Unintended Pathways: The Impact of High-Speed Rail on Gender Differences in the Local Labor Market - Evidence from South Korea*

Eunjee Kwon† Narae Lee‡

July 15, 2024

Abstract

This paper examines the role of High-Speed Rail (HSR) in reducing the gender gap in labor market outcomes in South Korea. The country's notable gender disparity, coupled with the extensive HSR network starting from 2004 that connects Seoul metropolitan areas with non-Seoul regions, presents a distinctive setting for this study. Using a staggered difference-in-differences method, we evaluate HSR's impact on district-level outcomes from 2000 to 2015. The empirical analysis demonstrates that the gender employment gap in labor markets of non-Seoul (i.e., less developed) regions connected by HSR has diminished following the rail expansion. We provide empirical evidence that this decline is mainly due to the increased labor demand in local service sectors around the stations, which rely more on the movement of people and predominantly employ women. Additionally, we provide evidence that improved local amenities, particularly childcare services, may have facilitated women's labor force participation in non-Seoul areas. Our results highlight how infrastructure projects like HSR can inadvertently benefit certain demographics by reshaping the local labor market and urban structures.

*This paper was originally distributed as 'Why do improvements in transportation infrastructure reduce the gender gap in South Korea?'.

Eunjee Kwon is grateful to Matthew E. Kahn, Richard K. Green, and Paulina Oliva for their guidance and encouragement as generous advisors. This paper has benefited substantially from Gabriel M. Ahlfeldt, Vittorio Bassi, Jan Brueckner, Juan Pablo Chauvin, John S. Earle, Matthew Freedman, Miquel-Àngel Garcia López, Marcel Henkel, Naoru Koizumi, Maurice D. Kugler, Jeffery Lin, Pierre Magontier, Andrii Parkomenko, Tanner Regan, Jorge de la Roca, Sita N. Slavov, Elisabet Viladecans-Marsal, and Siqi Zheng, as well as numerous seminar/conference participants. A research grant from Lusk Real Estate Center and the University of Cincinnati Real Estate Center financially supported this paper. All errors are our own.

†Assistant Professor in Real Estate, Lindner College of Business, University of Cincinnati, kwonee@ucmail.uc.edu

‡Impact Evaluation Specialist, Development Impact Evaluation (DIME), World Bank Group (IBRD), nlee3@worldbank.org

1 Introduction

The gender gap in the labor market narrows with economic development, but the disparities remain prominent even in developed countries such as the US, Japan, and South Korea ((Blau and Kahn, 2017)). South Korea, in particular, is known as one of the countries with the highest levels of gender inequality among the developed world, showing a 37.2% wage gap and a 22.2 percentage points difference in labor force participation rates between men and women (Figure A1, which is unprecedented given the narrowed education gap between genders. The costs associated with labor force participation for women are generally higher, mainly due to the substantial burden of childcare (Maurer-Fazio et al. (2011)) and responsibilities for home production (Greenwood, Seshadri, and Yorukoglu (2005)). Consequently, women tend to work fewer hours, seek jobs with more flexible arrangements, and choose different industries and occupations than men (Goldin and Katz (2016)).

Previous literature has shown that infrastructural and technological progress can improve the gender gap in labor market outcomes. For example, access to energy, such as electrification, in household sectors reduces the domestic burden on women (Dinkelman, 2011; Greenwood et al., 2016; Vidart, 2020) and helps decrease the gender gap in the labor market by encouraging them to join the labor force. On the other hand, robotization unintentionally increases the gender pay gap, since jobs that are male dominant disproportionately benefit from robots, ultimately increasing gender disparities in the labor market (Aksoy, Berkay, and Philip (2019)). These findings imply that technological affect men and women differently in the labor market, given the distinct labor market constraints they face and the different industries and occupations they choose.

Extensive literature has documented that transport infrastructure that reduces travel costs for people changes overall spatial distribution and local labor market landscapes, however, there is limited evidence regarding its effects on reshaping gender disparities in the labor market. This study expands the literature by investigating whether improvements in non-commuter trains could have unintentionally different labor market impacts on men and women. We investigate the case of South Korea's High-Speed Rail (HSR[1],[2]), which serves as inter-city transit and not primarily used as a commuter train in the South Korean context, facilitates most trips for visiting relatives and family, leisure and business, or enjoying services in other cities. Given the substantial government spending involved in such infrastructure investment[3], understanding its heterogeneous impacts across different demographic groups is necessary.

The construction of HSR in South Korea provides a distinct setting to explore its unintended impacts on gender disparities in the labor market. As an advanced economy still grappling with labor market gender disparities, South Korea offers an ideal testing ground to study whether transportation infrastructure can have gender-nuanced effects. In particular, the country has a highly gender-segmented labor market, where women predominantly occupy roles in local service sectors that generally provide flexible work arrangements and do not require particular work experience. Furthermore, the HSR connects the heavily concentrated and congested Seoul metropolitan area with non-Seoul metropolitan areas, adding significant external validity to our findings. The estimation results in this paper could be especially relevant for similarly urbanized and concentrated economies looking to understand the broader implications of such infrastructure projects.

We examine the impact of the installation of HSR stations on gender disparities in labor market outcomes, as well as on economic activities in districts closer to the new stations. HSR first launched in 2004, linking the Seoul metropolitan area with the rest of the peninsula, and has expanded its network over the years. In this analysis, we focus on the period from

[1]High-Speed Rail (HSR) is a fast-speed intercity railway mainly used for passenger carriage. Invented in 1964, the railway is expanding worldwide from advanced economies as first adopters. Although there is no international standard, a rail system with an average speed beyond 120 miles or 200 kilometers per hour is considered an HSR. Most newly built systems have speeds above 140 miles or 300 kilometers per hour.

