The Effect of Transportation on the Spatial Distribution of Jobs and firms: Evidence from South Korea

Eunjee Kwon*

July 2, 2018

Abstract

This research aims to improve upon our understanding of the effects of inter-city transits on spatial job distribution and their heterogeneous effects across firms and workers with different characteristics. To be more specific, I am going to investigate how the location decisions of workers and firms are affected by this transportation infrastructure, and in turn, how this reallocation changes the map of local economies. Despite the importance of transportation in local economies, little is known about its distributional/ aggregate impact on spatial labor markets due to the lack of detailed data with a large construction of inter-city transit systems. This paper combines extensive microdata collected by Statistics Korea and the massive expansion of high-speed train in South Korea to answer this question.

JEL Classification Codes: R40, R23

Keywords: Transportation, Regional Labor Market

*Ph.D. Student, Department of Economics, University of Southern California. email: eunjeekw@usc.edu

1 Introduction

How important are the impacts of transportation cost[1] on local labor markets and how are these impacts distributed across firms/ workers/ regions with different characteristics? How much are the efficiency gains that can be derived from this spatial reallocation of labor and the productivity improvement?

The increases in connectivity between cities brought about by the construction of inter-city transit can affect the location choice of firms and workers. Let us consider firms' locational choices with respect to their establishments[2]. When transportation costs from rural to urban areas are high, the disadvantages of firms locating their establishments in rural areas would be too large, as this would impede their access to good high-skilled labor pool (Lin (2017)) or agglomeration economies (Ellison et al. (2010)). However, once accessibilities between rural and urban areas improve, firms will have stronger incentives to locate their establishments in rural areas where they could benefit from not only cheaper rents or real estate prices but also lower cost of unskilled labor. Likewise, workers might have greater incentives to migrate to rural areas with the expansion of inter-city transit. When the costs of transportation from one city to other cities are too high, workers may have strong preferences for living in big cities, so as to enjoy better amenities, better schooling, and more job opportunities. Alternatively, if the bullet train expansion improves the amenities of less-developed cities, workers might have greater incentives to live in these less-developed cities where living costs (such as housing prices) are lower.

If firms and workers sort themselves spatially in response to the construction of the transportation infrastructure, the distribution of local economies would change as a consequence. For example, the distribution of population, labor market outcomes and the firms' productivity across cities would be affected by the locational choice of firms and workers. Moreover, the effects would differ by type of firm (e.g. industries, high-skilled/ low-skilled intensive jobs, sizes of establishments etc.), by type of worker (high-skilled/ low-skilled workers, household types with different demographic characteristics), and/or by region (rural/urban, population density, residential/industrial areas, etc).

This paper aims to investigate the effects of High-Speed Railroad(HSR)[3] on the local economies and the location choice of firms and workers. Specifically, I first examine how the bullet train expansion affects the distribution of the population and employment across cities. Next, I investigate the mechanism by looking at the entry and exit decision of firms and the workers' decisions of where to live and where to work. Finally,

[1]Marshall (1920) emphasized three different types of transport costs—the costs of moving goods, people, and ideas. In this study, I focus mainly on the costs of moving people and potentially, idea and technology, in case if people with idea or technology become more mobile due to the expansion of bullet train.

[2]Following Haltiwanger et al. (2013), I use the definitions of establishments and firms as defined by the U.S. Census Bureau. Specifically, an establishment is a specific physical location where business activity occurs, while a firm reflects all the establishments under common operational control.

[3]There is no official definition of HSR but in general it stands for a train whose speed can reach faster than 200km per hours.

this paper identifies the heterogeneity across different types of firms/ workers/ regions by considering types of industries, the workers' occupations, and the regional characteristics(e.g. sizes of the city, the closeness to Seoul, etc.).

This paper exploits the panel data sets at district(Gungu)-level which covers both before and after the opening of the HSR in South Korea. Korea Train eXpress (KTX) was first introduced in 2004 and experienced the expansions in 2010 and 2011. The ridership of KTX has increased significantly since its opening; which in 2004, accounted for 3.7% of interregional transportation modes (buses, cars, KTX or non-KTX trains), but reached 10% in 2015[4].

This massive construction of the inter-city transit together with the extensive micro-level panel data sets of workers and establishments collected by Statistics Korea becomes the laboratory where the questions previously asked can be empirically investigated. The census on Establishments and the Population and Housing Census enables me to investigate spatial labor market characteristics as well as reallocation patterns of individual firms and workers for 16 years (2000-2015). This period covers both before and after KTX expansion in all Gungus in Korea. The well-managed 16 years panel data make it possible for me to look at the aggregate and distributional effects of inter-city transit systems.

The contribution of my research will be three-fold. First, this paper provides a deeper understanding of the transportation effects. A few existing studies identify the significant effects of inter-city transportation on the different and distinctive aspects of local economies. Reductions in transportation costs can decrease trade costs and interregional price gaps, increase trade, and affect real income levels (Donaldson (2015)); reduce migration costs (Morten and Oliveira (2014)); improve market access, expand labor markets and enhance the spatial agglomeration of cities (Zheng and Kahn (2013)). Moreover, improvements in transportation infrastructure affect the locational choices of workers and firms as they change the incentives to migration by reducing commuting and shipping costs. (Baum-Snow and Turner (2017)). However, the lack of detailed data for both firms and workers coinciding with a massive construction of transit systems has impeded the analysis of the effects of improved transportation infrastructure on labor markets. To the best of my knowledge, at present, no studies consider the effects on both demand and supply sides of jobs and their incentives to move separately in transportation studies. Moreover, due to the lack of available data, few studies look at the heterogeneous effects across different types of jobs and cities.

Secondly, as the existing literature provides only an incomplete understanding of the general equilibrium effects of transportation infrastructure and hence a very little basis for welfare analysis, my research aims to analyze both partial and general equilibrium effects of the improvement in transportation, with the help

[4]Korea Transport Database of the Korea Transport Institute (KTDB)

of an appropriate empirical analysis. Tsivanidis (2018) builds a quantitative general equilibrium model of a closed city where low- and high-skill workers sort over where to live, where to work, and whether or not to own a car and evaluate the welfare gains of an 'intra-city' transit infrastructure, Bus Rapid Transit (BRT) system in Bogota, based on "commuter market access". I extend his study by considering the inter-city transit which can also affect the location choice of firms and the efficiency gains from the reallocation process.

