changes in market access. We combine the input-output network with the transportation network to develop the production network-based market access. This results in upstream, downstream, and horizontal market access. While upstream and downstream market access measures the access to suppliers and buyers, respectively, due to the transportation network, horizontal market access measures the number of within-industry competitors confronted through the connected expressway network.

These measurements have two key features. First, input-output-network-based market access combines linkages in the transportation and production networks. Second, it captures the effects of the whole network, so the firm may be influenced even if the transportation improvement happens in a county that is not directly linked to the firm's location. These features expanding the dimensions of the effects of market access on firm performance.

We examine the impact of market access on firm performance and regional economic growth, using the large-scale, rapid changes in China's transportation network during 1998–2007 as a natural experiment. By the end of this period, the length of expressways had increased to six times its length ten years earlier, covering over half of the counties in China. We document two facts of expressways' influence on the firm performance and distortions. First, after being connected to the expressways network, the counties that initially rank at the top in the industry (in terms of total sales, average output price, and average raw markup) drop in ranking. In contrast, those initially bottom-ranked counties rise. A similar pattern is observed within the same county across firms: initially high-ranked firms are caught up or replaced by the initially low-rank firms after the expressway connection. Second, expressway connections reduce distortions, as measured by the dispersion of revenue-to-variable costs ratio (raw markup) and subsidy-sales ratio as the proxies for distortions at the county-industry-year level. These patterns motivate us to investigate the influence of market access on firm performance by reducing distortions as an essential channel.

To understand the impact of upstream, downstream, and horizontal market access, we estimate firm-level productivity, input price, and markup from a large dataset of Chinese manufacturing firms' production information collected by the National Bureau of Statistics.

We extend the method of Grieco et al. (2016) and Li and Zhang (2022) to allow for variable markups and derive estimates for firms' heterogeneous productivity, quality-controlled input price, and markup. The key idea is to infer the state variables (productivity, input price, market power, etc.) that are observable to entrepreneurs (but not to researchers) from the optimal choices and the final expenditures that the researchers can observe. This method enables us to investigate multiple channels through which the upstream/downstream market access and market competition influence the firms.

We estimate the effect of market access on firm performance using the variation arising from the development of transportation networks and industrial distribution across regions. Because the development of the expressway network is endogenous, we use the recentered market access as IV for market access as the identification strategy following Borusyak and Hull (2023). The idea is to simulate the counterfactual connections between counties and calculate the expected market access of each county. Then, by subtracting the predicted market access from our market access measurements, we derive the recentered instruments that remove the bias from non-random shock exposure.

We have three major findings. First, increased upstream market access decreases input prices and increases firms' markup, and better downstream market access increases both productivity and markup. The decrease in input prices may be attributed to lower transportation costs. And the enlarged access to the upstream markets brings more affordable suppliers to the firms (e.g., Gümüş et al., 2012). The decrease in the marginal cost may further drive up the markup. Meanwhile, as firms have more exposure to the downstream market, the rise in productivity may be facilitated by easier and faster exchanges of ideas through the movement of goods and people (e.g., Dong et al., 2020). Also, the firms are able to match more customers willing to pay higher prices. Thus the revenue-based productivity and markup increases.

Second, while more market competition has opposite impacts on the firms, they are dominated by the market access effects. In addition to more suppliers and customers, the development of expressway network also brings more competitors to the firms. The more fierce competition decreases firms' revenue-based productivity and markup and increases input prices. However, through a simple counter-factual analysis, we find that these negative effects are dominated by the benefits brought by the increased upstream and downstream market access. Overall, the development of the expressway network contributes to approximately 40% increase of firms' productivity, 17% decrease in firms' input price, and 1.7% increase in firms' markup. Moreover, the heterogeneous effects of the expressway network lead to the convergence of firms' performance. We find that those firms initially performing well in terms of productivity, input price, and markup benefit less from the expressway network development. At the aggregate level, these heterogeneous effects contribute to about 10%, 15%, and 1.5% decrease in the interquartile range of firms' productivity, input price, and markup, respectively.

Third, we test the mechanism through which the expressway network influences the firm performance. We use the county-industry-year-level dispersions of markup and subsidy-sales ratio as proxies for the distortions in the local market. The dispersion of markup reflects the input misallocation (e.g., Edmond et al., 2015, 2023; Wu et al., 2023), and the dispersion of subsidy-sales ratio reflects the misallocation in terms of the government assistance (e.g., Aghion et al., 2015; Harris and Li, 2019). We find that increases in the upstream and downstream market access reduce the distortions. These decreases in the distortions both benefit firms by reducing their input price and increasing their productivity and markup. We also find that the decreases in the distortions help to reduce the dispersion of firm performance, implying that the reduction of the distortion caused by the increased market access is an important driver of market integration.

This paper makes a contribution to the literature on estimating firm-level heterogeneity. Firms' heterogeneity in performance may come from many sources, such as productivity (e.g., Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015), input prices (e.g., De Loecker and Warzynski, 2012; Grieco et al., 2016), and markup (e.g., De Loecker et al., 2016). The paper extends the estimation procedures from Li and Zhang (2022) to allow for the firms' variable markup. Under the newly extended framework, we are able to estimate firm-level productivity, input price, and markup without output price/quantity data.

Moreover, this paper also contributes to the literature exploring the impacts of transportation networks on reducing distortions. Many studies have investigated the influence of transportation infrastructure on China's regional development (e.g., Baum-Snow, 2007; Faber, 2014; Lin, 2017; Qin, 2017; He et al., 2020; Banerjee et al., 2020; Baum-Snow et al., 2020; Egger et al., 2023; Li et al., 2024), and there are also papers utilizing the various methods to capture the impact of transportation network construction (e.g., Fogel, 1964; Redding and Venables, 2004; Hanson, 2005; Donaldson and Hornbeck, 2016; Donaldson, 2018; Jaworski and Kitchens, 2019; Balboni, 2024). Our paper is mostly closed to Wu et al. (2023) and Hornbeck and Rotemberg (2024). Using the variations in the provincial road length data, Wu et al. (2023) studies the (mostly direct) impacts of road construction on the distortions and finds that road expansion reduces within-industry markup dispersions. Hornbeck and Rotemberg (2024) exploits the market access measurement to study how railroad expansion influences American manufacturing industries. Exploiting the county-by-industry data in the manufacturing sector, they estimate the county productivity and find evidence of indirect impacts of market access on aggregate productivity through the channel of expansion of economic activity in the distorted counties. In this paper, we use the expressway Geographic Information System (GIS) data to construct the market access measurement and introduce the I-O linkage into the market access measurement to jointly take the production network (through the cross-industry upstream-downstream correlations) and transportation network into consideration. Using the firm-level data, we provide new micro evidence on how changes in market access influence firm performance through the channel of reducing distortions.

Furthermore, this paper provides firm-level evidence supporting economic theories regarding misallocation and market integration. We test the distortion in terms of both input allocation (e.g., Hsieh and Klenow, 2009; Edmond et al., 2015, 2023) and the government assistence (e.g. Restuccia and Rogerson, 2008; Aghion et al., 2015; Harris and Li, 2019). Our empirical findings are consistent with the classical economic theories such as the law of one price by Marshall (1890), the zone pricing system by Stigler (1949), the spatial price equilibrium by Samuelson (1952), and the concept of external economies by Krugman (1991), all of which emphasize the importance of transportation's impact on market dynamics between different regions. However, past empirical literature has primarily concentrated on country-level evidence in international markets (e.g., Parsley and Wei, 1996; Atkeson and Burstein, 2008)[1]. Evidence regarding domestic market integration is more commonly found in studies of agricultural goods (e.g., Badiane and Shively, 1998; Costinot and Donaldson, 2016). Our paper provides compelling firm-level evidence from the manufacturing sector, demonstrating that market integration may be facilitated through the channel of transportation networks' heterogeneous effects on firms' performance. These results pave a novel path for further exploring the dynamics of market integration.

This paper is organized in the following way. In Section 2, we introduce the backgrounds and illustrate the motivating facts. In Section 3, we take a first glance at the features of the Chinese expressway network and its relationship with market integration. Section 4 introduces how we construct the measurements for the upstream/downstream market access and the

[1] See Donaldson (2015) for a review.

market competition and summarizes the data we use for the estimation. Section 5 develops a framework to jointly estimate firm-level productivity, input price, and markup based on revenue and input expenditure data. Section 6 reports our estimates of the average and heterogeneous influences of transportation development on the firms' performance. Section 7 reports the impacts of market access on distortions and the impacts of distortions on firm performance. Finally, Section 8 makes the conclusion.

2 Motivation

2.1 Background

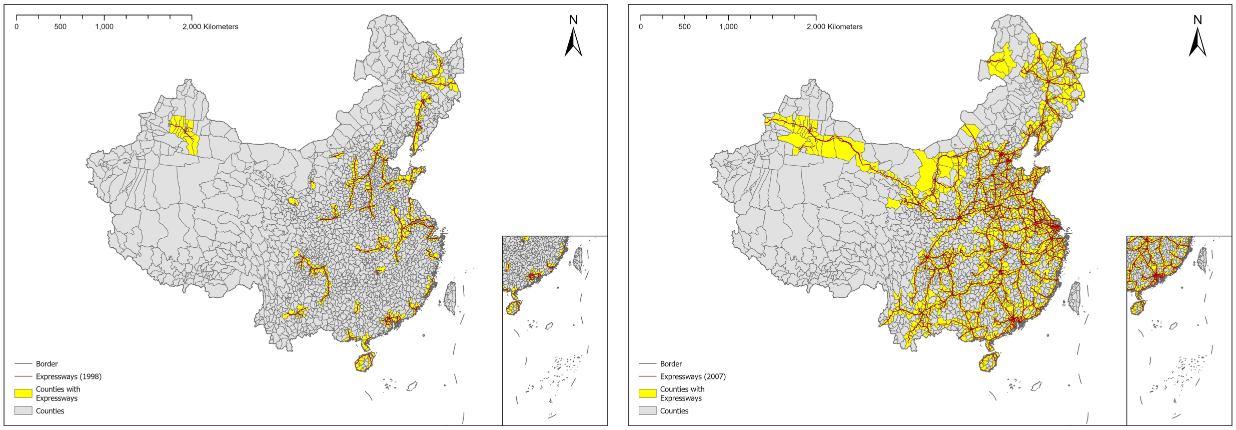
Over a considerable period, owing to its exceptional point-to-point capability, road transportation has maintained a dominant role in China's freight transportation modes. According to *China Traffic Statistics Yearbook*, from 1998 to 2007, the freight volume of motor vehicles contributed to about 60% of the gross freight volume of the years. And under the *Expressway Traffic Management Measures*[2] issued by the Chinese Ministry of Public Security and the *Expressway Engineering Construction Supervision Standards* issued by the Ministry of Transport in 1995, China's expressway systems have the conditions to provide high-quality (high speed, high carrying capacity) transportation. The extreme flexibility and high quality have made the expressway network a significant factor in influencing China's production network.

The reason why we focus on the time period is that in 2007, the State Council of China released a press conference confirming that the main trunk line of the "Five Verticals and Seven Horizontals" national main lines was basically completed by the end of the year. These 12 main lines are all high-grade roads at or above the second level, with expressways accounting for approximately 76% of the total mileage. They covered all mega cities with a population of over 1 million and 93% of large cities with a population of over 500 thousand in China[3]. The plan was implemented in 1993, with the period 1998–2003 to be regarded as the accelerated construction stage and the period 2003–2007 as the comprehensive construction stage. Figure 1 shows the expressways network' development from 1998 to 2007.

[2]This document set the prohibition of non-motorized vehicles, tractors, agricultural transport vehicles, and motor vehicles with speeds below 70 kilometers per hour from entering expressways.

[3]Specific illustration of the policy background can be found in Faber (2014).

Figure 1: China's Expressway Networks (1998 and 2007)



In Figure 1, the red lines refer to the expressways, the yellow polygons refer to the counties connected by the expressways, and the gray polygons refer to the unconnected counties. In 1998, the total length of expressways in China was 8.7 thousand kilometers, covering 639 counties (21.60%) nationwide. By 2007, China's expressways had reached 53.9 thousand kilometers, covering 1,567 counties (53.01%) nationwide.

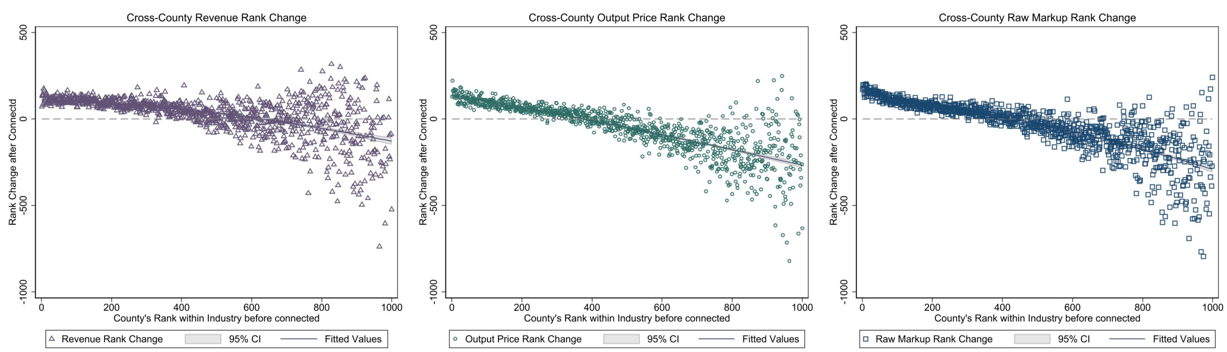
2.2 Motivational Facts

As China has experienced rapid development of the expressway network, massive counties and firms were connected. Did being connected to the expressways impact the firms in these counties? In this subsection, we show two motivating facts of the firms and counties before and after they are connected to the expressways. In the first fact, we find that after connecting the expressways, those counties initially occupying the high ranks in terms of aggregate performance (in terms of revenue, output price, and markup) were relatively worse off after connecting, while those initially low-rank counties were better off. In the second fact, we find the decline in the distortions, measured by the dispersions of subsidy-sales ratio and markup at the county-industry-year level, after connecting to the expressways.