[2]The literature finds that HSR enhances interaction among scientists living in different cities (Dong, Zheng, and Kahn (2020)) and enhances the spatial agglomeration of cities (Zheng and Kahn (2013); see Qin (2017) for Chinese cases; for German examples, see Ahlfeldt and Feddersen (2018)). HSR also affects the location decisions of firms (Charnoz, Lelarge, and Trevien (2018) for French firms) and has heterogeneous impacts across industries (Lin (2017)). Furthermore, expansion of the HSR can affect workers' decisions about where to live and where to work. For example, in Germany, workers prefer jobs in smaller cities while residing in larger towns (Heuermann and Schmieder (2019)).

[3]For example, the HSR in California, which links Los Angeles–San Francisco–Las Vegas, is estimated to cost $7 billion. The total cost to construct HSR in South Korea, linking Seoul to Busan and several other major cities, is $15 billion (2015).

2000 to 2015 at the district level. In our preferred specification, we define a 'treated' district as a district with a centroid within 30 minutes of any HSR station. Since HSR station installations have a 'staggered' nature as they are installed over the years, we mainly employ a staggered Difference-in-Differences (DID) design using Callaway and Sant'Anna (2021) as our main estimator.

Our estimation results indicate that South Korea's HSR played a significant role in decreasing the gender gap in employment in districts closer to the stations, particularly in the non-Seoul metropolitan area. This reduction is primarily driven by sectoral reallocations, with significant increases in local service sectors around the stations, which rely more on the movement of people and predominantly employ women. Conversely, male-majority industries such as manufacturing and construction have experienced declines near HSR stations. Additionally, improved local amenities are observed in both Seoul and non-Seoul areas; however, only in non-Seoul areas are amenities like childcare services available, which may have facilitated greater female labor force participation.

Additionally, we document a notable trend of women relocating closer to HSR-connected areas, suggesting that enhanced mobility and access to amenities play a crucial role in shaping labor market outcomes. These results underscore the broader implications of infrastructure projects, highlighting how they can inadvertently benefit specific demographics and reshape local economies.

The study aligns with and expands upon existing literature on the impact of inter-city transportation infrastructure and its varied effects across different spatial and socio-economic contexts (Duranton, Morrow, and Turner, 2014; Redding and Turner, 2014; Morten and Oliveira, 2016; Baum-Snow et al., 2017; Donaldson, 2018; Tsivanidis, 2018; Severen, 2018). It delves into the crucial, yet less examined aspect of cost of moving people, offering insights into how transportation infrastructure can influence labor market outcomes beyond the traditional focus on commuting costs. By integrating perspectives from urban economics and transportation literature on gender differences in location choices and commuting decisions (Rosenthal and Strange, 2012; Black, Kolesnikova, and Taylor, 2014; Chauvin, 2017; Kawabata and Abe, 2018; Farré, Jofre-Monseny, and Torrecillas, 2020; Le Barbanchon, Rathelot, and Roulet, 2020; Liu and Su, 2020), this research highlights the change-generating potential of HSR in reducing gender disparities through labor demand and supply mechanisms that extend beyond mere commuting.

Additionally, our study contributes to the discussion on gender disparities and economic structures by investigating the unintended effects of HSR advancements on gendered labor market dynamics. It extends the existing literature on the differential impacts of structural changes on male and female labor market outcomes, previously explored through the lens of technological shocks from both the supply (Greenwood et al., 2016; Vidart, 2020) and demand (Aksoy, Berkay, and Philip, 2019) perspectives. With the case of South Korean HSR, this research uncovers how additional transport modality mitigate gender disparities even when not targeting the improvement in commuting. Our results demonstrate that HSR bolsters labor demand in sectors with higher female employment, (akin to the mechanisms observed in the Bartik shock (Bartik, 1991)), while also facilitating women's labor supply by improving local amenities, especially in education and childcare. This dual focus on demand and supply sides enriches the understanding of transportation technological advancements' implications on gendered labor dynamics, a dimension that bridges transportation economics with gender-focused urban economic analysis in a novel way.

The rest of this paper is structured as follows. Section 2 provides contextual and infrastructure policy information, including the South Korean economic geography, gender gaps, and the expansion of HSR. Next, Sections 3 and 4 respectively detail the data and empirical strategies employed in this study. The estimation results are summarized in Section 5, and Section 6 empirically estimates and discusses some of the driving mechanisms behind these effects. Lastly, Section 7 concludes the paper.

2 Background

2.1 Gender Gap in Labor Market

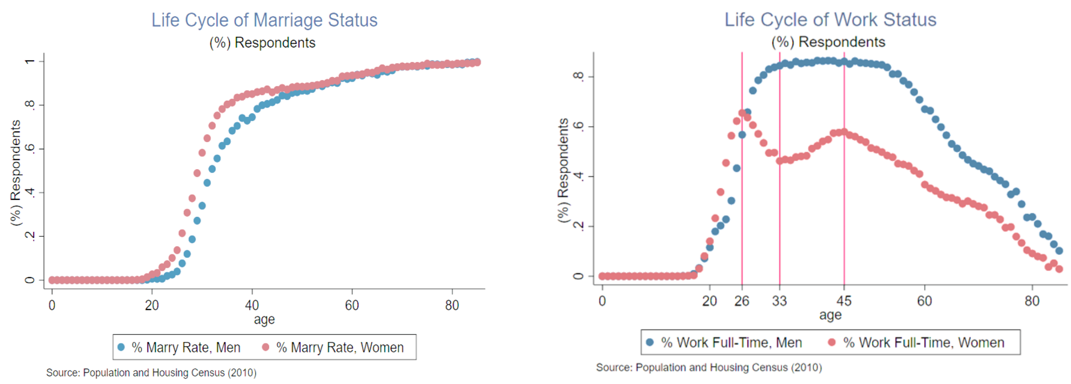
Among OECD countries, South Korea stands out as one of the most gender unequal countries in terms of labor market outcomes (Figure A1). Such disparity may result from complexsocioeconomic factors, such as heavier female domestic work expectation and childcare, discrimination, or social norms. Figure A2 presents the gender differences in time use over the years (Korea Income and Labor Panel Survey, 2000-2015). Female employment increased slightly, but the gender gap has not narrowed as much. Around 80% of men reported that they are 'working full-time' whereas only 40-50% of women responded likewise. Another 40-50% of women reported 'doing domestic work' as their main time use. Another 10% of women paralleled domestic work and part-time labor. Domestic burdens do not appear to be a significant concern for men, with less than 5% of men reporting domestic work as their primary activity.