Finally, this paper value-adds to the literature on the effects of HSR, which is a rising issue in the urban economics literature. HSR is regarded as relatively a new product amongst the inter-city transportation modes. Conventional transportation modes (e.g. trains, buses, cars etc.[5]) are for moving goods and people, whereas this new product is mostly used to move people. HSR can improve connectivity across cities, which increases the interactions between cities. Indeed, literature finds that HSR improves market access, expands labor markets, and enhances spatial agglomeration of cities in China (Zheng and Kahn, 2013). Moreover, Lin (2016) finds that the heterogeneous effects exist across industries. HSR expansion can also affect workers' decision of where to live and where to work. For example, in Germany, workers change jobs to smaller cities while keeping their place of residence in larger cities (Heuermann and Schmieder, 2018). This paper contributes to this literature by analyzing the mechanism behind these HSR effects across different types of regions, workers, and firms.

The plan of the paper is as follows. Section 2 provides background information on the bullet train expansion and the labor market in South Korea. In Section 3, data sets that I used are explained. Section 4 is about empirical methodologies, Section 5 presents the results and Section 6 concludes with my future research agenda.

2 KTX Expansion in South Korea and its Impact on the Labor Market

2.1 KTX Expansion in South Korea

Korea Train eXpress (KTX) was first introduced in 2004 and experienced massive expansions in 2010 and 2011. The construction plan of KTX was made in the 1980s, for the purpose of relieving road traffic congestions coming from the rapid economic growth in 1980s, and promoting the decentralization from Seoul[6]. This bullet train not only links Seoul and Busan[7] which are the first and second largest cities in South Korea

[5] In Korean context, domestic flights are not considered as the major inter-regional transits, as they are less efficient regarding the accessibility to the airport from each city and the size of the land. Indeed the ridership of any domestic flights (without in and out Jeju island) has been almost stagnant during the sample period of 2000-2015 (figure 2).

[6] The population of Seoul in 2016 is estimated at 10.29 million, although this is just the population of the Special City, which has a density of about 17,000 people per square kilometer (45,000/square mile). The sprawling metropolitan area is much larger at 25.6 million.

[7] Even though it is the second largest city in South Korea, the size is much smaller than Seoul, with the population at 3.4 million

but also links small cities which had poor connectivity to big cities.

Figure 1 about here

Table 1 about here

The KTX expansion process is depicted in Figure 1 and Table 1. In 2004 (stage 1), Gyeongbu-line which connects from Seoul to Daegu and the other small cities in between the two cities was introduced, and it was linked to the cities in southern part of the South Korean continent by using the electrified conventional railway of Gyeongbu- and Honam-line. In 2010-2011 (stage 2), Gyeongbu-line was extended from Daegu to Busan and some stations in Gyeongjeon-, and Jeonra-line began their operation. In 2018 (stage 3), another KTX extension is planned. In 2011, KTX is known to cover 22% of the total territory and 56% of total population of South Korea. (Data source: KOSIS).

KTX made a round-trip between any of the major cities feasible within a single-day; even travel from Seoul to Busan (417.4km), two cities located at the opposite edges of South Korea takes 6 hours by car or non-high-speed train, but takes only 2.5 hours by KTX. The train is mainly used for transporting people rather than goods. Among the KTX passengers, more than 80% of them are using KTX to visit friends or relatives or go for business trips. Less than 1% of passengers are using KTX as a commuting mode because of the expensive ticket price[8].

Figure 2 about here

As we can see from figure 2, the ridership of KTX has increased significantly since its opening; it accounted for 3.7% of interregional transportation modes (buses, cars, KTX or non-KTX train) in 2004, but reached 10% in 2015[9]. While ridership of other transportation modes (except for domestic flight) has decreased over time.

2.2 Impact on the Labor Market

Increases in connectivity between cities and reductions in transportation costs can affect the spatial distribution of firms and workers and as a result, labor market outcomes across cities. These results are derived from the responses of both demand (firms) and supply (workers) side of the labor market.

Considering the demand side of the labor market, in other word, firms. Before the KTX expansion, the disadvantages of firms locating their establishments in non-centralized areas (or rural areas) would have

[8] One-way train ticket from Seoul to Busan is around $50. Regarding that the average household net-adjusted disposable income per capita is USD 19,372 per year in 2015 (USD 1,614.33 per month), the train ticket is too expensive to be used as a daily commuting mode.

[9] Korea Transport Database of the Korea Transport Institute (KTDB)

induced most firms to locate in an urbanized and centralized area. Location in a rural area would impede firms' access to a good high-skilled labor pool or to the advantage of agglomeration economies (Ellison et al. (2010)). However, once KTX networks are expanded to rural areas, firms might consider relocating or establishing establishments in the rural areas with lower operation costs, including rents or unskilled labor costs. For example, going for a business trip from a rural area to an urban area within a day becomes possible with KTX, so cost-minimizing firms would re-consider their locational choice of establishments given the changes in their input prices (e.g. rent, transportation costs, etc).

Workers (supply side of jobs) also might have more incentive to migrate once KTX is expanded to a rural area. Before KTX expansion, workers would strongly prefer to live in a big city, due to better amenities in a big city and better job opportunities. Once KTX is introduced to a rural area, however, workers might less prefer to live in big cities where living costs including housing prices are too high, as they can live in rural areas with lower costs of living and visit the big cities more easily whenever they want thanks to KTX.

The effects of KTX expansion on the demand and the supply of jobs would change the spatial labor market equilibrium (e.g. the number of jobs and the wage level across cities). Whether the reduction in the transportation costs indeed raises the number of jobs and the wage of a local labor market depends on the magnitude of the effects. The effects might differ across different regions/ industries/ firms/ and workers. If the benefits coming from the reductions in transportation costs and are sufficiently large, then the local economy of the treated area[10] [11] is likely to be expanded.

Figure 3 about here

Figure 3 describes the trends of the outcome variables in Gungus with and without KTX. This includes the number of employment, log of employment, log of population, employment rate (=number of employment/population), log of sectoral employment (manufacturing, service, and high-skilled service), and land price of each Gungu.

As we can see in Figure 3.2 (the number of employment), before 2004, Gungus which have KTX stations and Gungus without KTX stations experienced the same increasing trends in employment. However, from 2004 when the KTX was first introduced, employment level in the Gungus with KTX stations started to increase faster than those without KTX stations. The gap between the two regions became greater in 2010 when there was an expansion in KTX stations again. From Figure 3, we can conjecture the positive effect of

[10]In this paper, the locality is defined as a "Gungu", which is an administrative unit and equivalent to a county in U.S. context

[11]The estimated results are sensitive to the market definition. According to Merfeld (2017), a district-level analysis of labor market outcome obscures substantial heterogeneity, and underestimate the policy effect in India. Regarding the availability, the local labor market is defined as a district-level. For the future work, a market definition using Geographic Information System(GIS) can be considered.