2.2.1 Market Access and Relative Changes in Performance

We first find evidence of expressways' influence on the counties' aggregate performance. Making use of the dataset from the Chinese Annual Surveys of Industrial Enterprises (ASIE), we first aggregate firms' performance to the county-industry-year level, in terms of revenue, revenue-to-variable cost ratio (raw markup), and the output price (from the sub-sample of ASIE, which account for about 1/3 of the sample). Then, we calculate each county's rank in each industry, and plot the rank changes after connection with respect to the counties' ranks before connection in Figure 2. The x-axis represents the counties' ranks within each industry before they are connected to the expressways. The y-axis reports the rank changes after connection. If the rank change after connection is greater than 0, it indicates that the county's rank drops after connection. The results show that for those initially top-ranked counties, their ranks within the industry fall after they are connected to the expressways. In contrast, the ranks of those initially bottom-ranked counties increase, though the dispersion is larger for these counties.

Figure 2: Counties' Aggregate Performance Rank Change



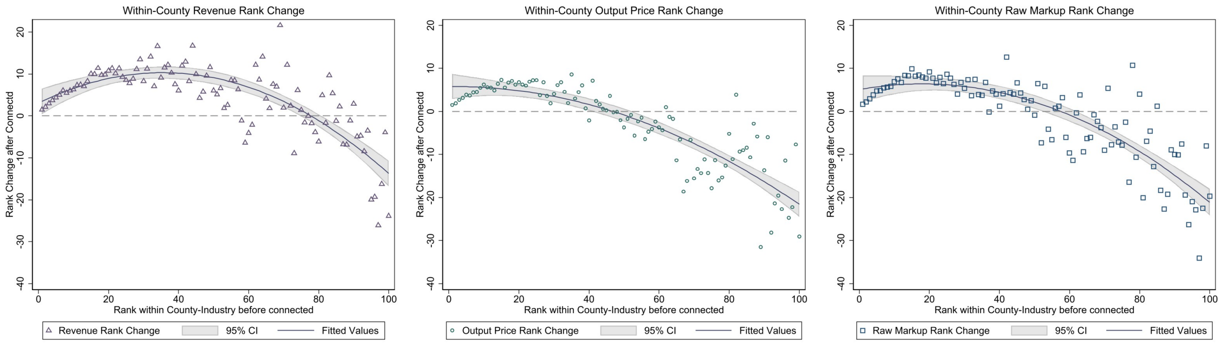
Note: Counties' revenue rank is based on the sum of the revenue conditional on each industry. Raw markup and output price are based on the sales-weighted average.

In Figure 3, we plot the ranks of firms' performance in the local market before and after their counties are connected to the expressways' network. Contradictory to the prediction generated by traditional literature (e.g., Melitz, 2003), the pattern surprisingly shows that the opening to the expressways network reshuffles firms' relative competitiveness in the local markets. After being connected to the expressway network, those high-performance firms seem caught up by those previously low-performance firms. Thus, the ranks of those high-rank firms dropped and were replaced by those low-rank firms, while those top firms still kept their premium. This indicates that the expressway's influence on the firms may not be homogeneous to all the firms in the same county and industry.

2.2.2 Decreases in Distortions

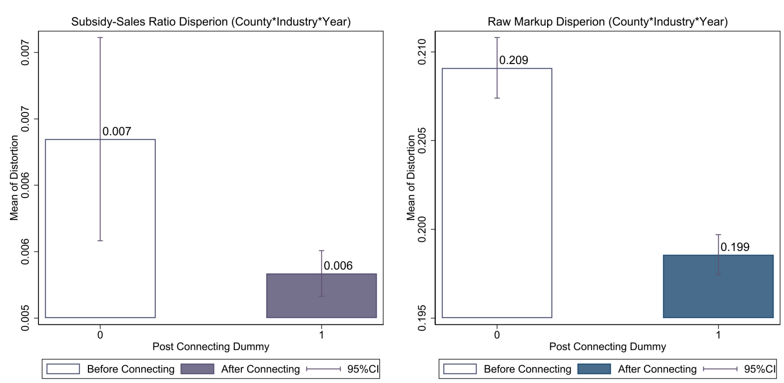
We cannot directly observe the distortions or allocation efficiency in each county. Given that the distortions may come from the government assistance misallocation (e.g., Restuccia and Rogerson, 2008; Aghion et al., 2015; Harris and Li, 2019) and resource misallocation (e.g.,

Figure 3: Firms' Performance Rank Change



Edmond et al., 2015, 2023; Wu et al., 2023; Hornbeck and Rotemberg, 2024), we take the dispersions of subsidy-sales ratio and (raw) markup at the county-industry-year level as the indicators for distortions. Figure 4 shows that connecting to the expressway network reduces the distortions in the region. Our findings are consistent with Wu et al. (2023), who use the provincial road length data as the measurement for China's road construction condition and find that the mass construction of roads significantly reduces markup dispersion within the same time period.

Figure 4: Distortions Change after Connection



Note: The distortions at the county-industry-year level are measured by the dispersions of firms' subsidy ratio and raw markup, which are both demeaned at the county-industry-year level when calculating the corresponding standard deviations. Both figures indicate that the distortions decrease after being connected.

3 Expressways Network

Exploiting the expressways geography information system data, we construct the minimal origin-destination distance between counties in China. Then we adopt the simple traditional measurements in network science (closeness centrality and mean geodesic distance) to take a glance at the features of Chinese expressways network development.

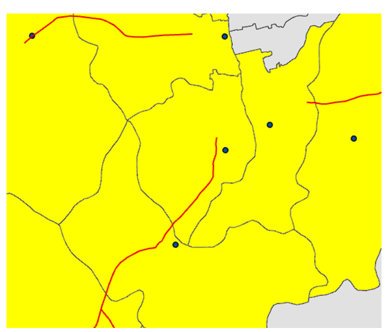
3.1 Expressways, Firms Locations, and County Entrants/Exits

As in He et al. (2020), we make use of the Historical geographic information systems (GIS) data on China's National Expressway Network of the years 1998, 2000, 2002, 2003, 2005, 2007[4]. The geographical locations of the firms in the counties are approximated by the longitude and latitude of the local governments' buildings in the counties. If there is no data on the local government building, the centroid of the polygonal data on the county map is used to represent the geographical location of the firm within the county[5]. Besides, the county entrants/exits on the expressways are approximated by the nearest points of the expressways within the counties to the local government buildings. The details about how we process the data is attached in the Appendix A.1 and A.2.

3.2 Minimal Origin-Destination Cost Distance

In order to capture the change in expressways network construction, we need to firstly calculate the minimal distance from counties to counties through expressways in China. Figure 5 displays the raw data sample with the county government locations (in blue dots), expressways across the counties (in red lines), and the shapes of counties connected to the expressways (in yellow polygons):

Figure 5: Expressways and Governments of Some Counties in Shanxi Province



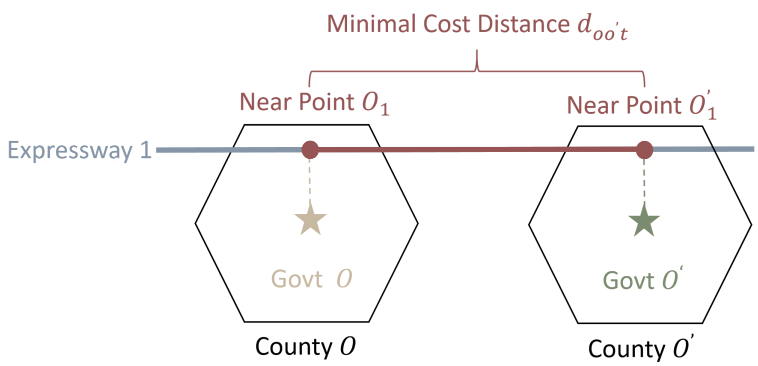
As shown in Figure 5, county governments are usually not exactly on the expressway. Thus,

[4]The data was collected from the China Administrative Spatio-Temporal Expressway Database (STED) from the Australian Centre of the Asian Spatial Information and Analysis Network (ACASIAN) Data Center, Griffith University. For details, please refer to He et al. (2020)

[5]Under this exercise, we potentially ignored the geographical differences of the firms within the county and adopted the local administrative center as an approximation of the firms' locations.

in order to generate the minimal OD cost Distance, we have to generate the nearest points on the expressways for the government of each county. The processing details are in Section A.2 in the appendix. These near-points can be interpreted as the proxy of the expressway entrance/exit in the county. We further assume that the counties can only be reached by expressways. Therefore, the minimum Origin-Destination (OD) Cost Distance between counties is defined as the shortest distance between the approximate entry and exit points within each county along the expressway network, which is as shown in Figure 6.

Figure 6: Minimal OD Cost Distance between Counties



3.3 Closeness Centrality and Mean Geodesic Distance

Given the complexity of expressway networks, which may contain multiple components, we follow the approach of Newman (2018) to define the Closeness Centrality (CC) of county o at time t. Closeness is conceptualized as the harmonic mean of the distances between counties. The formula for calculating the CC of a county is as follows:

$$CC_{ot} = \frac{\sum_{o' \in \mathcal{C}} \delta_{oo't}}{N}, \tag{1}$$

where $\mathcal{C}$ is the set of all counties, including those are not connected by the expressway, in China, and the total county number is a constant, denoted by $N = 2,956$. $\delta_{oo't}$ refers to the transportation weight of the destination county $o' \in \mathcal{C}$ for the origin county o, and it is calculated by:

$$\delta_{oo't} = \begin{cases} \frac{1}{\sqrt{S_o}/2} & \text{if } o' = o, \\ \frac{1}{d_{o'ot} + \left(\sqrt{S_o} + \sqrt{S_{o'}}\right)/2} & \text{if } o' \neq o, \end{cases} \tag{2}$$

where S_o and $S_{o'}$ represent for the area of counties o and o', respectively. $d_{o'ot}$ refers the minimum distance between counties as detailed in the preceding section.[6] If two counties are not connected by a road, the distance between them is considered infinite. And if a county is not connected with other counties, we assume they only have the access to the local market.[7] In this way, we treat all counties in China as part of the expressways network, regardless of whether they are connected to an expressway. We firstly construct $\delta_{oo't}$ for all the 2,956 counties in years 1998, 2000, 2002, 2003, 2005, and 2007. Then we use the linear interpolation to generate the $\delta_{oo't}$ for all the O-D pairs in all the years in 1998–2007. Eventually, we can calculate the CC_{ot} for all the counties in the expressways network in years 1998–2007. Since a county's area does not change over time (we use the map in 2000 for all the counties), improvements to this network are indicated by an increasing number of counties being connected by expressways or a reduction in the distances between counties.

We adopt this measure from network science because it captures the comprehensive effects of expressway network development on firms' accessibility. This includes changes in the scope of areas firms can reach via expressways, the extent of distances firms have to suffer to access other areas through expressways, and the influence of expressway development in other segments of the network that are not be directly linked to the firms. All these features help us to capture the development of expressways network.

Another reason for us to adopt the closeness centrality is that it naturally introduces the Mean Geodesic Distance (MGD) as a topological index for measuring the average cost (in terms of distance) for travelling between vertices within a network (e.g., Freiria et al., 2015; Newman, 2018; Ji et al., 2022). We can define the MGD for province p in year t, denoted by τ_{pt}, to measure the average distance for the counties of province p to travel both within and outside the province:

$$\tau_{pt} = \frac{N_p}{\sum_{o \in \mathcal{C}_p} CC_{ot}}, \tag{3}$$

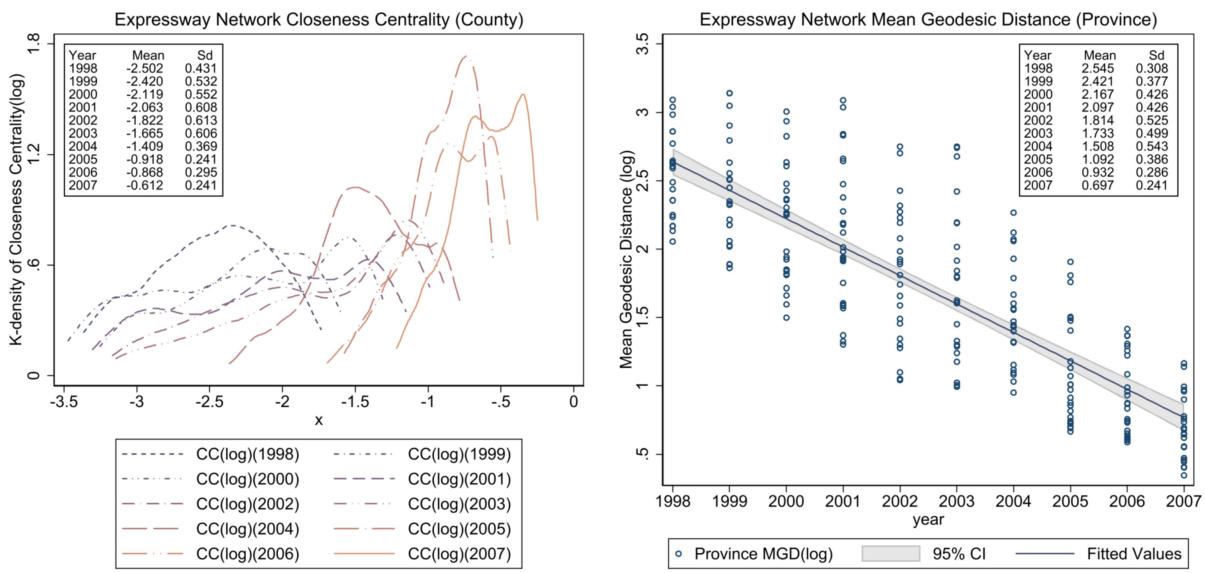
where CC_{ot} is the closeness centrality we just defined, and N_p referes to the number of counties in the province p. This measurement could be interpreted as an indicator for (average) transportation cost within an aggregated area, enable us to refer to various trade theories (e.g. Tinbergen, 1962; Anderson, 1979; Krugman et al., 1980; Eaton and Kortum,

[6]The distances we calculate are precise to meters. However, for the purposes of constructing a measure for expressway accessibility, we use 1,000 kilometers as the unit to prevent the measure from becoming too small.

[7]We are aware of that this assumption gives us a lower bound of the estimates on the gain from expressway connection.

2002; Arkolakis et al., 2012).