Marriage and Work Status Exploring the reasons for differing labor market decisions of women and men in South Korea is particularly intriguing, given the shrinking education gap between genders. Figure 1 illustrates the life cycle patterns of male and female marriage and work status. Initially, men and women exhibit similar labor force participation rates until their mid-twenties. However, a divergence occurs in their mid-twenties as women start to get married and have children, leading many to exit the labor market. Although some women re-enter the workforce in their mid-thirties, the numbers do not recover nor catch up to those of their male counterparts.

Furthermore, upon re-entering the labor market after marriage, most women find themselves being in different positions compared to when they exit the market. Figure 2 displays the occupational composition of each age group for men and women before their exit from the market (ages 25-34) and upon re-entry (ages 35-44). In the earlier stages of their lives (ages 25-34), men and women share a similar distribution across occupations, typically holding white-collar jobs such as managers, professionals, or office workers. However, upon their return to the workforce between the ages of 35-45, a notable shift occurs for women, with a significant increase in employment within sales and local services as illustrated in Figure 2, (b) and (d). This pattern is not observed among men, whose occupational distribution remains relatively unchanged. The occupational shift indicates a pronounced tendency for post-marriage women to experience career-disconnection or pursue flexible work arrangements.

Figure 1: Life Cycle of Work and Marriage Status

(a) Married (%) (b) Labor Force Participation (%)



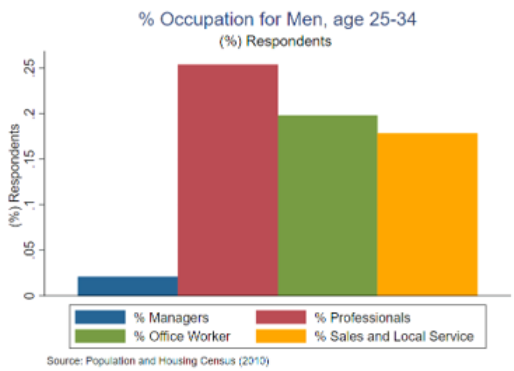
Notes: The graph plots the life cycle of men (blue dots) and women (red dots) with respect to marriage (Panel (a)) and labor force participation decisions (Panel (b)). Each dot represents the percentage of the birth cohort married and working, based on the Population and Housing Census (2010). "Married" in Panel (a) includes those ever married (including divorced or widowed), and labor force participation includes all types of work, including full-time and part-time jobs.

Gender-segmented Industry Structure Using Census on Establishment, Table A3 shows the gender-segmented labor market across sectors. There are significant differences in industry gender share (Column (3)) across sectors. The most male-dominant sector is transportation, with a 0.131 female-to-male employment ratio (i.e., there are 0.131 women per man), followed by construction, government, other services, real estate and rental services, manufacturing, and wholesale. On the other hand, more women are hired in local service sectors such as retail, banking (mostly insurance), education, medical, restaurants and bars, and welfare facilities.

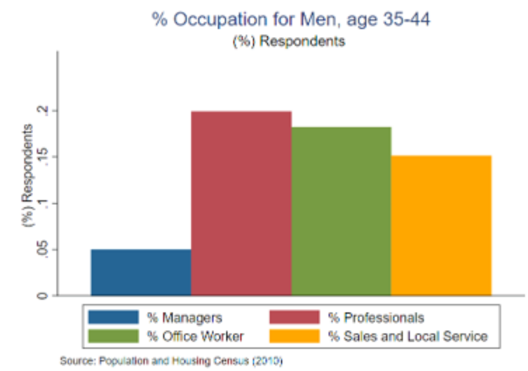
The divergent labor market trajectories of men and women, especially as influenced by marriage and working status, underscore the complex interplay between social norms and economic opportunities.

Figure 2: Occupation Composition for Gender and Age Groups

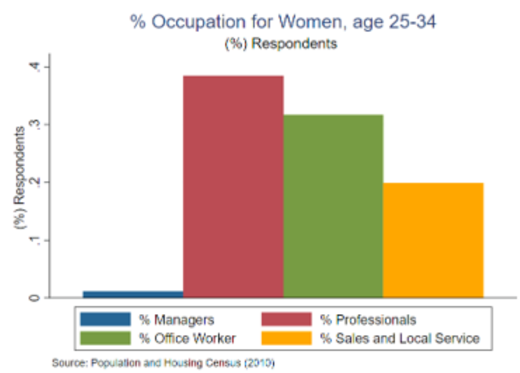
(a) Men (age 25-34)



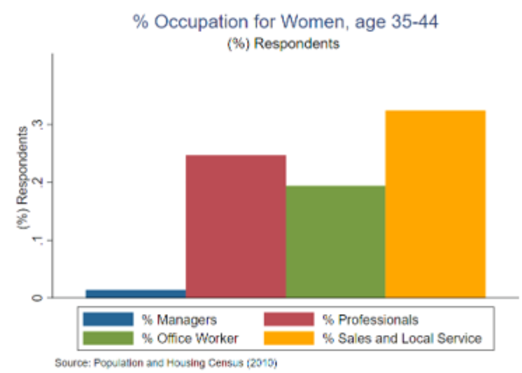
(b) Men (age 35-44)



(c) Women (age 25-34)



(d) Women (age 35-44)



Notes: The graph plots each gender and birth cohort's occupational compositions, based on the Population and Housing Census (2010). Each bar represents the share of gender-birth cohorts working in managerial positions (blue bar), white-collar professions (red bar), office workers (green bar), and the local service sector (orange bar). The sample includes workers in all kinds of jobs, including full-time and part-time positions.