KTX on a local labor market. However, Figure 3 just depicts the positive correlation between employment and KTX expansion. To analyze the causal relationship between the two more rigorously, I am going to use econometric models. The datasets and the econometric methods used are introduced in the following sections. Whether there is a significant effect of KTX on the local economy is empirically tested by the data sets explained in the following section.

3 Data

3.1 The Census on Establishment

The Census on Establishments is an annual population survey collected from 1994, that provides business information of all 3.3 million enterprises and establishments in Korea. This administrative dataset includes each establishment's business information, such as its' 5-digit industry code, number of employees, and geographic information. This detailed business information makes it possible to construct a good panel data set about job creation[12] and job destruction[13] with detailed geographic and establishment's information. Moreover, this detailed information allows me to detect heterogeneities across cities of different sizes, different industry compositions, or even distances from the major megacities. For this paper, I am going to start with the number of employees and the number of establishments at the Gungu-level. For the further extension, I am planning to construct job creation and destruction variables at the Gungu-level.

3.2 Other Data Sources

I exploit bullet train information from the Korea Transport Database of the Korea Transport Institute(KTDB). This information contains the exact location of bullet train stations, railroad connections, and the opening date of stations.

Other control variables (e.g. total population, male/female population, GRDP of the cities, etc) are collected from different surveys collected by Statistics Korea.

By combining the data sets mentioned above, I construct a panel data sets at the Gungu-level (with the number of Gungus = 228), which covers from 2000 to 2015 (t=16). Summary Statistics of the variables is presented in Table 2.

Table 2 about here

[12]The number of jobs created each year due to either firms' entrance or expansion of employment of incumbents

[13]The number of jobs disappeared each year due to either exit of firms or declines in employment of surviving firms

4 Empirical Strategy

In this section, I describe the empirical strategy used in this paper.

4.1 Difference-in-difference

The first specification considered in this paper is a difference-in-difference(DID) estimator, which has been one of the most popular tools in program evaluation. In DID estimation, a definition of "treated" and "control" is important. As of now, I define a Gungu 'treated' if the centroid of the Gungu is located within 10km from any KTX stations[14]. Table 3 presents the number of treated and control Gungus, based on the definition. As we can see from the table 3, the opening of the KTX stations happened only in 2004, 2010, and 2011. In 2011, 75 Gungus (around 38% of the total number of Gungus) are defined as "treated", and among those, 25 Gungus are defined as "near Seoul".

Table 3 about here

The consistency of DID estimator relies on the identifying assumptions which might not hold in a lot of policy intervention cases. Specifically, DID estimator requires that treated and control groups' average outcomes followed the similar trends, prior to the treatment.

Figure 3 about here

As we can see from Figure 3, the pre-trends of the a few outcomes (log of total employment (3-a), log of population (3-b), log of service employment (3-d), log of retail employment (3-e) of the control and treatment Gungus show the parallel pre-trends, whereas log of manufacturing employment (3-c) and log of high-skilled service employment (3-f) does not satisfy the DID assumption. As of now, I will use the same difference-in-difference specification for all the outcome variables, but notice that this would not be the correct specification, especially for the log manufacturing and high-skilled service employment.

In a Gungu j within a Sido s[15] at time t, the difference-in-difference model which estimates the impact of KTX expansion on outcome variables is as follows:

$$y_{j,s,t} = \alpha + \beta Treat_{j,s} * Post_t + \delta X_{j,s,t} + \eta_s * year_t + f_{j,s} + year_t + \epsilon_{j,s,t} \quad (1)$$

where $y_{j,s,t}$ is a dependent variable (number of employees, number of establishments or education level, wage etc.), $Treat_{j,s}$ takes 1 if the centroid of a Gungu j of a Sido s is located within 10km of any KTX stations and $Post_t$ takes 1 for the year when the KTX station is first opened. $X_{j,s,t}$ is a set of control

[14]The distance is a linear distance from a centroid of a Gungu to the nearest KTX station. The longitude and latitude information is obtained from Google Maps API.

[15]Sido is the bigger administrative district division in South Korea. There are 15 Sidos in South Korea and within Sidos, there are around 280 Gungus

variables. I am planning to estimate the model with more control variables (e.g. average income-level of a Gungu, housing prices of a Gungu, education-level, etc.) for robustness check. *Gungu fixed effect* ($f_{j,s}$) and *year fixed effect* ($year_t$) are included in all specifications. More importantly, I include the *Sido-time trend effect* ($\eta_s * year_t$) in all specification, which captures any time-varying omitted variables at the Sido-level.[16] Finally, all the standard errors are clustered at the Gungu-level.

Here, β captures a treatment effect. As it is described in section 2, if there is a sufficient reduction in transportation costs, migration of workers and reallocation of establishments would be likely to happen to induce a positive β coefficient. The treatment effects might differ across different industries, different sizes of firms, or rural and urban areas, and these potential heterogeneity effects are presented in the next section.

4.2 Market Potential Variable

As noted by a large body of literature, studying the effect of the transportation needs careful inspection because of the selection issues. The central inference problem that researchers are facing in the transportation literature is that transportation is not assigned randomly, but rather it is determined based on both observed or unobserved location characteristics and forecasting of expected benefits. (Redding and Turner (2015)). In our context, even if the location of the high-speed train stations in Korea was determined by the central government 14 years earlier than the actual opening and the plan was known as not endogenously modified afterward in response to location-specific economic conditions, still, possible endogeneity exists. For example, Gungus that expect to gain the most are more likely to seek treatment, by lobbying more aggressively which can raise the selection issues. This makes location choice non-random and these endogeneities are not fully controlled by observable characteristics.

A common approach to address this issue in the literature is to create a “Market Potential(or Market Access)”. The theoretical background of this measure is suggested by Donaldson and Hornbeck (2016) and according to their paper, the total impact of transportation on each region is captured by changes in that county’s “market access,” a reduced-form expression derived from general equilibrium trade theory.

According to the paper, the first order approxiamtion of market potential measure of a location j can be defined as:

$$MarketPotential_j = \Sigma_{k \neq j} Y_k e^{-d_{jk}} \tag{2}$$

where Y_k is population of a location k and d_{jk} is distance between j and k (or timed travel). This measure

[16] The ideal time trend effects which control for the unobserved variables are Gungu-time trend effect. However, due to the issue of the degrees of freedom, the Gungu-time trend effect cannot be included. Instead, I include the sido time trend effect to capture the time varying omitted variables which is common within a Sido.

is a gravity-based measure, which captures the nearness to the economic activities of a location j.