Figure 7: Closeness Centrality Dispersion & Mean Geodesic Distance Trend



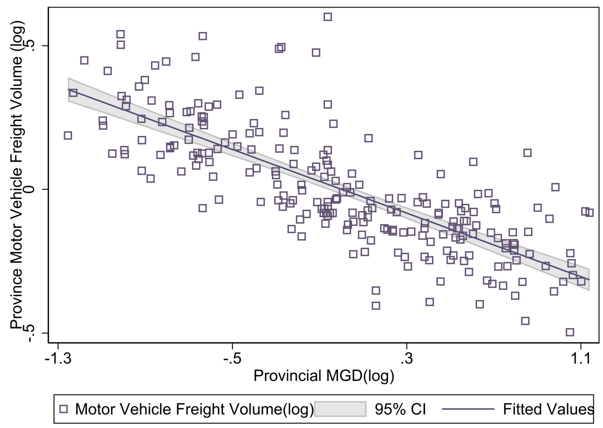
Note: Only counties/provinces with expressways are kept, with the top&bottom 10% trimmed. The dispersion for all the counties is reported in Figure B1 in Appendix.

The left panel of Figure 7 shows that the dispersion of Chinese counties' closeness centrality shrinks, and the mean shifted to the right. This indicates that for those counties connected to the expressways, their centrality increased, and the gap decreased. The right panel of Figure 7 illustrates the temporal dynamics of the Mean Geodesic Distance (MGD) within the network. Since MGD serves as a proxy for the overall transportation costs within the network, a decreasing trend in MGD over time would indicate that the expressway network's efficiency is improving. This, in turn, suggests that the network is enabling quicker and more cost-effective movement of goods and people, indicating the potential economic growth among the various regions of China. Meanwhile, the difference in the transportation conditions across provinces also decreases.

3.4 Expressways and Freight Volume

The measurements for the expressways network enable us to test the effect of the improvement of the transportation situation on the actual freight volume of the province in that year. This analysis is to take a first glance at how transportation improvements affect the regional economies (e.g., Allen and Arkolakis, 2014). We collect the province-level data of freight volume of automobiles and other motor vehicles and the mileage of ranked roads, including expressways, from the *China Traffic Statistics Yearbook*. Figure 8 plots the relationship between the gross motor vehicle freight volumes and the MGD of the provinces in different years, which shows an obvious negative correlations.

Figure 8: Provincial Motor Vehicle Freight Volume & MGD



Note: Normalized by the province mean and dropped the year 1998 due to the freight volume data quality.

To exclude the influence of other road transportation ways, it may be more appropriate for us to look at the change of expressways freight volume. Due to the lack of specific data on the actual carrying capacity of expressways, we approximate the carrying capacity of expressways in the province for that year by multiplying the total carrying capacity of motor vehicles by the ratio of the mileage of expressways to all levels of expressways in the province[8].

In Table 1, we report the effects of changes in the provincial MGD on the provincial freight volume. Column (1) presents the results while controlling for province fixed effects and year fixed effects. Considering the freight volume is estimated, we include an additional control for the length of other types of roads. The two sets of results are closed to each other, indicating that a decrease of 1% in the MGD, which can be interpreted as an improvement in the province's expressway network efficiency, is associated with an approximate 0.459% increase in freight volume, which is consistent with the trade literature that examines the relationship between distance and trade volumes across regions (e.g. Tinbergen, 1962; Anderson, 1979; Krugman et al., 1980; Eaton and Kortum, 2002; Arkolakis et al., 2012). This effect may be driven by both the extensive margin and the intensive margin. On the one

[8]Our exercise here is potentially under the assumption that expressways and other levels of expressways have equal carrying capacity per kilometer. Considering the high quality of expressway kilometers, our estimate is actually a lower-bound approximation of the fright volume of expressways.

Table 1: Province-level Expressway Freight Volume & MGD

	Expressway Freight Volume(log)	
	(1)	(2)
τ_{pt} (log)	-0.382***	-0.459***
	(0.1467)	(0.1365)
Length of Other Roads(log)		YES
Fixed Effects	YES	YES
Observations	226	226
Adjusted R^2	0.836	0.854

Note 1: Dropped the year 1998 due to the freight volume data quality.

Note 2: Fixed effects include province fixed effect and year fixed effect.

Note 3: Expressway freight volume is approximated.

Note 4: Standard errors (clustered at province-year level) in parentheses.

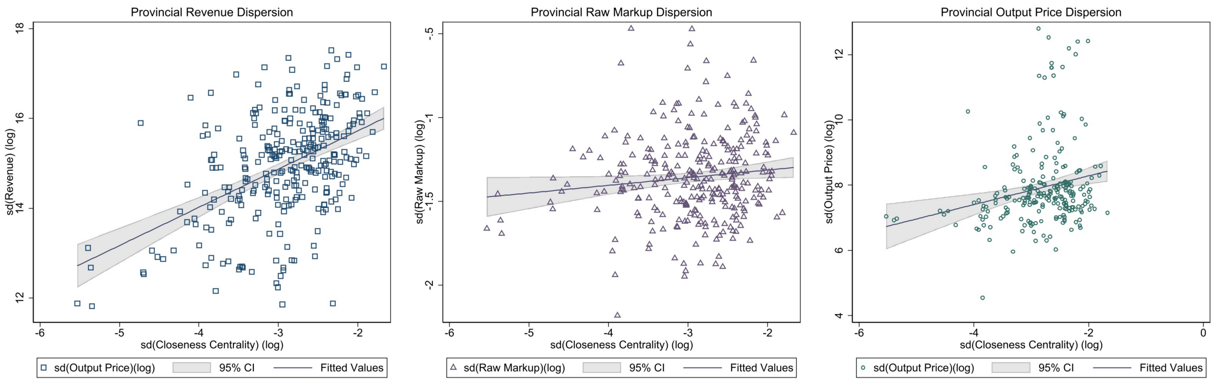
* $p < .10$, ** $p < .05$, *** $p < .01$.

hand, the high-quality traffic conditions, determined by such high construction standards and management level, provided by expressways may induce firms that previously relied on other modes of transportation to switch to expressways transportation. On the other hand, expressways can increase transportation speeds, which means that businesses that were previously restricted by road transportation, such as those with strict shelf life requirements or those that transport large materials and have specific road turning radius requirements, can now engage in long-distance transportation via expressways.

3.5 Expressways Network and Market Integration

We have seen in the previous data features that the dispersion of CC is decreasing for counties that are connected to the expressways network. The dispersion of CC within a province can be used as a measure of traffic friction in that province. In Figure 9, we report the relationship between the level of traffic friction in a province and the degree of market integration, measured by the dispersions of firms' revenue, raw markup, and output price within the province. The data patterns show that as the transportation friction of the province decreases, the market of the province becomes more integrated, which is a potential outcome of the decreased distortions in the local markets (e.g., Marshall, 1890; Stigler, 1949; Samuelson, 1952).

Figure 9: Provincial CC Dispersion & Within-Province Market Integration



Note: Firm indices are normalized by province&industry means. CC is normalized by the province mean.

4 Indices and Data

Although closeness centrality provides us with a criterion for evaluating the centrality of counties in the expressways network, it is still rough to evaluate firms' benefits. Firstly, closeness centrality can only capture changes in the pure transportation network but ignore the economic importance of different destinations. One solution is to construct indicators for market access (e.g. Donaldson and Hornbeck, 2016). This indicator can use the economic data of counties to weigh the economic significance of different destinations connected to an origin county. However, it still cannot fully capture the importance of each connected county for firms in the production supply chain. For firms belonging to different industries within the same county, the economic impact of connecting to another country is different. This is because the distribution of industries is not uniform among different counties. If a connected county has an industrial cluster of upstream firms that are very important for a certain industry in the origin county, then the impact of this connection on that industry is obviously greater than the impact on other industries within the same county. So we used an Input-Output (IO) table to construct the upstream and downstream market access indicators for firms belonging to different county-industry-year combinations.

4.1 Market Access and Market Competition

Define the upstream market access, denoted by MA^U_{okt}, and the downstream market access, denoted by MA^D_{okt}, for firms belongs to (sic-3-digit) industry k in county o as:

$$MA^U_{okt} = \frac{\sum_{o'\in\mathcal{C}}\left[\delta_{oo't}\left(\sum_{k'\in\mathcal{K}} R_{o'k',t=1998}\phi^U_{kk'}\right)\right]}{N},$$
$$MA^D_{okt} = \frac{\sum_{o'\in\mathcal{C}}\left[\delta_{oo't}\left(\sum_{k'\in\mathcal{K}} R_{o'k',t=1998}\phi^D_{kk'}\right)\right]}{N},$$

where $R_{o'k',t=1998}$ is the total sales of industry k' in county o' in 1998. $\phi^U_{kk'}$ and $\phi^D_{kk'}$ are the upstream and downstream coefficients between industries k and k' calculated from the I-O table in 2007. We fix the sales and IO-correlation in the certain years to avoid the endogeneity (mostly simultaneity) problem.

Similarly, we can define the market competition (MC), denoted by MA^H_{okt}, as:

$$MA^H_{okt} = \frac{\sum_{o'\in\mathcal{J}}\left(\delta_{oo't}N^F_{o'k,t=1998}\right)}{N}, \tag{4}$$

where $N^F_{o'k,t=1998}$ refers to the number of firms belonging to the same (sic-4-digit) industry in county o' in year 1998. We use this index to capture the change in competition intensity faced by the firms within the same industry. In Figure 10, we plot the dispersion of the market access and market competition. From the first two panels, we can find that both the upstream and downstream market accesses become less dispersed. However, the market competition, as shown in the third panel, becomes more dispersed. These patterns indicate that as being connected to more markets brings more potential suppliers and customers, more competitors also show up and harm firms' profit, making the influence of transportation construction on firms performance not that apparent. Thus the distinguishing of the upstream/downstream market accesses and market competition is necessary.

Since our indices are additive, we can make a decomposition to see how much effects are generated from nearby, medium, and remote markets. From Table 2, we can see that about 70% effects, no matter in terms of market access or market competition, are generated by the nearby markets, and the medium and remote markets are almost equally important.

Table 2: Indices Decomposition

	(1) Nearby Markets (<300km)	(2) Medium-distance Markets ($\geq$300km, $\leq$1000km)	(3) Remote Markets (>1000km)
MA_{okt}^{U}	68.136%	17.835%	14.029%
MA_{okt}^{H}	60.033%	22.753%	17.214%
MA_{okt}^{D}	68.087%	17.512%	14.401%

4.2 Data Summary

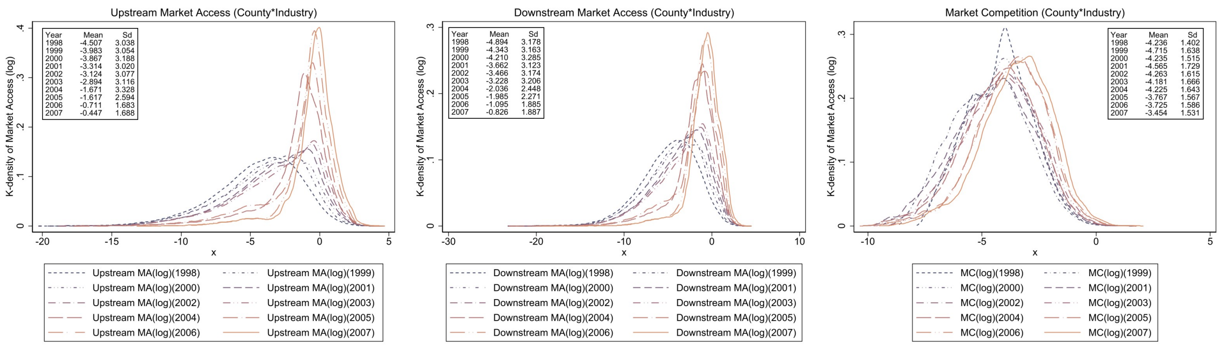
The dataset used in this study to analyze firm performance was sourced from the Chinese Annual Surveys of Industrial Enterprises (ASIE), collected annually by the National Bureau of Statistics (NBS) of China. This dataset encompasses production and sales information from 1998 to 2007 for both non-state-owned firms with annual revenues exceeding 5 million RMB (approximately 600 thousand USD) and all state-owned enterprises. It provides detailed firm-level information, including total sales, labor employed, wage expenditure, intermediate input expenditure, capital stock, etc. However, it does not include specific information on the prices or quantities of materials used. The summary statistics for the main variables we focus on are displayed in Table 3, where there are 406,101 firms, among which 78.32% are connected to the expressway network by 2007.

Table 3: Summary Statistics

Statistics	Median	Mean	SD	IQR	IDR
Total Sales (million USD)	2.271	8.880	75.128	4.383	12.979
Intermediate Input Expenditure (million USD)	1.703	6.704	58.204	3.276	9.720
Capital Stock (million USD)	0.497	3.132	35.706	1.295	4.395
Wage Expenditure (million USD)	0.143	0.434	2.472	0.258	0.732
Labor Employed	110	263.829	931.575	184	479
Intermediate Input Share over TVC	0.923	0.900	0.087	0.090	0.182
Counties with Expressways by 2007	1,567 (53.01%)				
Total Number of Counties	2,956				
Firms connected by Expressways by 2007	318,059 (78.32%)				
Total Number of Firms	406,101				
Observations with Expressways	406,335 (28.31%)				
Observations	1,435,295				

Note: All monetary values here are in millions of U.S. dollars in year 2000.

Figure 10: Upstream/Downstream Market Access & Market Competition



5 Estimation Method

To study the impact of the increase of county expressway network centrality on the firms' performance, we extend the estimation procedures as in Li and Zhang (2022) by allowing for variable markup. Under our newly extended framework, heterogeneity of firms' performance mainly comes from three components: input prices, productivity, and markup. This is because firstly, firms keep different input market power and face different market friction (e.g., localized market, transportation cost, supply network, and firm size), thus the input prices could be very different. Secondly, due to different production efficiency levels and the demand shocks (e.g., product appeal, customer network, and market size), an external manifestation is that the firms may have different levels of revenue-based total factor productivity (TFPR). Thirdly, firms may also keep heterogeneous output market power, causing the markup to vary across firms. Our extended estimation procedures aim to incorporate the above three sources of firms' heterogeneity and estimate the three key measures of firms' performance: TFPR, input price, and markup. Our method relies on the mappings from firms' heterogeneity, which is observed by the firms but unobservable to the researchers, to the input and output choices, which can be observed by the researchers[9], through firms' profit maximization.

5.1 Setup

Assume that a firm o produces output, denoted by Q_{jt}, with the output quality Φ_{jt}, which is affected by firm's productivity Ω_{jt}[10] and the (log) input quality h_{jt} in year t. Assume firms

[9] Except for the output quantity/price, which is only observed for 1/3 of our sample, we use them as important evidence in the main results but don't exploit them in the estimation.