2.2 Economic Concentration and the Expansion of KTX

Home to 51 million people, South Korea is renowned not only for its rapid economic growth but also for its high urbanization rate and significant geographical concentration of economic activities. The urbanization rate was 81.7% in 2018, similar to that of developed countries such as the United States, Canada, and France. Half of the South Korean population resides in the Seoul metropolitan areas (in Korean, 'Sudo-kwon')—the capital of the country (Seoul) and the surrounding areas that encompass the Gyeonggi and Incheon provinces (red-bordered area in Figure 3). The population of Seoul city was estimated at 10.29 million in 2016; however, the sprawling Seoul metropolitan area is significantly larger, encompassing 25.6 million people, which is slightly more than half of the total population. Hereafter, we define 'Seoul' areas to be the Seoul metropolitan area.

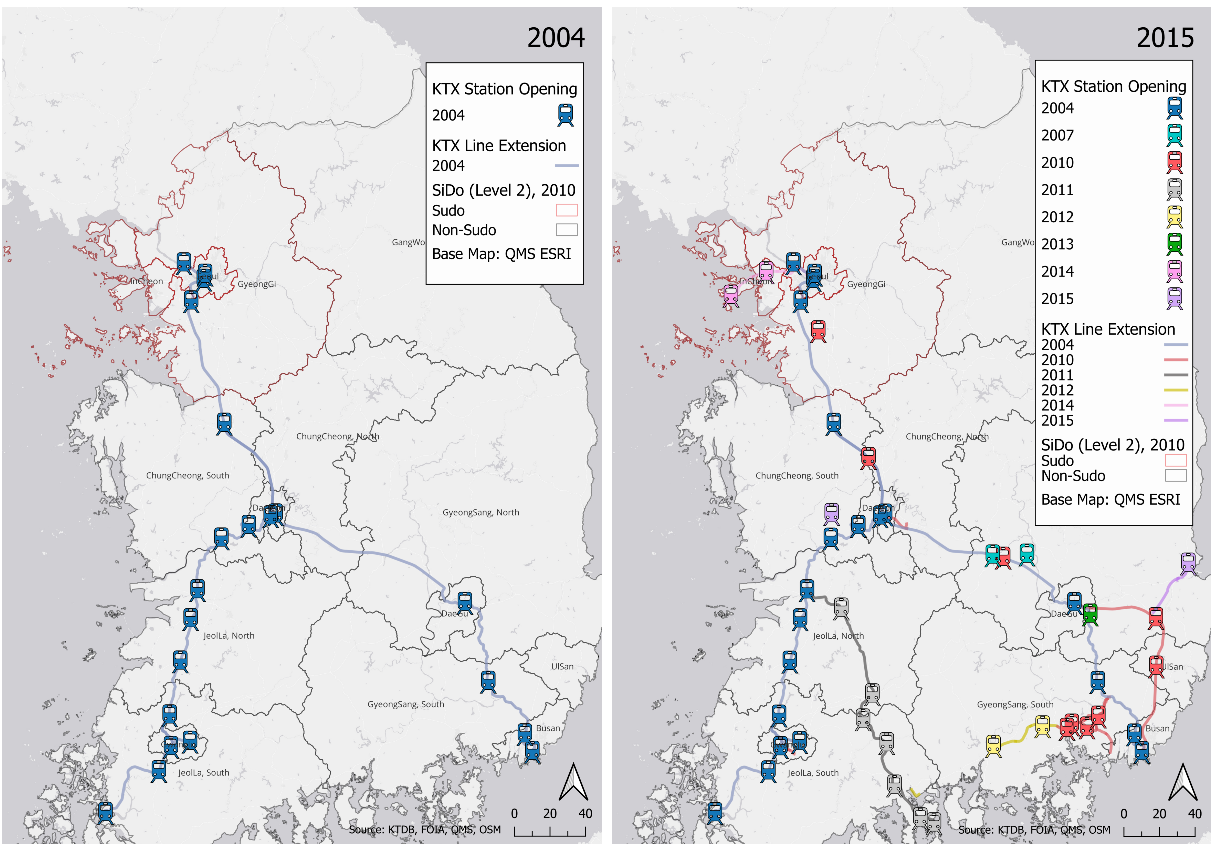
High-Speed Rail (HSR) Expansion The Korea Train eXpress (KTX), South Korea's HSR system, has constantly been evolving since its first introduction in 2004. The initial KTX planning was made in the 1980s to reduce regional economic disparities and alleviate acute urbanization problems near Seoul, including unaffordable housing and traffic congestion. The bullet train not only connected major cities such as Seoul and Busan, but also brought many smaller towns back then into the network improving their connectivity.