There are a few advantages of using this measure, compare to the DID estimate presented in the previous section. One advantage of the market potential measure is that this is a continuous measure. In my analysis, this can give some sense of heterogeneous effects of KTX expansion. In other words, this can capture the magnitude of the benefits that treated Gungu gets, which cannot be captured by the DID estimates.

Table 4 about here

The other advantage of this measure is that this does not make a clear distinction of the treated and control which frees us from the issue of SUTVA assumption. The violation of STUVA is often a concern when there are potential spillover effects to the control group. However, with this market potential measure, the indirect spillover effect to control groups can be captured. Indeed, as we can see from Table 4, control groups' market potential improve on average by 3.8% from 2011 to 2004. This is showing that there exists the spillover effect of the KTX station. Still, the magnitude of the benefits in treated Gungus is much bigger on average, which is consistent with our belief.

This measure enables me to investigate how much of the market potential in each Gungu were improved after the KTX expansion and what is its impact on the outcome of interests[17].

In a Gungu j within a Sido s at time t, the econometrics model estimates the impact of KTX expansion on outcome variables is as follows:

$$y_{j,s,t} = \alpha + \beta MarketPotential_{j,s,t} + \delta X_{j,s,t} + \eta_s * year_t + f_{j,s} + year_t + \epsilon_{j,s,t} \tag{3}$$

where $y_{j,s,t}$ is a dependent variable. For the time being, I regard the log of the number of employees, log of the population, and log of the sectoral employment of a Gungu j within a Sido s at time t. $Market\_Potentialj, s, t$ is calulated based on the equation (2). Same as previous specification in equation (1), a set of control variables ($X_{j,s,t}$), *Gungu fixed effect* ($f_{j,s}$), *year fixed effect* ($year_t$) and *Sido-time trend effect* ($\eta_s * year_t$) are included in all estimations. Again, all the standard errors are clustered at Gungu-level.

[17] As of now, I am using the market potential measure calculated by Kim and Sultana (2015). They construct a railroad by using ArcGIS network software and calculated the minimum timed travel from all cities to the other cities. In their study, however, they only include the period after 2004. Therefore, the treatment effects that are measured using this measure might be more conservative than the treatment effects estimated including periods before 2004.

5 Result

5.1 DID Result: Effect of KTX on Local Employment and Population

The DID estimation results are presented in table 5, table 6 and table 7. Start with table 5, overall, KTX does not have homogeneous effects on the local employmentor local population. (column (1), column (3)). However, once we include the interaction term with "near Seoul"(column (2), column (4)), the population of treated Gungus near Seoul decreases by 8.7%, once it is treated, whereas the employment is not significantly affected in the non-Seoul treated region. This means that the decentralization of population happens with the KTX expansion. If we see the heterogeneous effects across different industries in table 6, the sign of the coefficients are consistent with the decentralization stories, whereas the significances are only in manufacturing and high-skilled service. In table 7, the effects on the number of establishments show qualitatively the same pattern as the number of employment, with differences in the statistical significance.

Table 5 about here

Table 6 about here

Table 7 about here

5.2 Market Potential Results

The results using the market potential variables are presented in table 8. Each coefficient can be interpreted by the percentage changes of the outcome variables when the market potential measure increases by 1%[18]. The decentralization story can be applied to every variable (i.e. log(Market Potential) without the near Seoul interaction term shows positive coefficient and the interaction terms show negative with significance). However, the magnitudes and the statistical significances are different across different outcome variables. For example, the treated Gungus near Seoul lose employment (column (1)), manufacturing employment (column (3)), and skilled service employment (column (6)) whereas for those variables, the treated Gungus not near Seoul were not affected. Whereas population increases in non-Seoul area (column (2)), service employment increases in non-Seoul and decreases in Seoul area (column (4)) and finally, retail and wholesale employment is not affected by the changes in market potential.In table 9, the effects on the number of establishments shows qualitatively the same pattern as the number of employment, with differences in the statistical significance.

Table 8 about here

Table 9 about here

[18]Recall that in Table 4, the market potential of treated(control) groups on average increases by 5.7%(3.8 %) from 2004 to 2011.

6 Conclusion

This paper investigates the effect of interregional transportation on the local economies. So far, the empirical results show that the improvement in connectivity can change the distribution of employment and population. Some evidence of the decentralization pattern appeasr; the treatment effect is positive in the non-Seoul treated area, whereas is negative in near Seoul area.

So far, the analysis is just focusing on the "net" effect, which might mask the underlying mechanism of the KTX effect. For example, if the local economy of a treated Gungu goes through the creative destruction process of employment, (i.e. less competitive workers are replaced by the competitive workers after a KTX station is introduced in the Gungu) the number of employment might not be changed whereas the productivity of the economy would improve.

Disentangling the aggregated net effects from hidden establishments entry and exit decision will enhance the understanding of the distributional/ aggregate effects of transportation on local economies. Moreover, to understand the locational choice of workers with different occupations, education-level, and demographics will give us some evidence to understand the mechanism behind this. To implement the research plan, I am going to use the Microdata Integrated Service (MDIS), which enables me to access the administrative raw data sets collected by Statistics Korea. The underlying mechanism hidden under the "net" effect can be investigated as follows.

6.1 Entry and Exit of Firms and Productivity Changes in Service Sector

One of the key questions to be investigated is whether the KTX expansion creates creative destruction, by driving unproductive establishments out and bringing productive establishments into the local market. Foster, Haltiwanger, and Krizan (2006) find that the U.S. retail sector underwent a massive restructuring and reallocation in the 1990s with the technological advances, which induced the labor productivity growth by more productive entering establishments displacing much less productive exiting establishments. If this reallocation process happened accompanying the expansion of the KTX station, even if the number of establishments or employment might not be changed as a net sense, this could be viewed as a positive effect of KTX on the local economies.

The first step to examine this reallocation process is to see how this KTX expansion affects local establishments' entry and exit decision. The entry and exit variables can be constructed with the Census on Establishments, which is an annual establishment-level survey started in 1994. With the MDIS data, I can construct the panel data set at the establishment-level which tracts the annual economic activities of every establishment. Specifically, my interest lies in the entry, exit, and relocation decisions of each establishment in all Gungus, before and after KTX expansion.

The next step is whether the entrants are more productive ones compare to the exiters. This can be observed by a panel data set constructed based on the Service Industry Survey (1996,2001,2005) and Economic Census (2010,2015). These 5-year-gap surveys collect the detailed business information data on the revenue and input costs. With the information, the labor productivity can be measured by the log difference in establishment-level real revenue and the total labor hours used.