[10] Since we only observe the revenue output rather than the quantity output in the ASIE dataset, it is still the revenue-based total factor productivity (TFPR).

log productivity $\omega_{jt} \equiv \ln \Omega_{jt}$ evolves according to an AR(1) process:

$$\omega_{jt+1} = f_0 + f_1 \omega_{jt} + \epsilon^{\omega}_{jt+1}, \tag{5}$$

where ϵ^{ω}_{jt} is the current innovation shock on firm j's productivity and is assumed to be independently and identically distributed (i.i.d.) across firms and time. The market is monopolistic competitive and the firm's output price P_{jt} is determined by:

$$P_{jt} = (Q_{jt})^{1/\eta_{jt}}, \tag{6}$$

where η_{jt} is the demand elasticity, and markup μ_{jt} is determined by $\eta_{jt}/(1+\eta_{jt})$. Assuming that a firm's market power evolves as a persistent process, the markup, which measures this market power, follows:

$$\ln \mu_{jt} = b_0 + b_1 \ln \mu_{jt-1} + \epsilon^{\mu}_{jt} \tag{7}$$

The output with quality ($\tilde{Q}_{jt} \equiv \Phi_{jt} Q_{jt}$) is produced through a CES production function by using the input of labor L_{jt}, intermediate input M_{jt}, and capital K_{jt}, with the distribution parameters α_L, α_M, and α_K, respectively[11]:

$$\tilde{Q}_{jt} = \tilde{\Omega}_{jt} \left(\alpha_L L^{\gamma}_{jt} + \alpha_M M^{\gamma}_{jt} + \alpha_K K^{\gamma}_{jt} \right)^{1/\gamma}, \tag{8}$$

where $\gamma \equiv (\sigma - 1)/\sigma$, and σ is the elasticity of substitution among the inputs. $\tilde{\Omega}_{jt}$ is the Hicks neutral TFPR which contains the information of firms productivity ω_{jt} and input quality h_{jt}. Following the same spirit of De Loecker et al. (2016), we further assume that $\tilde{\Omega}_{jt}$ is determined by the the following function:

$$\tilde{\Omega}_{jt} = \left(e^{\theta \omega_{jt}} + e^{\theta h_{jt}} \right)^{\frac{1}{\theta}}, \tag{9}$$

where $\theta \neq 0$, and $1/(1-\theta)$ refers to the elasticity of substitution between the firm's productivity and the input quality. We assume the (input-quality-adjusted and logged) input price p^M_{jt} also follows an AR(1) process:

$$p^M_{j,t} = \beta^M_0 + \beta^M_1 p^M_{jt-1} + \epsilon^M_{jt}, \tag{10}$$

where ϵ^M_{jt} is the current i.i.d. shock on firm o's input price, and the quality-involved input

[11] By normalisation, the sum of α_L, α_M, and α_K is one.

price $\tilde{P}_{jt}^{M} = exp(p_{jt}^{M} + h_{jt})$.

Basing on the observation of firm productivity (ω_{jt}), capital (K_{jt}), and the price of intermediate input (p_{jt}^{M}) and labor (P_{jt}^{L}), firm o chooses the optimal output ($\tilde{Q}_{jt}$), labor (L_{jt}), intermediate input quantity (M_{jt}), and intermediate input quality (h_{jt}) by profit maximization:

$$\max_{\tilde{Q}_{jt}, L_{jt}, M_{jt}, h_{jt}} P_{jt}\tilde{Q}_{jt} - \tilde{P}_{jt}^{M} M_{jt} - P_{jt}^{L} L_{jt}. \tag{11}$$

By making use of the first order conditions w.r.t. the output, the labor, and the intermediate input quantity form the above profit maximization problem, we can derive the function to be estimated as:

$$R_{jt} = \frac{\eta_{jt}}{1+\eta_{jt}} \left[E_{jt}^{M} + E_{jt}^{L} \left(1 + \frac{\alpha_K}{\alpha_L} \left(\frac{K_{jt}}{L_{jt}} \right)^{\frac{\sigma-1}{\sigma}} \right) \right] e^{u_{jt}}, \tag{12}$$

where R_{jt} is the revenue, E_{jt}^{M} and E_{jt}^{L} are the expenditures on intermediate input and labor, and u_{jt} is the i.i.d. measurement error. Before doing the first-stage estimation, we first use the NLLS to estimate the u_{jt} by exploiting the firms' selling expenses as the proxy variable (and other control variables) to control for the unknown μ_{jt}. After that, we can get rid of the measurement error term to derive the measurement-error-excluded revenue $\hat{R}_{jt}$ to be used in the estimation afterward.

5.2 Estimation Procedures

In the first stage, we use revenue production function, i.e., equation (12), and the first order condition of labor and material to recover markup (μ_{jt}) and quality-inclusive measures $\left(\tilde{\Omega}_{jt}, \tilde{P}_{Mjt}\right)$:

$$\begin{aligned}
\tilde{P}_{Mjt} &= \left[\frac{\alpha_M}{\alpha_L}\right]^{\frac{1}{\gamma}} \left[\frac{E_{Mjt}}{E_{Ljt}}\right]^{1-\frac{1}{\gamma}} P_{Ljt}, \\
\tilde{\Omega}_{jt} &= \frac{\mu_{jt}}{\alpha_L} L_{jt}^{-\gamma} E_{L_{jt}} \left[\alpha_L L_{jt}^{\gamma} \left(1 + \frac{E_{M_{jt}}}{E_{L_{jt}}}\right) + \alpha_K K_{jt}^{\gamma}\right]^{1-\frac{1}{\gamma\mu_{jt}}}, \\
\mu_{jt} &= \hat{R}_{jt} / \left[E_{M_{jt}} + E_{L_{jt}} \left(1 + \frac{\alpha_K}{\alpha_L}\left(\frac{K_{jt}}{L_{jt}}\right)^{\gamma}\right)\right].
\end{aligned} \tag{13}$$

By extending Li and Zhang (2022)'s method to allow for the variable markup, we substitute μ_{jt} into the AR(1) process to construct moment conditions to estimate the production parameters $(\alpha_L, \alpha_M, \alpha_K, \sigma)$.

In the second stage, we account for the correlation between quality-inclusive firm capability $\tilde{P}_{Mjt}$ and quality-inclusive input price $\tilde{\Omega}_{jt}$ through firms' endogenous choice of input quality, which can be inverted to recover quality-adjusted productivity and quality-adjusted input price. The first order condition of input quality implies that input quality is a monotone function of productivity (in logs):

$$h_{jt} = \frac{1}{\theta} \ln \frac{\sigma_{Mjt}}{1 - \sigma_{Mjt}} + \omega_{jt}, \tag{14}$$

where $\sigma_{Mjt} = (\partial F(\cdot)/\partial M_{jt}) \cdot (M_{jt}/F(\cdot))$ is the output elasticity of material, with $F(\cdot)$ to be the production function. Then we can substitute equation (14) into the capability function and input price to recover:

$$\begin{aligned} \omega_{jt} &= \tilde{\omega}_{jt} + \frac{1}{\theta} \ln \left(1 - \sigma_{Mjt}\right), \\ p_{Mjt} &= \tilde{p}_{Mjt} - \tilde{\omega}_{jt} - \frac{1}{\theta} \ln \left(\sigma_{Mjt}\right). \end{aligned} \tag{15}$$

The estimation strictly follows the second stage of Li and Zhang (2022), and we can get the estimate of θ, with σ_{Mjt}, $\tilde{\omega}_{jt}$, and $\tilde{p}_{Mjt}$ computed from data and the first stage, using AR(1) processes of ω_{jt} and p_{Mjt} to construct moment conditions a la Olley and Pakes (1996).

6 Empirical Results

After deriving the estimates of the (log) productivity ω_{jt}, the (log) input-quality-adjusted input price p^M_{jt}, and the (log) markup $ln\left(\mu_{jt}\right)$, we exploit the market access and market competition measurements to study the development of transportation on firms' performance.

6.1 Benchmark Results

Our benchmark estimations are based on the following specifications:

$$\ln FP_{jt} = \alpha_0 + \alpha_{up} \ln MA^U_{jkt} + \alpha_{hori} \ln MA^H_{jkt} + \alpha_{down} \ln MA^D_{jkt} + \alpha^X \mathbf{X}_{jt} + \xi^{FP}_{jt}, \tag{16}$$

where $\ln FP_{jt}$ refers to the log productivity ω_{jt}, the log)input-quality-adjusted input price p^M_{jt}, and the log markup $ln\left(\mu_{jt}\right)$ we have structurally estimated from the firm-level production dataset. $\mathbf{X}_{jt}$ are the vector of control variables that include firm features (e.g., firm age, firm size, capital intensity, research and development, and ownership), local industry feature (industry cluster), and fixed effects at industry and county-year levels. Table 4 reports how the change of upstream/downstream market accesses and market competitions affects firms' performance. There are three main findings.

Firstly, better upstream market access decreases input prices (columns 3 and 4, panel A) and increases markup (columns 5 and 6, panel A). In panel B, we make use of the small sample of firms with the price information to see how the output prices are influenced and find that increased upstream market access lowers the output price (columns 1 and 2, panel B). Given that the output price is decreased, the increase in markup may be because firms can source the suppliers with cheaper input prices from the enlarged upstream markets, and the marginal cost of production becomes lower. Thus the firms can have a higher markup even if the output price becomes lower.

Secondly, better downstream market access increases productivity and markup. Since our estimate of productivity is TFPR, the productivity may be raised by both physical productivity increases and the output price increases (columns 1 and 2, panel B). Being close to the downstream market may have better knowledge spillover, and the output price is also increased because of the increased market size. By constructing the raw markup index (revenue and total variable cost ratio) as in De Loecker and Warzynski (2012), we can see that the increase in downstream market access raises firms' raw markup, and the results are consistent with our estimated markup (columns 5 and 6, panel A).

Thirdly, increased market competition harms firms. The better connection also brings more competitors, and firms' markup decreases (columns 5 and 6). Under the more fierce competition situation, firms have to lower the output prices to compete in the market. Meanwhile, these firms also need to compete for the suppliers, thus we can observe the input price increases (columns 3 and 4, panel A).

As both the market access and market competition increase, we make a simple counter-factual analysis to evaluate which effects dominate. We assume the upstream/downstream market accesses and the market competition remain the same as the level in 1998. In Figure 11, we plot the percentage change of the mean w.r.t. firms' productivity, input price, and markup. The results show that if there's no expressways improvement during the 10 years, firms will have a 40% decrease in productivity, 17% increase in input price, and 1.7% decrease in markup. Thus, we can see that while more market competition has opposite impacts, they are dominated by market access effects. Firms eventually benefit from the expressway network improvements.

Table 4: Influences of Expressway Networks Construction

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Influences on the productivity, input price, and markup						
	Productivity(log)		Input Price(log)		Markup(log)	
MA^U_{okt} (log)	0.064***	0.056***	-0.064***	-0.061***	0.006***	0.006***
	(0.0084)	(0.0086)	(0.0064)	(0.0064)	(0.0006)	(0.0006)
MA^H_{okt} (log)	-0.019***	-0.020***	0.019***	0.015***	-0.002***	-0.001***
	(0.0027)	(0.0027)	(0.0022)	(0.0022)	(0.0002)	(0.0002)
MA^D_{okt} (log)	0.131***	0.126***	-0.031***	-0.025***	0.003***	0.003***
	(0.0077)	(0.0078)	(0.0056)	(0.0056)	(0.0005)	(0.0005)
Observations	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295
Adjusted R^2	0.712	0.713	0.587	0.589	0.439	0.440
Mean of Dep. Var.	3.618	3.618	1.395	1.395	0.067	0.067
SD of Dep. Var.	2.514	2.514	1.573	1.573	0.122	0.122
Panel B: Influences on the output price, revenue, and raw markup						
	(1)	(2)	(3)	(4)	(5)	(6)
	Output Price(log)		Revenue(log)		Raw Markup(log)	
MA^U_{okt} (log)	-0.163***	-0.165***	0.154***	0.132***	0.004***	0.004***
	(0.0213)	(0.0213)	(0.0065)	(0.0066)	(0.0008)	(0.0008)
MA^H_{okt} (log)	-0.112***	-0.102***	-0.017***	0.005***	-0.003***	-0.002***
	(0.0121)	(0.0122)	(0.0018)	(0.0017)	(0.0003)	(0.0003)
MA^D_{okt} (log)	0.064***	0.054**	0.211***	0.175***	0.008***	0.007***
	(0.0220)	(0.0219)	(0.0059)	(0.0061)	(0.0008)	(0.0008)
Observations	454,313	454,313	1,435,295	1,435,295	1,435,295	1,435,295
Adjusted R^2	0.517	0.518	0.203	0.295	0.181	0.186
Mean of Dep. Var.	-0.227	-0.227	10.005	10.005	0.185	0.185
SD of Dep. Var.	3.396	3.396	1.250	1.250	0.216	0.216
Capital Intensity(log), R&D		YES		YES		YES
Industry Cluster		YES		YES		YES
Firm Age(log), Firm Size	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES

Note 1: (Mean, SD) for MA^U_{okt} (log), MA^H_{okt} (log), MA^D_{okt} (log) are (0.241, 1.705), (-1.794, 1.451), (-0.171, 1.821) resp.

Note 2: Fixed effects include industry, ownership, and county×year fixed effects.

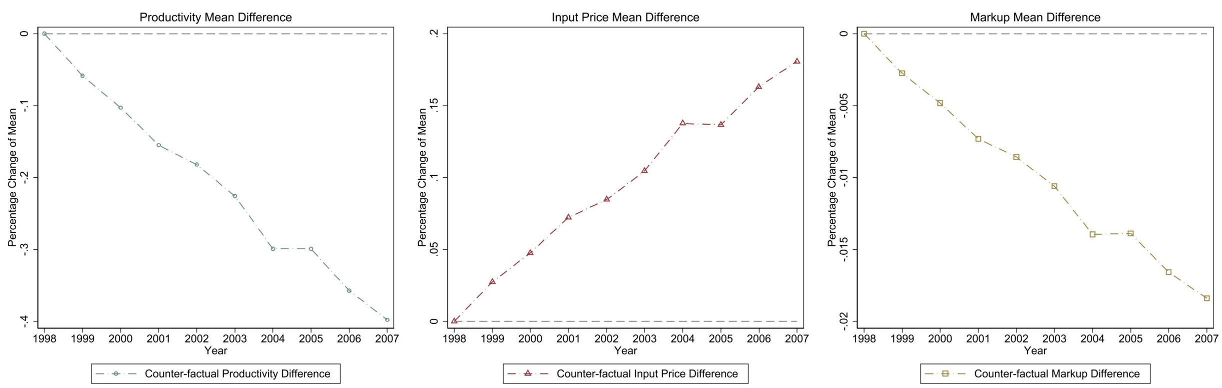
Note 3: The industry cluster is measured by county-industry-year level local quotient.