Figures 3 and Table A1 depict the main expansion phases of the KTX system. The system continuously evolved with major segment expansion in 2004, 2010, and 2012.[4] In April 2004, KTX began its first operation in twenty-one stations along the Gyeongbu line, connecting Seoul to the southeastern region, and the Honam line, connecting Seoul to the southwestern part of South Korea. Between November and December 2010, nine more stations in the Gyeongbu and Gyeongui lines opened extending the service from Daegu to Busan. In October 2011, seven more stations in the Jeonla line began operation. By 2012, KTX covered 22% of the total territory and served 56% of the total South Korean population (KOSIS).

[4]There were other major expansions in later years, but our data points ended in 2015. The present information is from the Korea Railroad Corporation (KORAIL), upon the authors' information release request through the Korea Ministry of Interior and Safety (MOIS).

Figure 3: KTX Network Expansion

(a) 2004 (b) 2015



Notes: The map depicts the Korea Train eXpress (KTX) network in 2004 (Panel (a)) and in 2015 (Panel (b)). The red-bordered area in the northwest corner of the South Korean peninsula represents the Seoul metropolitan areas, which encompass Seoul city, Gyeonggi province, and Incheon province. The 2015 map shows the staggered expansion of the network system. Source: Korea Transportation Database (KTDB).

KTX has distinctive features as an inter-city, non-commuter train. First, it has significantly shortened inter-city travel times. For example, a one-way trip from Seoul to Busan takes six hours by car or traditional rail, but only two to two and a half hours via KTX. Secondly, it is a non-commuter passenger train, with more than 80% of KTX passengers using the service to visit friends or relatives or for business trips. According to transportation statistics, fewer than 1% of passengers use KTX for daily commutes due to the high ticket price.[5]

As illustrated in Figure A3 Panel (a), KTX passenger numbers have increased significantly since its inception. KTX accounted for 3.7% of interregional ridership (i.e., buses, cars, KTX, or non-KTX trains) in 2004 and reached 10% by 2015,[6] while ridership on other modes of transportation (except for domestic flights[7]) has decreased over time. Considering the total passenger-kilometers in Figure A3 Panel (b), the increase for KTX is even more significant. Total passenger-kilometers for KTX surpassed those of non-KTX trains by 2006, contributing to the overall increase in total passenger-kilometers from 2011. KTX is used for longer distances than the average commuting distance.[8].

The shift in transportation infrastructure preferences underscores the transformative role of KTX in South Korea's economic geography and sets the stage for examining its broader socio-economic impacts. Among these impacts, the potential influence of HSR expansion on labor market dynamics, particularly in terms of gender inequality, warrants closer examination.

[5] To be specific, a one-way train ticket from Seoul to Busan costs around $40. With the average household net-adjusted disposable income per capita at USD 19,372 per year in 2015 (USD 1,614.33 per month), train tickets were prohibitively expensive for daily commutes.

[6] Korea Transport DataBase (KTDB), Korea Transport Institute (KOTI).

[7] KTX competes primarily with intercity buses and existing train networks, but not with domestic flights. Domestic flights are not significant interregional transit options within the Korean peninsula due to airports being distant from city centers, making them less efficient. Indeed, ridership on domestic flights (excluding flights to Jeju Island) remained stagnant from 2000 to 2015, while ridership for non-KTX trains and intercity buses significantly decreased over the same period.

[8] Combining Figures A3 (a) and (b), in 2015, the average travel distance per ticket was 15,000,000,000 passenger-kilometers / 60,000,000 tickets = 250 km, which far exceeds the average commute distance in South Korea.

3 Data

This section explains how we collated multiple data sources into a geo-coded balanced panel dataset, covering the period from 2000 to 2016 annually (t=16) for 228 districts (x=228)(For a detailed description of the data processing steps, see Appendix A.). To evaluate the socio-economic impacts of the KTX, we harmonized the Census on Establishments, the Population and Housing Census, and the Korea Labor and Income Panel Study (KLIPS). The two censuses are collected by Statistics Korea (KOSTAT) and managed in the Micro Data Integrated System (MDIS), which can be accessed with subscription. The Korea Labor Institute (KLI) publishes and manages KLIPS. The two censuses and KLIPS are released on an annual basis.

To ensure consistency between districts that have changed over time, we standardized yearly district variations to align with the 2010 administrative guidelines. Since our analysis focuses on the railway system, we excluded island districts from our sample. For example, Jeju Gun and Ulleung Gun were excluded from the analysis.

For the treatment specification, we calculate a travel time analysis from each district centroid to the closest KTX station for the period from 2000 to 2015. Therefore, with the expansion of the KTX rail network the closest station for districts changed over time. If the closest KTX station was within a 30-minute driving distance from a district centroid in a given year, we defined the district as treated. Until a new KTX station opened within this proximity, the district maintained its control group status.

4 Empirical Strategy

Does the construction of HSR influence gender disparities in labor market outcomes, such as wages and employment levels? Additionally, how does HSR shift the population and economic activity across spaces in Seoul and non-Seoul? This section connects the questions into testable empirical designs.

4.1 Staggered Difference-in-Differences (DID) Design

As our main empirical strategy, we use the difference-in-differences (DID) research design to study the impact of the HSR expansion on outcome variables. The HSR station installments have a 'staggered' nature as they roll out over the years. The DID literature discusses potential identification problems when treatment timing varies between units and periods and suggests a number of solutions (Borusyak, Jaravel, and Spiess, 2022; Sun and Abraham, 2021). To address the potential identification issues, we adopt the estimator proposed by Callaway and Sant'Anna (2021), estimating dynamic effects of HSR across different durations after interventions.