Figure 1: KTX network in South Korea (Kim and Sultana, 2015)

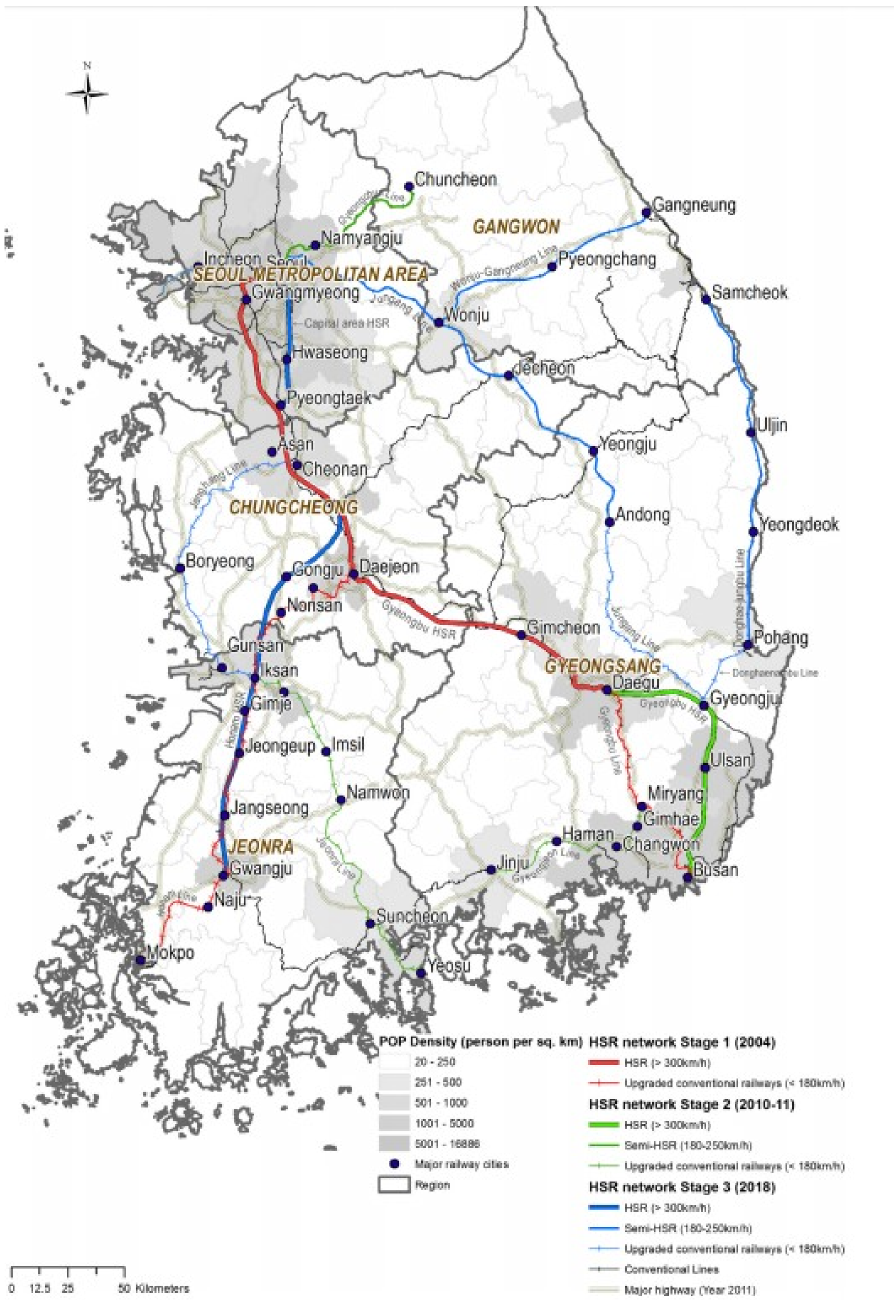


Fig. 1. Railway network in South Korea by each stage of HSR.

Figure 2: Annual Ridership across Different Transportation Modes

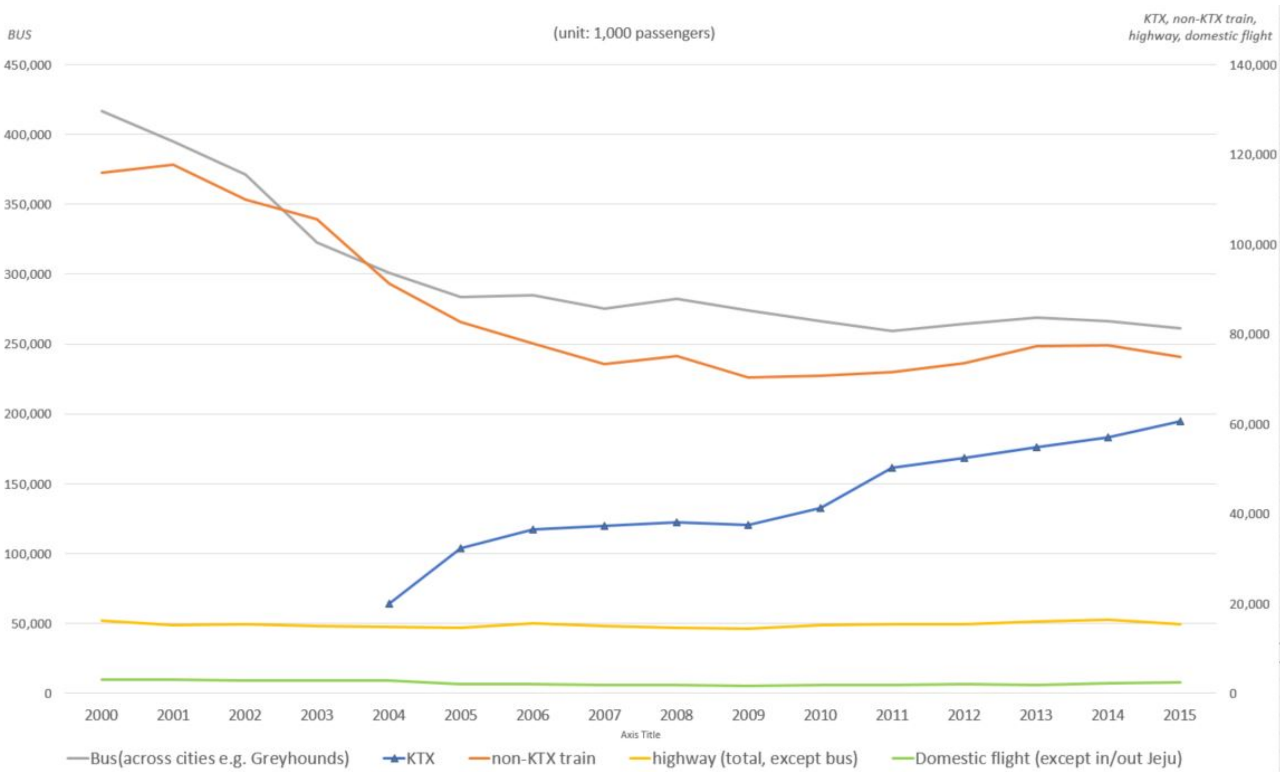
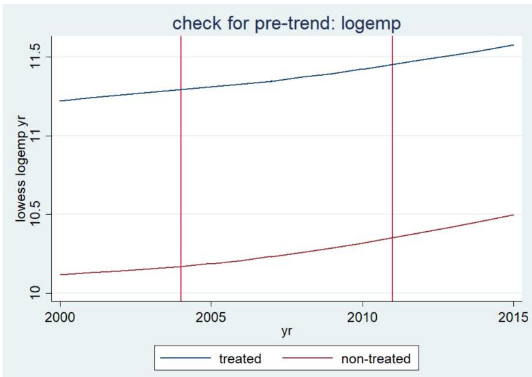
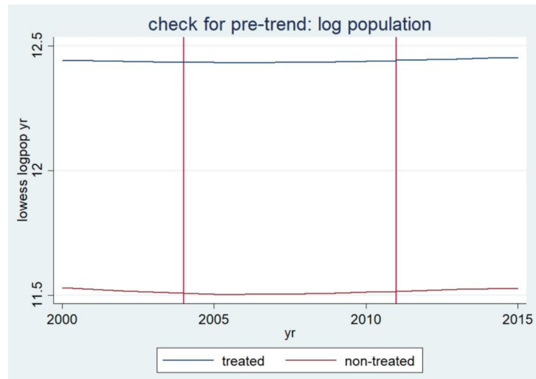