Note 4: The firm size is captured by the firm's capital level.

Note 5: Standard errors (clustered at county-year level) in parentheses.

* $p < .10$, ** $p < .05$, *** $p < .01$.

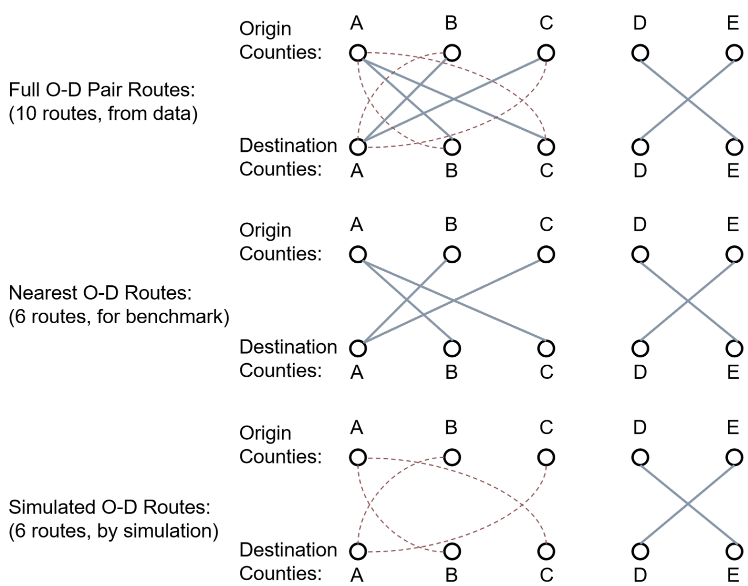
Figure 11: Counterfactual Firms' Performance



6.2 Identification

The endogeneity concern is that regions that are more central in the economic geography are more likely to grow more in market access and market competition. So we adopt the method proposed in Borusyak and Hull (2023) to construct the recentered instruments for upstream/downstream market access and market competition. In Figure 12, we illustrate how we simulate the O-D routes based on the full routes between each origin and destination pair from the data in 2007. In the expressway GIS data in 2007, we can generate multiple routes in terms of each pair of origin and destination counties, because there could be multiple expressways across one county. When generating the market access and competition indices, we only keep the shortest route of each O-D pair. In each simulation, we randomly draw the same number of routes as in the data. Then, we identify the shortest routes based on the simulated sample and recalculate the indices for market access and competition.

Figure 12: Simulated O-D Routes



After simulating 100 times, we can estimate the expected upstream market access ($\bar{MA}^U_{okt}$), expected downstream market access ($\bar{MA}^D_{okt}$), and market competition ($\bar{MA}^H_{okt}$) by simply calculating the means of the simulated indices. By subtracting the expected market access and market competition from the indices generated from the real data, we get the recentered instruments for the upstream market access, downstream market access, and market competition, denoted by Z^U_{okt}, Z^D_{okt}, and Z^H_{okt}, respectively:

$$Z^U_{okt} = \ln\left(MA^U_{okt}\right) - \ln\left(\bar{M}A^U_{okt}\right), \tag{17}$$

$$Z^H_{okt} = \ln\left(MA^H_{okt}\right) - \ln\left(\bar{M}A^H_{okt}\right), \tag{18}$$

$$Z^D_{okt} = \ln\left(MA^D_{okt}\right) - \ln\left(\bar{M}A^D_{okt}\right), \tag{19}$$

In Table 5, we report the IV estimation using the recentered instruments. The results show that after addressing the endogeneity problem, the effects of expressways construction are slightly larger than the benchmark results.

Table 5: Recentered Estimates Results

Panel A: Influences on the productivity, input price, and markup						
	Productivity(log)		Input Price(log)		Markup(log)	
	(1)	(2)	(3)	(4)	(5)	(6)
MA^U_{okt} (log)	0.079***	0.079***	-0.088***	-0.086***	0.008***	0.008***
	(0.0106)	(0.0107)	(0.0078)	(0.0077)	(0.0007)	(0.0007)
MA^H_{okt} (log)	-0.039***	-0.063***	0.051***	0.056***	-0.004***	-0.005***
	(0.0074)	(0.0077)	(0.0054)	(0.0053)	(0.0005)	(0.0005)
MA^D_{okt} (log)	0.127***	0.127***	-0.021***	-0.017***	0.002***	0.002***
	(0.0090)	(0.0091)	(0.0067)	(0.0067)	(0.0006)	(0.0006)
Observations	1,333,306	1,333,306	1,333,306	1,333,306	1,333,306	1,333,306
Adjusted R^2	0.008	0.010	0.005	0.010	0.007	0.009
1^{st} Step Kleibergen-Paap rk Wald F stat.	513.514	506.577	513.514	506.577	513.514	506.577
Panel B: Influences on the output price, revenue, and raw markup						
	Output Price(log)		Revenue(log)		Raw Markup(log)	
	(1)	(2)	(3)	(4)	(5)	(6)
MA^U_{okt} (log)	-0.083***	-0.087***	0.165***	0.151***	0.005***	0.005***
	(0.0241)	(0.0240)	(0.0083)	(0.0080)	(0.0010)	(0.0010)
MA^H_{okt} (log)	-0.023	-0.011	-0.002	-0.026***	-0.003***	-0.003***
	(0.0251)	(0.0249)	(0.0051)	(0.0060)	(0.0008)	(0.0008)
MA^D_{okt} (log)	-0.053**	-0.061**	0.218***	0.196***	0.007***	0.006***
	(0.0256)	(0.0256)	(0.0068)	(0.0067)	(0.0010)	(0.0010)
Observations	430,220	430,220	1,333,306	1,333,306	1,333,306	1,333,306
Adjusted R^2	0.006	0.008	0.062	0.117	0.001	0.007
1^{st} Step Kleibergen-Paap rk Wald F stat.	339.243	330.086	513.514	506.577	513.514	506.577
Capital Intensity(log), R&D		YES		YES		YES
Industry Cluster		YES		YES		YES
Firm Age(log), Firm Size	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES

Note 1: To generate the recentered instruments, we have simulate the routes for 100 times to calculate the expected market access.
Note 1: Fixed effects include industry, ownership, and county×year fixed effects.
Note 2: The industry cluster is measured by county-industry-year level local quotient.
Note 3: The firm size is captured by the firm's capital level.
Note 4: Standard errors (clustered at county-year level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$.

6.3 Robustness

To address the concern that the changes in market access/competition are driven by the areas of the counties, we adjust the transportation weights of the counties. We multiply the transportation weight by 0.5 for the local market to decrease the weight of the local market, which is not affected by the construction of expressways. For the transportation weights of counties $o' \neq o$, we remain the same as in the benchmark results, because the distance across counties ($d_{o'ot}$) dominates the denominator of the $\tilde{\delta}_{oo't}$. So the adjusted weights are calculated by:

$$\tilde{\delta}_{oo't} = \begin{cases} 0.5 * \frac{1}{\sqrt{S_o}/2} & \text{if } o' = o, \\ \frac{1}{d_{o'ot} + \left(\sqrt{S_o} + \sqrt{S_{o'}}\right)/2} & \text{if } o' \neq o, \end{cases} \tag{20}$$

Then we recalculate the market access/competition indices and repeat the benchmark analysis. The results are reported in Table B1 in Appendix B.1. All the results are highly robust to the main results, indicating that the effects of market access/competition are not driven by the areas of counties.

6.4 Heterogeneous Effects

Although we have evaluated firms' average gain from the enhanced transportation connections, we cannot explain why the firms' rank in the local market is reshuffled after the connection. Therefore, we test the heterogeneous effects of the increased market access and market competition to different groups of firms.

We define high-performance firms in three dimensions–high-productivity firms, low-input price firms, and high-markup firms. The estimation is based on the specifications:

$$\begin{aligned} \ln FP_{jt} = {} & \beta_0 + \beta_{up}^{High} D_{jk,t=0}^{High} \times \ln MA_{jkt}^{U} + \beta_{hori}^{High} D_{jk,t=0}^{High} \times \ln MA_{jkt}^{H} + \beta_{down}^{High} D_{jk,t=0}^{High} \times \ln MA_{jkt}^{D} \\ & + \beta^{High} D_{jk,t=0}^{High} + \beta_{up} \ln MA_{jkt}^{U} + \beta_{hori} \ln MA_{jkt}^{H} + \beta_{down} \ln MA_{jkt}^{D} \\ & + \beta_{X}^{High} \mathbf{X}_{jt} + \epsilon_{jt}^{FP}, \end{aligned} \tag{21}$$

where $D_{jk,t=0}^{High}$ indicates whether a firm is high-performance is determined by whether a firm's productivity/markup is higher than the median of the four-digit industry in the first year of its appearance in the sample, and whether the input price is lower than the corresponding median, respectively. Thus, the sign and significance of the coefficients $\left(\beta_{up}^{High}, \beta_{hori}^{High}, \beta_{down}^{High}\right)$

implicate whether the high-performance firms can benefit more in the increases in market access/competition.

Table 6 reports the main heterogeneous effects, and the tables with the full information are reported in Tables B2 to B4 in appendix B.1. We find that no matter what index we choose as the classification standard, firms initially at the higher end of the performance always gain less than the low-performance firms. These may explain why those high-performance firms will be caught up or even replaced by the low-performance firms.

Will these heterogeneous effects drive market integration? Similarly, we make a simple counter-factual analysis to see that if the market accesses and the market competition remain unchanged as in 1998, what will the gap between high-performance firms and low-performance firms be? In Figure 13, we plot the percentage change of the interquartile range (IQR) w.r.t. productivity, input price, and markup. The results show that, there will be approximately 10% increase in the IQR of productivity, 15% increase in the IQR of the input price, and 1.5% increase in the IQR of the markup by 2007.

Table 6: Heterogeneous Effects of Expressway Networks

	(1)	(2)	(3)	(4)	(5)	(6)
	Productivity(log)		Input Price(log)		Markup(log)	
MA^{U}_{okt} (log) $\times D^{High\omega}_{jk,t=0}$	0.030***	0.031***				
	(0.0041)	(0.0041)				
MA^{H}_{okt} (log) $\times D^{High\omega}_{jk,t=0}$	-0.092***	-0.092***				
	(0.0030)	(0.0030)				
MA^{D}_{okt} (log) $\times D^{High\omega}_{jk,t=0}$	-0.086***	-0.086***				
	(0.0031)	(0.0031)				
$D^{High\omega}_{jk,t=0}$	1.146***	1.141***				
	(0.0083)	(0.0083)				
MA^{U}_{okt} (log) $\times D^{LowPM}_{jk,t=0}$			0.012***	0.015***		
			(0.0035)	(0.0035)		
MA^{H}_{okt} (log) $\times D^{LowPM}_{jk,t=0}$			0.010***	0.010***		
			(0.0026)	(0.0026)		
MA^{D}_{okt} (log) $\times D^{LowPM}_{jk,t=0}$			0.057***	0.055***		
			(0.0029)	(0.0028)		
$D^{LowPM}_{jk,t=0}$			-0.923***	-0.919***		
			(0.0080)	(0.0080)		
MA^{U}_{okt} (log) $\times D^{High\mu}_{jk,t=0}$					-0.001**	-0.001***
					(0.0003)	(0.0003)
MA^{H}_{okt} (log) $\times D^{High\mu}_{jk,t=0}$					-0.000	-0.000
					(0.0002)	(0.0002)
MA^{D}_{okt} (log) $\times D^{High\mu}_{jk,t=0}$					-0.007***	-0.007***
					(0.0003)	(0.0003)
$D^{High\mu}_{jk,t=0}$					0.083***	0.083***
					(0.0007)	(0.0007)
Observations	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295
Adjusted R^2	0.775	0.776	0.661	0.662	0.535	0.535
Mean of Dep. Var.	3.618	3.618	1.395	1.395	0.067	0.067
SD of Dep. Var.	2.514	2.514	1.573	1.573	0.122	0.122
Capital Intensity(log), R&D		YES		YES		YES
Industry Cluster		YES		YES		YES
MA^{U}_{okt} (log)	YES	YES	YES	YES	YES	YES
MA^{H}_{okt} (log)	YES	YES	YES	YES	YES	YES
MA^{D}_{okt} (log)	YES	YES	YES	YES	YES	YES
Firm Age(log), Firm Size(K)	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES

Note 1: (Mean, SD) for MA^{U}_{okt} (log), MA^{H}_{okt} (log), MA^{D}_{okt} (log) are (0.191, 1.706), (-1.936, 1.546), (-0.221, 1.823) resp.

Note 2: Fixed effects include industry, ownership, and county$\times$year fixed effects.

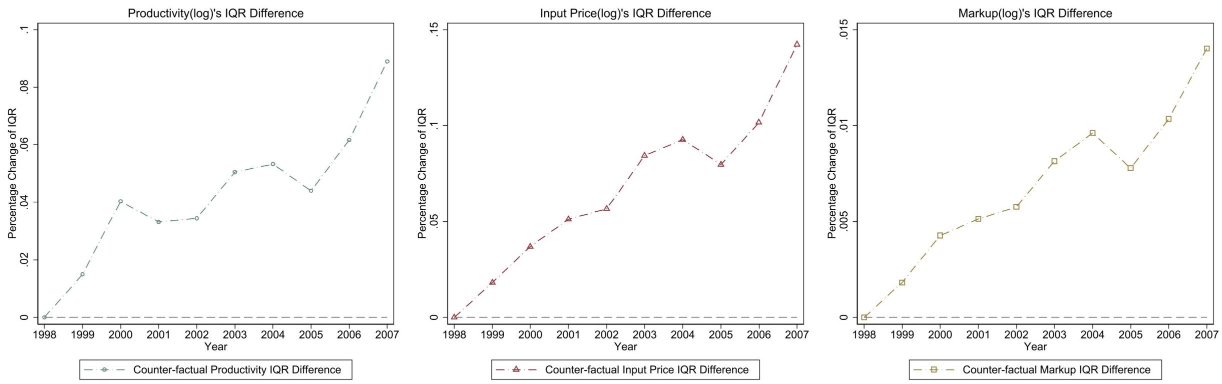
Note 3: The industry cluster is measured by county-industry-year level local quotient.