For a district i at time t, our main specification is as follows.

$$Y_{it} = \alpha_i + \gamma_t + \sum_l \beta_l \cdot \mathbb{1}\{t - E_i = l\} + \epsilon_{it}, \tag{1}$$

where Y_{it} represents the outcome of interest in district i in year t, $\mathbb{1}$ is an indicator function for a district i defined as 'treated'. As explained before, the treated districts are the ones with centroids within 30 minutes of driving time from the closest HSR station, and E_i is the year of the first HSR station opening in district i during the sample period, thus l constitutes the duration after the initial intervention. District fixed effects (α_i) and year fixed effects (γ_t) are included in the regression, with standard errors clustered at the district level. β coefficients can be interpreted as the change in the outcome of interest following the construction of a KTX station in a district, relative to the control group. The map of the districts treated is shown in Figure A5, treated regions are colored with Pinks.

To identify the heterogeneous impacts between core and non-core areas, in our main regression, we separate the districts located in core versus non-core areas. Core areas are defined as districts located in the Seoul metropolitan areas (Seoul, Gyeonggi, and Incheon provinces), and non-core areas are districts in the rest of the provinces.

4.2 Selection of Control Groups

To establish suitable control groups for the treated districts, we select districts located within 100 km of old railroad stations constructed during the Japanese Colonial Era. Beginning in 1894, Japan initiated railroad construction in Korea as part of its modernization efforts. The Gyeongin line, the first railroad network in Korea, opened in 1899. Subsequently, the Gyeongbu and Gyeongui lines were added between 1904 and 1906 to facilitate military logistics to China and Russia. By 1910, the Honam lines were constructed to transport natural resources and agricultural products from South Korea to Japan.

We argue that using historical railroads to select control groups ensures that both the treated and control districts share similar characteristics in terms of population dynamics, employment patterns, and more importantly, gender distribution, providing a robust comparison for our analysis. Specifically, first, the old railroad stations, constructed during the Japanese Colonial Era, have influenced the development patterns of their surrounding districts for more than a century, leading to development trajectories similar to those near the new HSR stations. Secondly, these historical stations were strategically placed to optimize connectivity across Korea's mountainous terrain, ensuring that districts near them have comparable geographic and topological characteristics to those chosen for new HSR stations. This similarity ensures that both treated and control groups have faced analogous geographic constraints and opportunities. Lastly, given that the mountainous terrain of the Korean Peninsula, which covers approximately 70% of the area, using existing rail corridors for the construction of the new HSR reflects practical considerations such as cost and feasibility, ensuring that both treated and control groups are selected based on similar strategic considerations. Integrating these historical sites as control groups aligns our treated and control districts on key dimensions, enhancing the credibility of our difference-in-differences estimates and allowing for more accurate conclusions about the impact of HSR expansion[9].

[9] The existing literature addresses endogeneity issues in transportation infrastructure using railroad planning maps (Baum-Snow, 2007; Duranton, Morrow, and Turner, 2014), historical railroads (Duranton, Morrow, and Turner, 2014; Morten and Oliveira, 2016; Garcia-López, Holl, and Viladecans-Marsal, 2015; Tsivanidis, 2018), or topological traits (Faber, 2014) as instrumental variables (IV). By adopting a similar approach, we ensure that our control groups offer a valid counterfactual to the treated districts, thereby strengthening the validity of our DID estimates.

5 Results

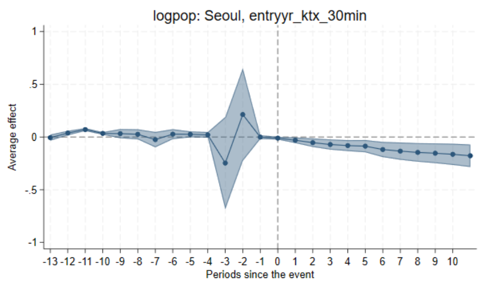
5.1 KTX and the Changes in Local Employment and Population

Figures 4 and Table 1 illustrate the impact of HSR on the relative population, employment, and employment-to-population ratio changes for the treated districts. In Figure 4, we do not observe clear pre-trend patterns between treated and control groups before the HSR station installment for all variables, which reinforces our identification strategies. Overall, we see that the dynamics of population and employment in treated districts in Seoul decreases relative to control groups in Seoul (Panels (a) and (b)). We do not observe employment-to-population ratio changes in Seoul areas relative to control districts in Seoul areas. On the other hand, the non-Seoul population marginally increases, and the employment level does not seem to be affected by the entry of HSR stations (Panels (d) and (e)). The employment-to-population ratio is also not statistically significantly affected by HSR (Panel (f)). Table 1 synthesizes the event study estimates into post-Average Treatment Effect for Treated (ATT) results, which reinforces the results in Figures 4.

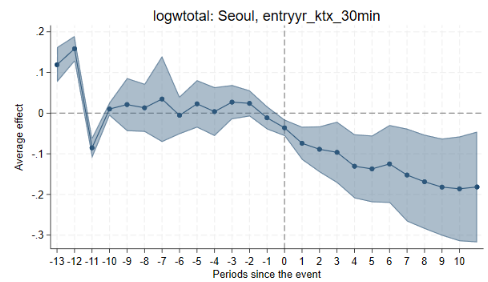
It is important to note that the results do not imply absolute population and employment level changes in the treated regions. For example, in Seoul, the coefficients are relative to the control group in Seoul (i.e., those who are further away from train stations), which experiences rapid growth during this period (2000-2015). The same caution applies to non-Seoul areas where the population and employment typically decline. Furthermore, with the publicly available Census on Establishment data or population data, we cannot observe the dynamics of entry and exit or population inflow and outflow, which could be an important margin to investigate.