Figure 3: DID: Check for pre-trend

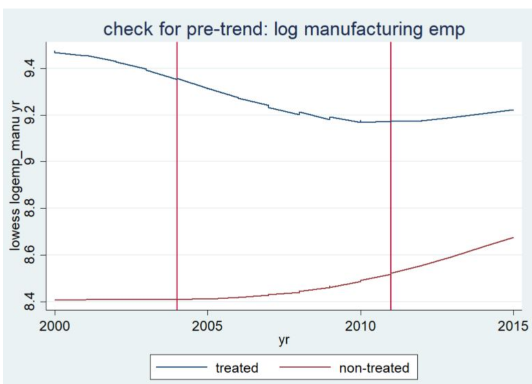
(a) Log Employment



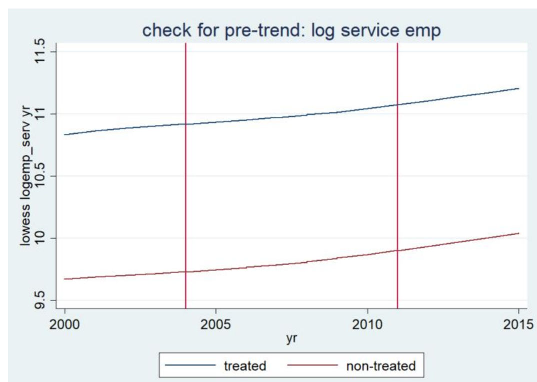
(b) Log Population



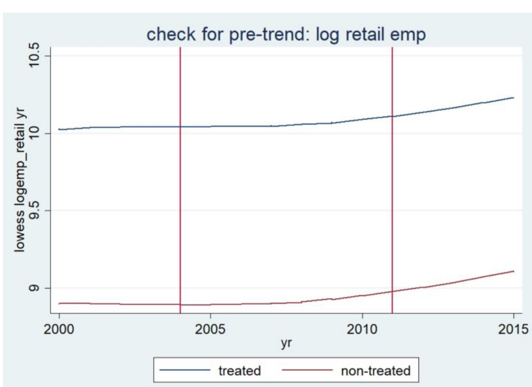
(c) Log Employment: Manufacturing



(d) Log Employment: Service



(e) Log Employment: Retail



(f) Log Employment; High-Skilled Service(Finance, Real-estate, Science)

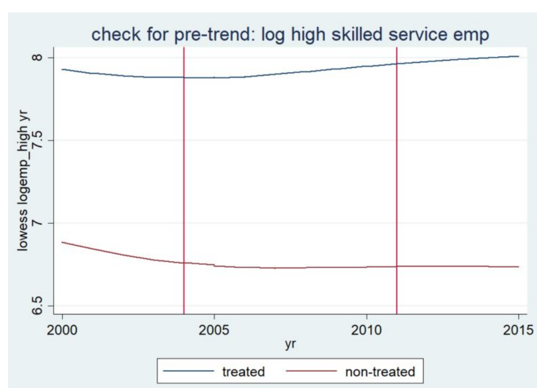
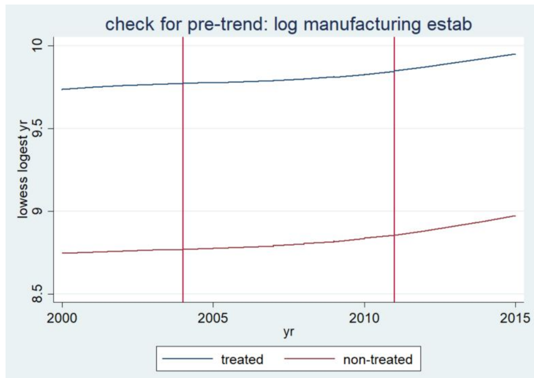
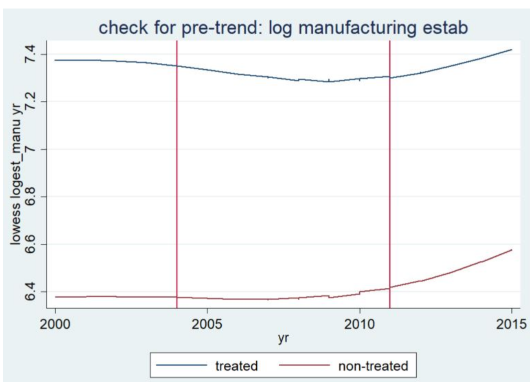