Note 4: The firm size is captured by the firm's capital level.

Note 5: Standard errors (clustered at county-year level) in parentheses.

* $p < .10$, ** $p < .05$, *** $p < .01$.

Figure 13: Counterfactual High-Low-Performance Firms' Gap



7 Mechanisms

Why may increases in market access benefit firm performance? In this section, we show evidence that increases in market access reduce the distortions in the county and that the decreases in the distortions benefit the firms.

7.1 Transportation Network and Distortions

In Table 7, we report how changes in market access influence the distortions. We calculate the county-industry-year level dispersions of firms' subsidy-sales ratio and our estimated firms' markup. Columns (1) and (3) report the estimation results after controlling the industry and county-year fixed effects, and columns (2) and (4) further control the corresponding industry cluster within the counties. The results of the subsidy-sales ratio and markup all show that the increases in market access significantly reduce the county-industry distortions.

These results are highly consistent with the findings in Wu et al. (2023). Their paper investigates the effect of China's road expansion on allocation efficiency in the same time period as ours, and they find that the improved transportation infrastructure helps to reduce the distortions caused by resource misallocation. In addition to the resource misallocation, the other sources of distortions may come from government assistance (e.g. Harris and Li, 2019). The subsidy can also be interpreted as the resource allocated to the firms. In a world without distortions, the subsidy should be allocated to the most competitive firms (e.g. Restuccia and Rogerson, 2008; Aghion et al., 2015). So, the dispersion of subsidies within the same industry also reflects the allocation efficiency. Being connected means more exposure to external monitoring, and this may improve local government's behavior in subsidizing firms (e.g. Li and Zhang, 2022). In addition, firms initially harmed by the distortions can now source from and sell to the newly connected remote markets to counteract or bargain with the local government, lowering the distortions.

Table 7: Expressways and Distortions

	(1)	(2)	(3)	(4)
	$\sigma_{okt}^{SSratio}$ (log)	$\sigma_{okt}^{SSratio}$ (log)	σ_{okt}^{μ} (log)	σ_{okt}^{μ} (log)
MA_{okt}^{U} (log)	-0.177***	-0.162***	-0.017**	-0.014*
	(0.0289)	(0.0291)	(0.0079)	(0.0079)
MA_{okt}^{H} (log)	0.106***	0.099***	0.096***	0.094***
	(0.0122)	(0.0123)	(0.0035)	(0.0035)
MA_{okt}^{D} (log)	-0.150***	-0.139***	-0.019**	-0.016**
	(0.0272)	(0.0273)	(0.0075)	(0.0075)
Observations	71,981	71,981	182,886	182,886
Adjusted R^2	0.193	0.194	0.106	0.106
Industry Cluster		YES		YES
Fixed Effects	YES	YES	YES	YES

Note 1: Fixed effects include industry and county×year fixed effects.

Note 2: The firm size is captured by the firm's capital level.

Note 3: Standard errors (clustered at county-year level) in parentheses.

* $p < .10$, ** $p < .05$, *** $p < .01$.

7.2 Distortions and Firm Performance

What are the outcomes of distortions decreasing? In panel A of Table 8, we report the average influences of the changes in distortions on firms' productivity, input price, and markup. In panel B, we report the influences on the dispersions of firms' performance. In columns (1), (3), and (5) of both panels, we use the dispersion of subsidy-sales ratio as the measurement for distortion, while in columns (2), (4), and (6), we use the dispersion of markup as the measurement for distortion.

Panel A shows that, regardless of which distortion measurement we use, the decreases in distortions averagely raise firms' productivity and markup while reducing firms' input prices. As stated in the previous subsection, the enhanced economic performance may be driven by the improved allocation efficiency of both inputs and government assistance. In panel B, we show that the decreased distortions also reduce the dispersion of firm performance. This may also explain why the increases in market access drive the market to be more integrated. Ideally, in a world of no distortions, firms may have the same performance as predicted in the market integration theories (e.g., Marshall, 1890; Stigler, 1949; Samuelson, 1952).

Table 8: Distortions' Influence on County-Industry Aggregate Performance

Panel A: Average Influences on the firms' productivity, input price, and markup						
	(1)	(2)	(3)	(4)	(5)	(6)
	Productivity(log)		Input Price(log)		Markup(log)	
$\sigma^{SSratio}_{okt}$ (log)	-0.018***		0.006***		-0.001***	
	(0.0019)		(0.0014)		(0.0001)	
σ^{μ}_{okt} (log)		-0.194***		0.185***		-0.020***
		(0.0034)		(0.0027)		(0.0002)
Capital Intensity(log), R&D	YES	YES	YES	YES	YES	YES
Industry Cluster	YES	YES	YES	YES	YES	YES
Firm Age(log), Firm Size	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES
Observations	850,957	1,351,218	850,957	1,351,218	850,957	1,351,218
Adjusted R^2	0.697	0.712	0.570	0.585	0.425	0.445
Panel B: Influences on the dispersions of productivity, input price, and markup						
	(1)	(2)	(3)	(4)	(5)	(6)
	Productivity SD(log)		Input Price SD(log)		Markup SD(log)	
$\sigma^{SSratio}_{okt}$ (log)	0.007**		0.073***		0.051***	
	(0.0035)		(0.0032)		(0.0020)	
σ^{μ}_{okt} (log)		0.420***		1.117***		1.000***
		(0.0047)		(0.0055)		(0.0000)
Industry Cluster	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES
Observations	71,981	182,886	71,981	182,886	71,981	182,886
Adjusted R^2	0.734	0.736	0.543	0.755	0.192	1.000

Note 1: (Mean, SD) for firms' subsidy-sales ratio and markup are (0.003, 0.044) and (1.077, 0.124), respectively.

Note 2: Fixed effects include industry, ownership, and county×year fixed effects.

Note 3: The industry cluster is measured by county-industry-year level local quotient.

Note 4: The firm size is captured by the firm's capital level.

Note 5: Standard errors (clustered at county-year level) in parentheses.

* $p < .10$, ** $p < .05$, *** $p < .01$.

8 Conclusion

To improve the firm performance and the aggregate economy in the local market, Chinese governments have launched huge plans for transportation construction. However, the empirical literature has not comprehensively investigated the channels through which the transportation network influences firm performance. This paper provides new firm-level evidence on how the developments of expressway networks affect the performance of Chinese manufacturing firms and drive market integration through the channel of reducing distortions in the local markets.

We first introduce the I-O linkage into the traditional market access index to incorporate the heterogeneity of the upstream, downstream, and horizontal markets, which enables us to distinguish the different impacts brought about by the expansion of upstream and downstream markets and the intensification of competition within the same industry. Then we extend the estimation framework offered by Li and Zhang (2022) to allow for the variable markup so that we can investigate how firms' performance is influenced, in terms of productivity, input price, and markup.

Our paper presents three main findings. First, increased upstream and downstream market access benefits firm performance by reducing their input prices and increasing productivity and markup. Second, more connections also bring more market competition. Increased market competition forces firms to face lower productivity and markup, and they have to pay higher input prices to the suppliers. However, the competition effects are dominated by the market access effects. Using the recentered instruments proposed in Borusyak and Hull (2023), we address the endogeneity concern for the non-random expressway constructions, and the results show that the effects are slightly larger than the benchmark results. Moreover, we discover that the construction of the expressway network has heterogeneous effects on different groups of firms. Specifically, firms that initially perform worse within their industries benefit significantly more from the improved transportation conditions. Through the simple counterfactual analysis, we show that the heterogeneous effects of the transportation improvements help to reduce about 10% of the gap in productivity, 15% of the gap in input price, and 1.5% of the gap in markup, which helps us to explain the fact of firms' rank reshuffling and the market integration after connecting to the expressway network. Finally, we investigate the mechanisms through which transportation network development influences firm performance. Increased market access reduces the distortions in the local markets, and the improved allocation efficiency benefits firms' performance.

Our findings provide strong evidence of the channels through which transportation networks reduce distortions and influence firm performance. The transportation network has direct impacts on the distortions of the local market, which influences firm performance, and the heterogeneity of these impacts contributes to firm convergence. All the findings support market integration theories and deepen our understanding of the dynamics involved.

References

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

Aghion, P., Cai, J., Dewatripont, M., Du, L., Harrison, A., and Legros, P. (2015). Industrial policy and competition. *American Economic Journal: Macroeconomics*, 7(4):1–32.

Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3):1085–1140.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American economic review*, 69(1):106–116.

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

Asturias, J., García-Santana, M., and Ramos, R. (2019). Competition and the welfare gains from transportation infrastructure: Evidence from the golden quadrilateral of India. *Journal of the European Economic Association*, 17(6):1881–1940.

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, 98(5):1998–2031.

Badiane, O. and Shively, G. E. (1998). Spatial integration, transport costs, and the response of local prices to policy changes in Ghana. *Journal of Development Economics*, 56(2):411–431.

Bai, C.-E., Du, Y., Tao, Z., and Tong, S. Y. (2004). Local protectionism and regional specialization: evidence from China's industries. *Journal of International Economics*, 63(2):397–417.

Bai, C.-E. and Qian, Y. (2010). Infrastructure development in China: the cases of electricity, highways, and railways. *Journal of Comparative Economics*, 38(1):34–51.

Balboni, C. A. (2024). In harm's way? Infrastructure investments and the persistence of coastal cities. *American Economic Review*, forthcoming.

Banerjee, A., Duflo, E., and Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics*, 145:102442.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of economics*, 122(2):775–805.

Baum-Snow, N. (2020). Urban transport expansions and changes in the spatial structure of us cities: Implications for productivity and welfare. *Review of Economics and Statistics*, 102(5):929–945.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics*, 99(3):435–448.

Baum-Snow, N., Henderson, J. V., Turner, M. A., Zhang, Q., and Brandt, L. (2020). Does investment in national highways help or hurt hinterland city growth? *Journal of Urban Economics*, 115:103124.

Borusyak, K. and Hull, P. (2023). Nonrandom exposure to exogenous shocks. *Econometrica*, 91(6):2155–2185.

Calamunci, F. and Lonsky, J. (2024). The road to crime: An unintended consequence of the interstate highway system. *The Economic Journal*, page ueae068.

Chakrabarti, S., Kushari, T., and Mazumder, T. (2022). Does transportation network centrality determine housing price? *Journal of Transport Geography*, 103:103397.

Chaney, T. (2018). The gravity equation in international trade: An explanation. *Journal of Political Economy*, 126(1):150–177.

Costinot, A. and Donaldson, D. (2016). How large are the gains from economic integration? theory and evidence from US agriculture, 1880-1997. Technical report, National Bureau of Economic Research.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

De Loecker, J. and Warzynski, F. (2012). Markups and firm-level export status. *American economic review*, 102(6):2437–2471.

Donaldson, D. (2015). The gains from market integration. *Annual Review of Economics*, 7(1):619–647.

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934.

Donaldson, D. and Hornbeck, R. (2016). Railroads and American economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Dong, X., Zheng, S., and Kahn, M. E. (2020). The role of transportation speed in facilitating high skilled teamwork across cities. *Journal of Urban Economics*, 115:103212. Cities in China.

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

Edmond, C., Midrigan, V., and Xu, D. Y. (2015). Competition, markups, and the gains from international trade. *American Economic Review*, 105(10):3183–3221.

Edmond, C., Midrigan, V., and Xu, D. Y. (2023). How costly are markups? *Journal of Political Economy*, 131(7):1619–1675.

Egger, P. H., Loumeau, G., and Loumeau, N. (2023). China's dazzling transport-infrastructure growth: Measurement and effects. *Journal of International Economics*, 142:103734.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from China's national trunk highway system. *Review of Economic Studies*, 81(3):1046–1070.

Fan, J. (2019). Internal geography, labor mobility, and the distributional impacts of trade. *American Economic Journal: Macroeconomics*, 11(3):252–288.

Fan, J., Lu, Y., and Luo, W. (2023). Valuing domestic transport infrastructure: A view from the route choice of exporters. *Review of Economics and Statistics*, 105(6):1562–1579.

Fleming, D. K. and Hayuth, Y. (1994). Spatial characteristics of transportation hubs: centrality and intermediacy. *Journal of Transport Geography*, 2(1):3–18.

Fogel, R. W. (1964). *Railroads and American economic growth*. Johns Hopkins Press Baltimore.

Freiria, S., Ribeiro, B., and Tavares, A. O. (2015). Understanding road network dynamics: Link-based topological patterns. *Journal of Transport Geography*, 46:55–66.

Grieco, P. L., Li, S., and Zhang, H. (2016). Production function estimation with unobserved input price dispersion. *International Economic Review*, 57(2):665–690.

Gümüş, M., Ray, S., and Gurnani, H. (2012). Supply-side story: Risks, guarantees, competition, and information asymmetry. *Management Science*, 58(9):1694–1714.

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

Harris, R. and Li, S. (2019). Government assistance and total factor productivity: firm-level evidence from china. *Journal of Productivity Analysis*, 52(1):1–27.

He, G., Xie, Y., and Zhang, B. (2020). Expressways, GDP, and the environment: The case of China. *Journal of Development Economics*, 145:102485.

Head, K. and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics*, volume 4, pages 131–195. Elsevier.

Holmes, T. J., Hsu, W.-T., and Lee, S. (2014). Allocative efficiency, mark-ups, and the welfare gains from trade. *Journal of International Economics*, 94(2):195–206.

Hornbeck, R. and Rotemberg, M. (2024). Growth off the rails: Aggregate productivity growth in distorted economies. *Journal of Political Economy*, 132(11):3547–3602.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in China and India. *The Quarterly journal of economics*, 124(4):1403–1448.

Jaworski, T. and Kitchens, C. T. (2019). National policy for regional development: Historical evidence from Appalachian highways. *Review of Economics and Statistics*, 101(5):777–790.

Ji, S., Wang, X., Lyu, T., Liu, X., Wang, Y., Heinen, E., and Sun, Z. (2022). Understanding cycling distance according to the prediction of the xgboost and the interpretation of shap: A non-linear and interaction effect analysis. *Journal of Transport Geography*, 103:103414.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–499.

Krugman, P. et al. (1980). Scale economies, product differentiation, and the pattern of trade. *American economic review*, 70(5):950–959.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2):317–341.

Li, S. and Zhang, H. (2022). Does external monitoring from the government improve the performance of state-owned enterprises? *The Economic Journal*, 132(642):675–708.