Figure 4: Impact of HSR on Local Population and Employment, DID Estimates

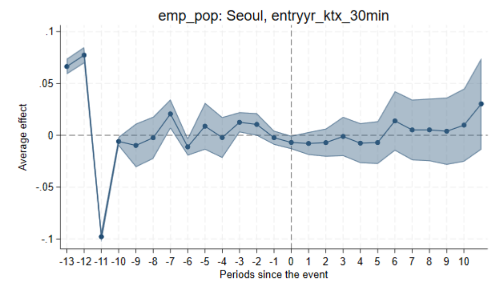
(a) Seoul (log Population)



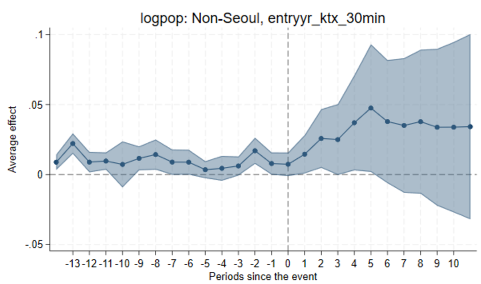
(b) Seoul (log Employment)



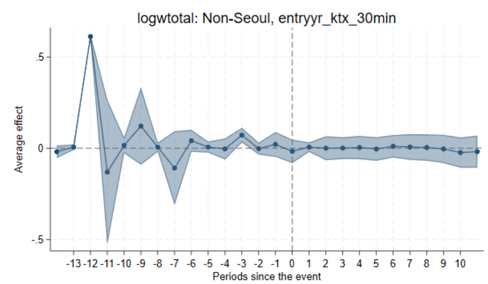
(c) Seoul (Emp-to-pop)



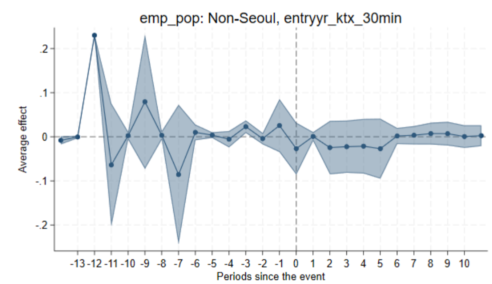
(d) Non-Seoul (log Population)



(e) Non-Seoul (log Employment)



(f) Non-Seoul (Emp-to-pop)



Notes: The graphs are event study plots for Average Treatment effect for Treated (ATT), using (Callaway and Sant'Anna, 2021) estimator with 'never-treated' and 'not-yet-treated' districts as control groups. The dependent variable for Panel (a) and (d) is log population and for Panel (b) and (e) is log employment, and (c) and (f) is employment-to-population ratio. Panel (a), (b), and (c) depict the estimates from the Seoul metropolitan area sample, which includes Incheon, Seoul, and Gyeonggi provinces. Panel (c), (d), and (e) illustrate the estimates from the non-Seoul area, including all districts other than the Seoul metropolitan and a few districts in island. Standard errors are clustered at the district level.

Table 1: **Impact of HSR on Local Population and Employment**

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Population)		Log(Employment)		Emp-to-Pop	
	Seoul	Non-Seoul	Seoul	Non-Seoul	Seoul	Non-Seoul
Treat (ktx< 30min)	-0.100***	0.030*	-0.128***	-0.003	0.002	-0.01
	(0.032)	(0.018)	(0.042)	(0.028)	(0.011)	(0.017)
District FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Observations	1,168	2,144	1,160	2,141	1,158	2,141

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The coefficients are Average Treatment effect for Treated (ATT) using (Callaway and Sant'Anna, 2021) estimator, with 'never-treated' and 'not-yet-treated' districts as control groups. The dependent variable for columns (1)-(2) is log population, for columns (3)-(4) is log employment, and for columns (5)-(6) is employment-to-population ratio. Seoul is defined as districts within the Seoul metropolitan area of Incheon, Seoul, or Gyeonggi provinces. Non-Seoul encompasses all other districts except a few islands. Standard errors are in parentheses, clustered at the district level.

5.2 KTX and the Gender Gap in Employment and Population

Figure 5 and Table 2 reveal the impact of HSR on the female-to-male population and employment ratios in Seoul and non-Seoul areas. Again, the staggered DID results in Figure 5 show the absence of a pre-trend, suggesting that the observed changes are attributed to the HSR station installments.

Overall, relative to the control districts, women tend to reside more in districts closer to HSRs in both Seoul (Panel (a)) and non-Seoul (Panel (c)). In the Seoul area, the female-to-male population ratio increases immediately with the HSR station's entry, whereas in the non-Seoul areas, the effects come with about a five-year lag after the installment of HSR. Focusing on the female-to-male employment ratio, in Seoul, there is no significant post-trend observed (Panel (b)). However, as shown in Panel (d), districts in non-Seoul areas show an immediate increase in the female-to-male employment ratio once an HSR station is installed.

Table 2 presents the one-shot ATT estimates on the gender gap. Columns (1) and (2) show that proximity to an HSR station is associated with a 1.2% increase in the female-to-male population ratio in the Seoul areas and a 0.5% increase in non-Seoul areas. The impact on the female-to-male employment ratio is more pronounced in non-Seoul areas, with a 4.77% (3.8 percentage points) increase (Column (4)). The impact on the female-to-male

Table 2: Impact of HSR on Population and Employment Gender Ratio, DID Estimates

	(1)	(2)	(3)	(4)
	Female-to-Male Pop Ratio		Female-to-Male Emp Ratio	
	Seoul	Non-Seoul	Seoul	Non-Seoul
Treat (KTX< 30min)	0.012***	0.005**	-0.015	0.038***
	(0.003)	(0.002)	(0.017)	(0.01)
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Pre-treatment (2003) Mean	0.986	1	0.744	0.791
Observations	1,166	2,144	1,160	2,141

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The coefficients are Average Treatment effect for Treated (ATT) using (Callaway and Sant'Anna, 2021) estimator, with 'never-treated' and 'not-yet-treated' districts as control groups. The dependent variable for columns (1)-(3) is the female-to-male population ratio and for columns (4)-(6) is the female-to-male employment ratio. Seoul is defined as districts within the Seoul metropolitan area of Incheon, Seoul, or Gyeonggi provinces. Non-Seoul encompasses all other districts except a few islands. Standard errors are in parentheses, clustered at the district level.

employment ratio in the Seoul area is not statistically or economically significant (Column (3)).