Figure 4: DID: Check for pre-trend

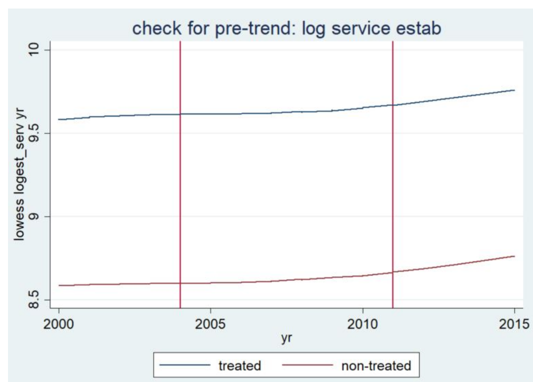
(a) Log Establishment



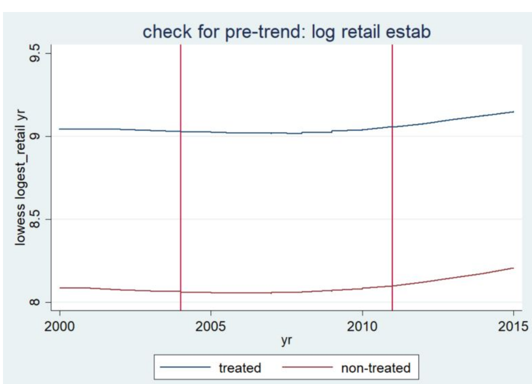
(b) Log Establishment: Manufacturing



(c) Log Establishment: Service



(d) Log Establishment: Retail



(e) Log Establishment; High-Skilled Service(Finance, Real-estate, Science)

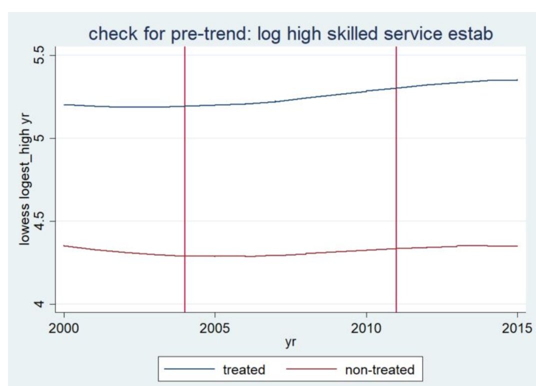


Table 1: History of the KTX Station Expansion

Date	Line	From/To	Project Type	Station
Apr/2004	Gyeongbu*	Seoul/Daegu	New line	Seoul, Hangshin, Youngdengpo, Gwangmeng Suwon, Cheonan-Asan, Daejeon, Dong-Daegu
		Daegu/Busan	New line	Gupo, Milyang, Busan
	Honam**	Daejeon/Mokpo	New line	Yongsan, Seo-Daejeon, Gyeryong, Nonsan Iksan, Jeongeup, Gwangju-Songjeong, Naju, Mokpo
Dec/2010	Gyeongbu*	Seoul/Daegu	Add-station	Osong
		Daegu/Busan	Add-station Improve speed	Shin-Gyeongju, Gimcheon(Gumi), Ulsan Busan, Milang, Gupo
	Gyeongjeon**	Milyang/Masan	New line	Jinyoung, Changwon-Joongang, Changwon, Masan
Oct/2011	Jeonla**	Ilsan/Yeosu	New line	Jeonju, Namwon, Guryegu, Sooncheon, Yeocheon Yeosu-expo, Goksung
Dec/2012	Gyeongjeon**	Masan/Jingu	Add-station	Jinju

*: High Speed Railroad (speed more than 200km)
**: Electrified conventional railway directly connected with HSR
Source: Korail

Table 2: Summary Statistics: Dependent Variables

Variable		Mean	Std.Dev.	Min	Max
Population	Level (Population)	215212.2	205098.7	9191	1184624
	Log (Population)	11.82129	1.001277	9.12598	13.98494
No. of Employment	Level (Total)	72054.07	80359.93	2847	711278
	Log (Total)	10.63328	1.102006	7.954021	13.47482
	Level (Manufacturing)	15367.1	22253.47	69	211329
	Log (Manufacturing)	8.73744	1.426526	4.234107	12.26117
	Level (Service)	47157.99	53052.03	2246	481289
	Log (Service)	10.21048	1.090624	7.716906	13.08422
	Level (Retail&WholeSale)	19160.36	21621.98	919	207900
	Log (Retail&WholeSale)	9.320025	1.082501	6.823286	12.24481
	Level (High skilled serv)	2852.754	5643.144	57	69872
	Log (High skilled serv)	7.142252	1.201655	4.043051	11.15442
No. of Establishments	Level (Total)	14504.72	13337.32	908	83187
	Log (Total)	9.148477	.9780885	6.811244	11.32885
	Level (Manufacturing)	1496.667	1856.202	31	16481
	Log (Manufacturing)	6.713264	1.114403	3.433987	9.709964
	Level (Service)	12177.61	11159.37	836	67223
	Log (Service)	8.971045	.9829047	6.728629	11.11577
	Level (Retail&WholeSale)	6707.412	6186.377	527	45750
	Log (Retail&WholeSale)	8.406758	.936166	6.2672	10.73095
	Level (High skilled serv)	163.6816	191.5186	7	1757
	Log (High skilled serv)	4.626556	.9446152	1.94591	7.471363

Table 3: Number of Treated and Control Gungus

Year	Control	Treat		Total
2004	163	65		228
		25 (Near Seoul)	40 (Not Seoul)	
2011	153	75		228
		25 (Near Seoul)	50 (Not Seoul)	

Table 4: Changes in Market Potential: Treated and Control Gungus

Year	Control	Treat	Difference (T-C)
2004	593322	729014.7	22.8%
2011	615879.4	770214	25%
Change (2011-2004)	3.8%	5.7%	

Table 5: KTX Effect on Local Employment and Local Population

VARIABLES	(1) log(employment)	(2) log(employment)	(3) log(population)	(4) log(population)
Dummy=1 if KTX stn in 10km	0.005	0.025	0.004	0.032
	(0.023)	(0.029)	(0.023)	(0.027)
* Near Seoul		-0.073*		-0.100**
		(0.043)		(0.041)
Year Fixed Effect	yes	yes	yes	yes
Gungu Fixed Effect	yes	yes	yes	yes
*Sido * Year Fixed Effect*	yes	yes	yes	yes
Observations	3,642	3,642	3,642	3,642
R-squared	0.693	0.695	0.321	0.330
Number of code2011n	228	228	228	228
adj_r2	0.673	0.675	0.276	0.285

Notes: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable of column (1), (2) is log of the number of employment of each Gungu and column (3), (4) is log of the population of each Gungu. Treatment (Dummy=1 if a station is in 10km) equals to one if a linear distance from a Gungu centroid to any KTX stations is within 10km in a given year. Near Seoul = 1 if a Gungu is within 25km from Seoul. The standard errors are clustered at Gungu-level.