Li, X., Ma, L., and Tang, Y. (2024). Migration and resource misallocation in China. *Journal of Development Economics*, 167:103218.

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

Marshall, A. (1890). *The Principles of Economics*. Macmillan.

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

Newman, M. (2018). *Networks*. Oxford university press.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

Parsley, D. C. and Wei, S.-J. (1996). Convergence to the law of one price without trade barriers or currency fluctuations. *The Quarterly Journal of Economics*, 111(4):1211–1236.

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

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

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

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics*, 11(4):707–720.

Rodrigue, J.-P. (2020). *The geography of transport systems*. Routledge.

Samuelson, P. A. (1952). Spatial price equilibrium and linear programming. *The American economic review*, 42(3):283–303.

Song, Z., Storesletten, K., and Zilibotti, F. (2011). Growing like China. *American economic review*, 101(1):196–233.

Stigler, G. J. (1949). A theory of delivered price systems. *The American Economic Review*, 39(6):1144–1159.

Tinbergen, J. (1962). An analysis of world trade flows. *Shaping the world economy*, 3:1–117.

Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of China. *American Economic Review*, 109(5):1843–1872.

Wang, J., Mo, H., Wang, F., and Jin, F. (2011). Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. *Journal of Transport Geography*, 19(4):712–721.

Wu, M., Yu, L., and Zhang, J. (2023). Road expansion, allocative efficiency, and pro-competitive effect of transport infrastructure: Evidence from China. *Journal of Development Economics*, 162:103050.

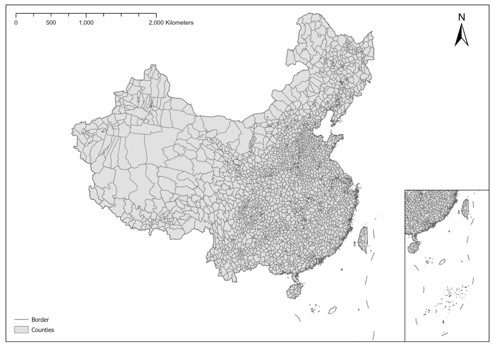
Zhang, X., Miller-Hooks, E., and Denny, K. (2015). Assessing the role of network topology in transportation network resilience. *Journal of Transport Geography*, 46:35–45.

A Appendix - Data Proccessing

A.1 Firm Location

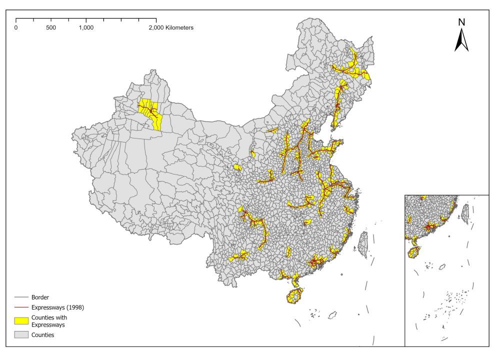
- **Polygon-type**: Map of China (County-level, fixed)

Figure 14: Map of China



- **Line-type**: Expressway Data (of years 1998, 2000, 2002, 2003, 2005, 2007)

Figure 15: Expressways(1998)



- **Point-type**: County-level Government Coordinate Data (As a proxy for the firms' locations within this county, fixed)

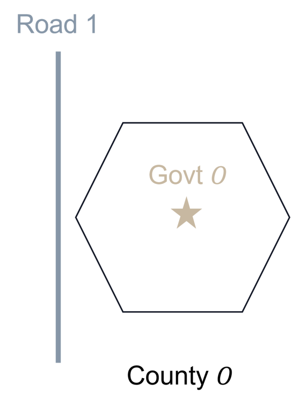
Figure 16: Government Coordinates of Some Counties in Shanxi Province

A.2 Entrants & Exits of Expressways

- **Type I. No Road Cross**

  The most common case is that there's no express way cross the county:

  Figure 17: No Road Cross

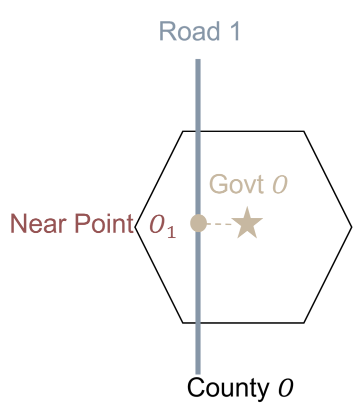
  

  For this case, we don't generate the near point. Thus, there's no minimal OD cost distance for this type of counties.

- **Type II. One Road Cross**

  For those counties with roads cross, the most common case is that there's only one road in the county:

  Figure 18: One Road Cross

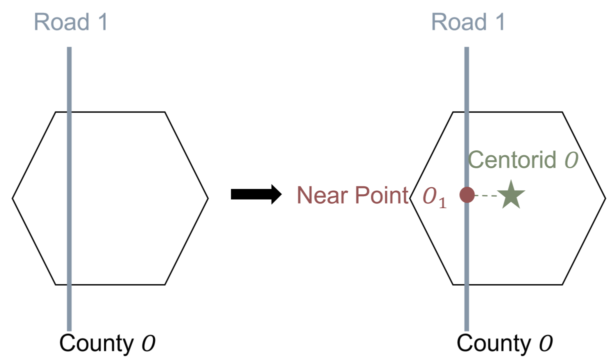
  

  For this kind of counties, we just generate one near-point directly as the entrance of that specific expressway in the county.

- **Type III. Road(s) Cross (No Government Coordinates)**

There're some counties with road(s) that don't have the county government coordinates:

Figure 19: One Road Cross (No Government Coordinates)

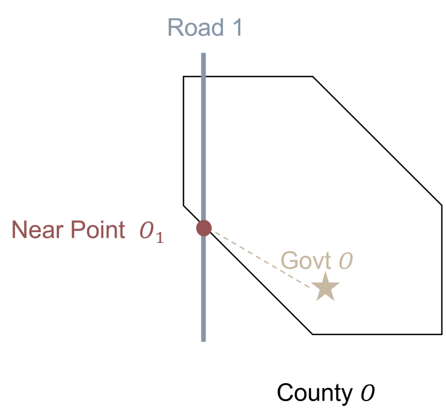


For this type of counties, we first need to generate the centroid of the county as the proxy for the firms' location in the county. And then make use of the newly generated centroid to generate the near-point(s) as the entrance(s) of the expressway(s) in the county.

- **Type IV. One Road Cross (Euclidean Near Point Outside County)**

  There has also been a kind of situation when generating the near-points - the euclidean near point on the expressway is not within the county:

Figure 20: One Road Cross (Euclidean Near Point Outside County)

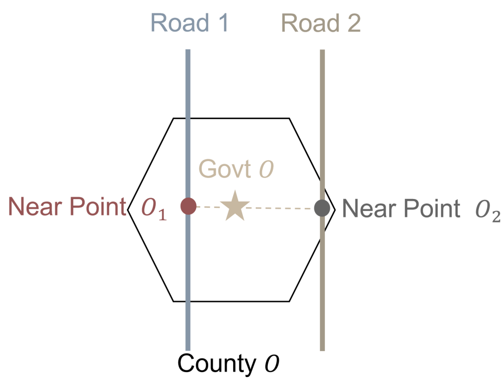


Since we interpret the near-point as the entrance/exit of the expressway in the county, this type of near-points are not that consistent with our assumption. Therefore, for this type of counties, we first cut off the roads outside the county and only keep the roads inside the county, and then generate the nearest points on the roads inside the county.

- **Type V. Multiple Roads Cross**

  For many important traffic hubs, there are multiple expressways cross:

  Figure 21: Multiple Roads Cross

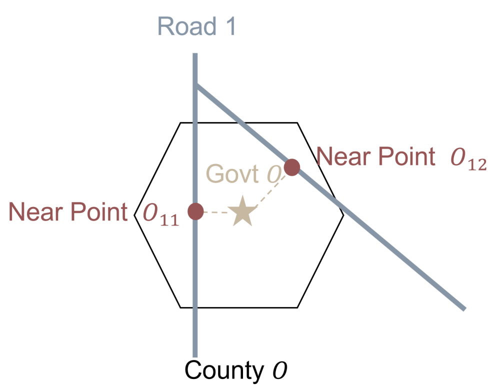


  For this kind of counties, we generate the corresponding near-point on each road that cross the county. Thus there will be multiple near-points for this type of counties.

- **Type VI. One Road Cross (Multiple Branches Cross)**

  Sometimes there're different branches of the same expressways in the county:

  Figure 22: One Road Cross (Multiple Branches Cross)



  Consider that the different branches usually heading to different directions, so there could be multiple entrance for different branches within one county. So we just treated as the former case to generate multiple near points.

B Appendix - Figures & Tables

B.1 Tables

Table B1: Robustness Check

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Influences on the productivity, input price, and markup						
	Productivity(log)		Input Price(log)		Markup(log)	
$\tilde{MA}^U_{okt}$ (log)	0.065***	0.056***	-0.063***	-0.060***	0.006***	0.006***
	(0.0088)	(0.0089)	(0.0066)	(0.0066)	(0.0006)	(0.0006)
$\tilde{MA}^H_{okt}$ (log)	-0.019***	-0.019***	0.018***	0.014***	-0.001***	-0.001***
	(0.0028)	(0.0028)	(0.0023)	(0.0022)	(0.0002)	(0.0002)
$\tilde{MA}^D_{okt}$ (log)	0.131***	0.126***	-0.032***	-0.026***	0.003***	0.003***
	(0.0080)	(0.0081)	(0.0058)	(0.0058)	(0.0005)	(0.0005)
Observations	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295
Adjusted R^2	0.711	0.713	0.587	0.589	0.439	0.440
Mean of Dep. Var.	3.618	3.618	1.395	1.395	0.067	0.067
SD of Dep. Var.	2.514	2.514	1.573	1.573	0.122	0.122
Panel B: Influences on the output price, revenue, and raw markup						
	(1)	(2)	(3)	(4)	(5)	(6)
	Output Price(log)		Revenue(log)		Raw Markup(log)	
$\tilde{MA}^U_{okt}$ (log)	-0.152***	-0.153***	0.147***	0.126***	0.004***	0.004***
	(0.0221)	(0.0222)	(0.0065)	(0.0067)	(0.0009)	(0.0009)
$\tilde{MA}^H_{okt}$ (log)	-0.114***	-0.104***	-0.019***	0.005***	-0.003***	-0.002***
	(0.0127)	(0.0128)	(0.0018)	(0.0017)	(0.0003)	(0.0003)
$\tilde{MA}^D_{okt}$ (log)	0.060***	0.050**	0.208***	0.172***	0.008***	0.007***
	(0.0228)	(0.0227)	(0.0061)	(0.0063)	(0.0009)	(0.0009)
Observations	454,313	454,313	1,435,295	1,435,295	1,435,295	1,435,295
Adjusted R^2	0.517	0.517	0.202	0.294	0.181	0.186
Mean of Dep. Var.	-0.227	-0.227	10.005	10.005	0.185	0.185
SD of Dep. Var.	3.396	3.396	1.250	1.250	0.216	0.216
Capital Intensity(log), R&D		YES		YES		YES
Industry Cluster		YES		YES		YES
Firm Age(log), Firm Size	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES

Note 1: (Mean, SD) for MA^U_{okt} (log), MA^H_{okt} (log), MA^D_{okt} (log) are (0.241, 1.705), (-1.794, 1.451), (-0.171, 1.821) resp.
Note 2: Fixed effects include industry, ownership, and county×year fixed effects.
Note 3: The industry cluster is measured by county-industry-year level local quotient.
Note 4: The firm size is captured by the firm's capital level.
Note 5: Standard errors (clustered at county-year level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Table B2: Heterogeneous Effects (in terms of High Productivity)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Productivity(log)		Input Price(log)		Markup(log)		Output Price(log)		Revenue(log)		Raw Markup(log)	
MA^U_{okt} (log) $\times D^{High\omega}_{jk,t=0}$	0.030***	0.031***	0.016***	0.020***	-0.000	-0.000	0.038***	0.034***	0.038***	0.022***	0.000	-0.000
	(0.0041)	(0.0041)	(0.0033)	(0.0033)	(0.0003)	(0.0003)	(0.0086)	(0.0086)	(0.0028)	(0.0026)	(0.0005)	(0.0005)
MA^H_{okt} (log) $\times D^{High\omega}_{jk,t=0}$	-0.092***	-0.092***	-0.031***	-0.035***	0.002***	0.002***	-0.017**	-0.013*	-0.014***	-0.000	-0.001***	-0.000
	(0.0030)	(0.0030)	(0.0025)	(0.0025)	(0.0002)	(0.0002)	(0.0074)	(0.0074)	(0.0022)	(0.0020)	(0.0003)	(0.0003)
MA^D_{okt} (log) $\times D^{High\omega}_{jk,t=0}$	-0.086***	-0.086***	0.042***	0.039***	-0.006***	-0.006***	-0.063***	-0.061***	-0.066***	-0.056***	-0.004***	-0.003***
	(0.0031)	(0.0031)	(0.0029)	(0.0029)	(0.0002)	(0.0002)	(0.0083)	(0.0083)	(0.0024)	(0.0022)	(0.0004)	(0.0004)
$D^{High\omega}_{jk,t=0}$	1.146***	1.141***	-0.589***	-0.606***	0.059***	0.060***	-0.010	0.008	0.196***	0.245***	0.023***	0.026***
	(0.0083)	(0.0083)	(0.0071)	(0.0072)	(0.0007)	(0.0007)	(0.0155)	(0.0156)	(0.0047)	(0.0045)	(0.0008)	(0.0008)
MA^U_{okt} (log)	0.016**	0.010	-0.057***	-0.056***	0.004***	0.004***	-0.179***	-0.182***	0.130***	0.115***	0.004***	0.003***
	(0.0075)	(0.0076)	(0.0063)	(0.0063)	(0.0006)	(0.0005)	(0.0215)	(0.0215)	(0.0065)	(0.0066)	(0.0009)	(0.0009)
MA^H_{okt} (log)	0.025***	0.026***	0.036***	0.032***	-0.003***	-0.002***	-0.106***	-0.095***	-0.010***	0.006***	-0.002***	-0.002***
	(0.0030)	(0.0030)	(0.0027)	(0.0026)	(0.0002)	(0.0002)	(0.0127)	(0.0127)	(0.0019)	(0.0018)	(0.0003)	(0.0003)
MA^D_{okt} (log)	0.140***	0.136***	-0.039***	-0.031***	0.005***	0.004***	0.094***	0.077***	0.235***	0.193***	0.009***	0.007***
	(0.0066)	(0.0067)	(0.0054)	(0.0054)	(0.0005)	(0.0005)	(0.0225)	(0.0224)	(0.0061)	(0.0062)	(0.0009)	(0.0009)
Observations	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	454,313	454,313	1,435,295	1,435,295	1,435,295	1,435,295
Adjusted R^2	0.775	0.776	0.613	0.616	0.486	0.488	0.516	0.518	0.212	0.305	0.184	0.190
Mean of Dep. Var.	3.618	3.618	1.395	1.395	0.067	0.067	-0.227	-0.227	10.005	10.005	0.185	0.185
SD of Dep. Var.	2.514	2.514	1.573	1.573	0.122	0.122	3.396	3.396	1.250	1.250	0.216	0.216
Capital Intensity(log), R&D		YES		YES		YES		YES		YES		YES
Industry Cluster		YES		YES		YES		YES		YES		YES
Firm Age(log), Firm Size(K)	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note 1: (Mean, SD) for MA^U_{okt} (log), MA^H_{okt} (log), MA^D_{okt} (log) are (0.241, 1.705), (-1.794, 1.451), (-0.171, 1.821), respectively.
Note 2: Fixed effects include industry, ownership, and county×year fixed effects.
Note 3: The industry cluster is measured by county-industry-year level local quotient.
Note 4: The firm size is captured by the firm's capital level.
Note 5: Standard errors (clustered at county-year level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Table B3: Heterogeneous Effects (in terms of Low Input Price)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Productivity(log)		Input Price(log)		Markup(log)		Output Price(log)		Revenue(log)		Raw Markup(log)	
MA^U_{okt} (log) $\times D^{LowP^M}_{jk,t=0}$	0.027***	0.030***	0.012***	0.015***	-0.001***	-0.001***	-0.035***	-0.040***	0.031***	0.018***	-0.000	-0.001
	(0.0044)	(0.0043)	(0.0035)	(0.0035)	(0.0003)	(0.0003)	(0.0092)	(0.0091)	(0.0029)	(0.0027)	(0.0004)	(0.0004)
MA^H_{okt} (log) $\times D^{LowP^M}_{jk,t=0}$	-0.029***	-0.029***	0.010***	0.010***	0.001***	0.001***	0.037***	0.036***	-0.008***	-0.009***	-0.000	-0.000
	(0.0030)	(0.0030)	(0.0026)	(0.0026)	(0.0002)	(0.0002)	(0.0082)	(0.0081)	(0.0022)	(0.0021)	(0.0003)	(0.0003)
MA^D_{okt} (log) $\times D^{LowP^M}_{jk,t=0}$	-0.068***	-0.068***	0.057***	0.055***	-0.007***	-0.007***	-0.004	-0.002	-0.036***	-0.026***	-0.004***	-0.004***
	(0.0039)	(0.0039)	(0.0029)	(0.0028)	(0.0003)	(0.0003)	(0.0085)	(0.0085)	(0.0025)	(0.0023)	(0.0004)	(0.0004)
$D^{LowP^M}_{jk,t=0}$	0.657***	0.660***	-0.923***	-0.919***	0.080***	0.080***	0.148***	0.137***	0.144***	0.122***	0.032***	0.031***
	(0.0081)	(0.0081)	(0.0080)	(0.0080)	(0.0007)	(0.0007)	(0.0172)	(0.0172)	(0.0051)	(0.0047)	(0.0008)	(0.0008)
MA^U_{okt} (log)	0.037***	0.028***	-0.049***	-0.048***	0.005***	0.005***	-0.146***	-0.148***	0.137***	0.121***	0.004***	0.003***
	(0.0083)	(0.0084)	(0.0058)	(0.0057)	(0.0005)	(0.0005)	(0.0217)	(0.0217)	(0.0065)	(0.0066)	(0.0008)	(0.0008)
MA^H_{okt} (log)	-0.005*	-0.006**	0.015***	0.011***	-0.002***	-0.002***	-0.133***	-0.119***	-0.013***	0.010***	-0.003***	-0.002***
	(0.0029)	(0.0029)	(0.0026)	(0.0025)	(0.0002)	(0.0002)	(0.0131)	(0.0131)	(0.0019)	(0.0018)	(0.0003)	(0.0003)
MA^D_{okt} (log)	0.147***	0.143***	-0.038***	-0.033***	0.005***	0.005***	0.067***	0.051**	0.224***	0.183***	0.009***	0.008***
	(0.0076)	(0.0077)	(0.0051)	(0.0051)	(0.0005)	(0.0005)	(0.0224)	(0.0222)	(0.0060)	(0.0062)	(0.0008)	(0.0008)
Observations	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	454,313	454,313	1,435,295	1,435,295	1,435,295	1,435,295
Adjusted R^2	0.729	0.730	0.661	0.662	0.527	0.527	0.516	0.518	0.207	0.298	0.186	0.191
Mean of Dep. Var.	3.618	3.618	1.395	1.395	0.067	0.067	-0.227	-0.227	10.005	10.005	0.185	0.185
SD of Dep. Var.	2.514	2.514	1.573	1.573	0.122	0.122	3.396	3.396	1.250	1.250	0.216	0.216
Capital Intensity(log), R&D		YES		YES		YES		YES		YES		YES
Industry Cluster		YES		YES		YES		YES		YES		YES
Firm Age(log), Firm Size(K)	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note 1: (Mean, SD) for MA^U_{okt} (log), MA^H_{okt} (log), MA^D_{okt} (log) are (0.241, 1.705), (-1.794, 1.451), (-0.171, 1.821), respectively.
Note 2: Fixed effects include industry, ownership, and county×year fixed effects.
Note 3: The industry cluster is measured by county-industry-year level local quotient.
Note 4: The firm size is captured by the firm's capital level.
Note 5: Standard errors (clustered at county-year level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Table B4: Heterogeneous Effects (in terms of High Markup)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Productivity(log)		Input Price(log)		Markup(log)		Output Price(log)		Revenue(log)		Raw Markup(log)	
MA^{U}_{okt} (log) $\times D^{High\mu}_{jk,t=0}$	0.033***	0.036***	0.010***	0.013***	-0.001**	-0.001***	-0.023**	-0.028***	0.039***	0.024***	-0.000	-0.001*
	(0.0042)	(0.0041)	(0.0035)	(0.0035)	(0.0003)	(0.0003)	(0.0090)	(0.0090)	(0.0029)	(0.0027)	(0.0005)	(0.0005)
MA^{H}_{okt} (log) $\times D^{High\mu}_{jk,t=0}$	-0.054***	-0.054***	0.003	0.004	-0.000	-0.000	0.037***	0.036***	0.001	-0.002	-0.001	-0.001*
	(0.0030)	(0.0030)	(0.0026)	(0.0026)	(0.0002)	(0.0002)	(0.0081)	(0.0080)	(0.0023)	(0.0021)	(0.0003)	(0.0003)
MA^{D}_{okt} (log) $\times D^{High\mu}_{jk,t=0}$	-0.077***	-0.077***	0.058***	0.056***	-0.007***	-0.007***	-0.018**	-0.014*	-0.060***	-0.046***	-0.004***	-0.003***
	(0.0035)	(0.0035)	(0.0028)	(0.0028)	(0.0003)	(0.0003)	(0.0084)	(0.0084)	(0.0025)	(0.0023)	(0.0004)	(0.0004)
$D^{High\mu}_{jk,t=0}$	0.757***	0.760***	-0.898***	-0.895***	0.083***	0.083***	0.125***	0.116***	0.078***	0.065***	0.034***	0.033***
	(0.0076)	(0.0075)	(0.0079)	(0.0079)	(0.0007)	(0.0007)	(0.0171)	(0.0170)	(0.0051)	(0.0048)	(0.0008)	(0.0008)
MA^{U}_{okt} (log)	0.031***	0.022***	-0.049***	-0.047***	0.004***	0.004***	-0.151***	-0.153***	0.135***	0.120***	0.004***	0.003***
	(0.0082)	(0.0083)	(0.0058)	(0.0057)	(0.0005)	(0.0005)	(0.0216)	(0.0216)	(0.0064)	(0.0066)	(0.0009)	(0.0009)
MA^{H}_{okt} (log)	0.009***	0.007**	0.017***	0.013***	-0.001***	-0.001***	-0.132***	-0.119***	-0.017***	0.006***	-0.003***	-0.002***
	(0.0029)	(0.0029)	(0.0026)	(0.0026)	(0.0002)	(0.0002)	(0.0132)	(0.0132)	(0.0019)	(0.0018)	(0.0003)	(0.0003)
MA^{D}_{okt} (log)	0.149***	0.145***	-0.041***	-0.034***	0.005***	0.005***	0.073***	0.057**	0.235***	0.193***	0.009***	0.007***
	(0.0074)	(0.0075)	(0.0051)	(0.0051)	(0.0005)	(0.0005)	(0.0224)	(0.0223)	(0.0061)	(0.0063)	(0.0008)	(0.0008)
Observations	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	1,435,295	454,313	454,313	1,435,295	1,435,295	1,435,295	1,435,295
Adjusted R^2	0.736	0.738	0.655	0.657	0.535	0.535	0.516	0.518	0.205	0.296	0.187	0.192
Mean of Dep. Var.	3.618	3.618	1.395	1.395	0.067	0.067	-0.227	-0.227	10.005	10.005	0.185	0.185
SD of Dep. Var.	2.514	2.514	1.573	1.573	0.122	0.122	3.396	3.396	1.250	1.250	0.216	0.216
Capital Intensity(log), R&D		YES		YES		YES		YES		YES		YES
Industry Cluster		YES		YES		YES		YES		YES		YES
Firm Age(log), Firm Size(K)	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note 1: (Mean, SD) for MA^{U}_{okt} (log), MA^{H}_{okt} (log), MA^{D}_{okt} (log) are (0.241, 1.705), (-1.794, 1.451), (-0.171, 1.821), respectively.
Note 2: Fixed effects include industry, ownership, and county×year fixed effects.
Note 3: The industry cluster is measured by county-industry-year level local quotient.
Note 4: The firm size is captured by the firm's capital level.
Note 5: Standard errors (clustered at county-year level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$.

Table B5: Distortions' Influence on County-Industry Aggregate Performance

Panel A: Influences on the means of county-industry-year-level aggregate performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Productivity(log)		Input Price(log)		Markup(log)		Output Price(log)		Revenue(log)		Raw Markup(log)	
$\sigma^{SSratio}_{okt}$ (log)	-0.041***	-0.036***	0.047***	0.042***	-0.002***	-0.002***	0.037***	0.040***	-0.151***	-0.125***	-0.003***	-0.002***
	(0.0028)	(0.0028)	(0.0022)	(0.0022)	(0.0002)	(0.0002)	(0.0071)	(0.0071)	(0.0032)	(0.0028)	(0.0003)	(0.0003)
Observations	71,465	71,465	71,465	71,465	71,465	71,465	47,421	47,421	71,465	71,465	71,465	71,465
Adjusted R^2	0.817	0.819	0.692	0.698	0.613	0.618	0.446	0.447	0.375	0.524	0.345	0.351

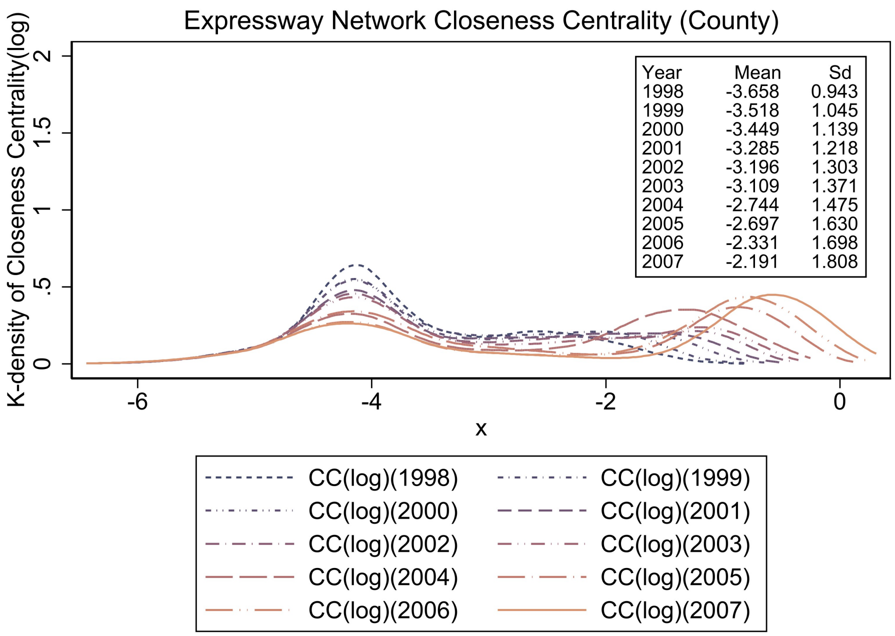
Panel B: Influences on the dispersions of county-industry-year-level aggregate performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Productivity(log)		Input Price(log)		Markup(log)		Output Price(log)		Revenue(log)		Raw Markup(log)	
$\sigma^{SSratio}_{okt}$ (log)	0.005	0.007**	0.070***	0.068***	0.049***	0.047***	0.054***	0.057***	-0.144***	-0.116***	0.046***	0.044***
	(0.0035)	(0.0035)	(0.0032)	(0.0032)	(0.0020)	(0.0020)	(0.0094)	(0.0094)	(0.0035)	(0.0032)	(0.0022)	(0.0022)
Observations	71,465	71,465	71,465	71,465	71,465	71,465	33,958	33,958	71,459	71,459	71,465	71,465
Adjusted R^2	0.737	0.739	0.544	0.546	0.197	0.201	0.478	0.478	0.334	0.458	0.222	0.223
Average Capital Intensity(log)		YES		YES		YES		YES		YES		YES
Industry Cluster		YES		YES		YES		YES		YES		YES
Average Firm Size, Firm Age(log)	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note 1: Fixed effects include industry, ownership, and county×year fixed effects.
Note 2: The industry cluster is measured by county-industry-year level local quotient.
Note 3: The firm size is captured by the firm's capital level.
Note 4: Standard errors (clustered at county-year level) in parentheses.
* $p < .10$, ** $p < .05$, *** $p < .01$.

B.2 Figures

Figure B1: All Counties' Closeness Centrality Dispersion



Note: All counties are kept here.