Table 6: KTX Effect on Local Employment: Heterogeneity across Industries

VARIABLES	(1) log(employment) (Manufacturing)	(2) log(employment) (Service)	(3) log(employment) (Retail&Wholesale)	(4) log(employment) (Skilled Service)
Dummy=1 if KTX stn in 10km	-0.051	0.030	0.028	0.101**
	(0.046)	(0.032)	(0.031)	(0.043)
* Near Seoul	-0.261***	-0.084*	-0.093*	-0.139
	(0.097)	(0.051)	(0.048)	(0.086)
Year Fixed Effect	yes	yes	yes	yes
Gungu Fixed Effect	yes	yes	yes	yes
*Sido * Year Fixed Effect*	yes	yes	yes	yes
Observations	3,642	3,642	3,642	3,642
R-squared	0.462	0.695	0.483	0.215
Number of code2011n	228	228	228	228
adj_r2	0.427	0.675	0.449	0.162

Notes: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable of column (1), (2) is log of the number of employment of each Gungu and column (3), (4) is log of the population of each Gungu. Treatment (Dummy=1 if a station is in 10km) equals to one if a linear distance from a Gungu centroid to any KTX stations is within 10km in a given year. Near Seoul = 1 if a Gungu is within 25km from Seoul. High skilled service is defined as finance, insurance, and science and skilled service(KSIC 9th wave: K,m). The standard errors are clustered at Gungu-level.

Table 7: KTX Effect on the Number of Establishments

VARIABLES	(1) log(no. estab) (Total)	(2) log(no. estab) (Manuf)	(3) log(no. estab) (Service)	(4) log(no. estab) (Retail&Wholesale)	(5) log(no. estab) (Skilled Serv)
Dummy=1	0.040	0.049	0.038	0.036	0.097**
if KTX stn in 10km	(0.030)	(0.041)	(0.030)	(0.028)	(0.038)
* Near Seoul	-0.097**	-0.157***	-0.099**	-0.086**	-0.141**
	(0.042)	(0.059)	(0.042)	(0.042)	(0.068)
Year Fixed Effect	yes	yes	yes	yes	yes
Gungu Fixed Effect	yes	yes	yes	yes	yes
*Sido * Year Fixed Effect*	yes	yes	yes	yes	yes
Observations	3,642	3,642	3,642	3,642	3,642
R-squared	0.552	0.369	0.502	0.435	0.279
Number of code2011n	228	228	228	228	228
adj_r2	0.522	0.327	0.469	0.397	0.231

Notes: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable of log of the number of establishments (total, manufacturing, service, retail and wholesale, skilled service for column (1)-column (5), respectively) of each Gungu. Treatment (Dummy=1 if a station is in 10km) equals to one if a linear distance from a Gungu centroid to any KTX stations is within 10km in a given year. Near Seoul = 1 if a Gungu is within 25km from Seoul. High skilled service is defined as finance, insurance, and science and skilled service(KSIC 9th wave: K,m). The standard errors are clustered at Gungu-level.

Table 8: Effect of Market Potential on Local Employment, Population and Sectoral Employment

VARIABLES	(1) log(emp)	(2) log(pop)	(3) log(emp) (Manu)	(4) log(emp) (Serv)	(5) log(emp) (Retail&Wholesale)	(6) log(emp) (Skilled Serv)
log(Market Potential)	0.127	0.272***	0.183	0.179*	0.070	0.069
	(0.092)	(0.099)	(0.179)	(0.097)	(0.087)	(0.195)
* Near Seoul	-0.688**	-0.145	-2.104**	-0.743**	-0.106	-1.434***
	(0.288)	(0.294)	(0.922)	(0.329)	(0.228)	(0.479)
Year Fixed Effect	yes	yes	yes	yes	yes	yes
Gungu Fixed Effect	yes	yes	yes	yes	yes	yes
*Sido * Year Fixed Effect*	yes	yes	yes	yes	yes	yes
Observations	2,706	2,706	2,706	2,706	2,706	2,706
R-squared	0.763	0.258	0.460	0.753	0.591	0.161
Number of code2011n	226	226	226	226	226	226
adj_r2	0.748	0.209	0.425	0.737	0.565	0.106

Notes: *** p<0.01, ** p<0.05, * p<0.1. the dependent variable of column (1), (2) is log of the number of employment of each Gungu and column (3), (4) is log of the population of each Gungu. Treatment (Dummy=1 if a station is in 10km) equals to one if a linear distance from a Gungu centroid to any KTX stations is within 10km in a given year. Near Seoul = 1 if a Gungu is within 25km from Seoul. High skilled service is defined as finance, insurance, and science and skilled service(KSIC 9th wave: K,m). The standard errors are clustered at Gungu-level.

Table 9: Effect of Market Potential on the Number of Establishments

VARIABLES	(1) log(estab) (Total)	(2) log(estab) (Manu)	(3) log(estab) (Serv)	(4) log(estab) (Retail&Wholesale)	(5) log(estab) (Skilled Serv)
log(Market Potential)	0.112	0.189	0.123	0.089	0.062
	(0.102)	(0.199)	(0.094)	(0.088)	(0.191)
* Near Seoul	-0.088	-0.482	-0.089	-0.077	-0.813**
	(0.214)	(0.330)	(0.223)	(0.235)	(0.325)
Year Fixed Effect	yes	yes	yes	yes	yes
Gungu Fixed Effect	yes	yes	yes	yes	yes
*Sido * Year Fixed Effect*	yes	yes	yes	yes	yes
Observations	2,706	2,706	2,706	2,706	2,706
R-squared	0.631	0.452	0.576	0.539	0.345
Number of code2011n	226	226	226	226	226
adj_r2	0.606	0.416	0.548	0.509	0.302

Notes: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable of log of the number of establishments (total, manufacturing, service, retail and wholesale, skilled service for column (1)-column (5), respectively) of each Gungu. Treatment (Dummy=1 if a station is in 10km) equals to one if a linear distance from a Gungu centroid to any KTX stations is within 10km in a given year. Near Seoul = 1 if a Gungu is within 25km from Seoul. High skilled service is defined as finance, insurance, and science and skilled service(KSIC 9th wave: K,m). The standard errors are clustered at Gungu-level.