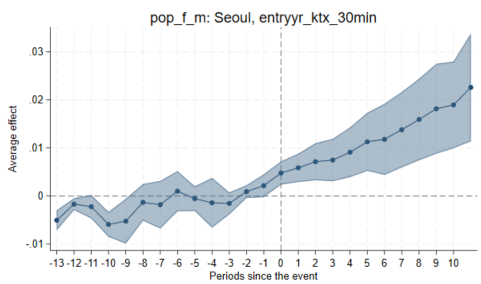
These findings indicate that while HSR stations contribute to an increased female-to-male population ratio in both Seoul and non-Seoul areas, their impact on female employment is significantly more pronounced in non-Seoul areas. This suggests potential differences in labor market dynamics and the role of HSR accessibility in influencing gender-specific employment opportunities in urban versus non-urban regions. Further investigation into these dynamics could provide deeper insights into how infrastructure developments like HSR stations affect gender disparities in the labor market.

6 Mechanism

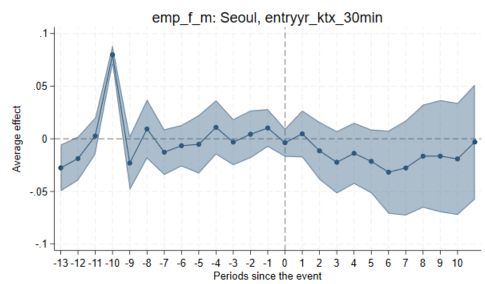
In this section, we further investigate what drives the gender-nuanced effects of HSR on labor market opportunities, which initially appear to be a gender-neutral technology. To understand the mechanisms, we explore the gender-segmented labor market dynamics by examining both the labor demand and labor supply channels.

Figure 5: Impact of HSR on Population and Employment Gender Ratio, DID Estimates

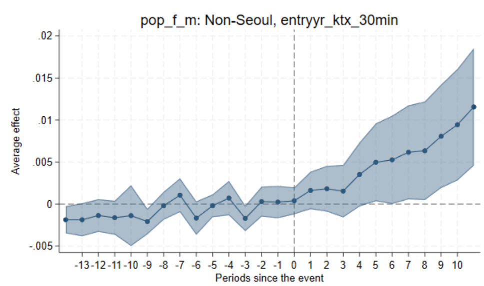
(a) Seoul (Female-to-Male Pop Ratio)



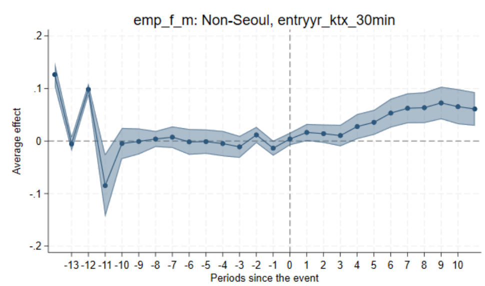
(b) Seoul (Female-to-Male Emp Ratio))



(c) Non-Seoul (Female-to-Male Pop Ratio)



(d) Non-Seoul (Female-to-Male Emp Ratio)



Notes: The graphs are event study plots for Average Treatment effect for Treated (ATT), using (Callaway and Sant'Anna, 2021) estimator with 'never-treated' and 'not-yet-treated' districts as control groups. The dependent variable for Panel (a) and (c) is the female-to-male population ratio, and for Panel (b) and (d) is the female-to-male employment ratio. Panel (a) and (b) depict the estimates from the Seoul metropolitan area sample, which includes Incheon, Seoul, and Gyeonggi provinces. Panel (c) and (d) illustrate The estimates from the non-Seoul area, including all districts other than the Seoul metropolitan and a few island districts. Standard errors are clustered at the district level.

6.1 Labor Demand Channel

As highlighted in previous literature (Lin, 2017; Dong, Zheng, and Kahn, 2020), HSR tends to stimulate industries that rely more on the transportation cost of moving people rather than goods. From a gender-specific labor demand perspective, new train stations could induce a disproportionate increase or decrease in labor demand for certain genders, as South Korea's sectors tend to be gender-segmented, as documented in Section 2. In this section, we empirically examine whether HSR stations affect employment differently for male and female-intensive sectors, and how the gender shares in each industry are affected.

Effect on Reallocation of Industries Figure 6 shows the difference-in-difference estimation's ATT results for Seoul (upper panel) and non-Seoul (lower panel). Each dot and bar indicate the regression coefficients and standard errors, with sector-specific employment share within a district as the dependent variable. The X-axis indicates the sector-level 2003 female-to-male employment ratio, and the Y-axis shows the regression coefficients. The red dots indicate statistically significant regression coefficients at the 10% level. (The full regression results for each dot are presented in Table A3). We refer to 'male-intensive' sectors as those with a female-to-male employment ratio of less than one.

In Seoul (Panel a, Figure 6), significant negative effects are observed in the employment shares of real estate, manufacturing, and lodging sectors, indicating a decline in these sectors relative to the control group. Conversely, banking and finance, as well as restaurant and bar sectors, show significant positive coefficients, suggesting an increase in employment shares in these service-oriented, and also more female-intensive industries. This aligns with the hypothesis that HSR tends to boost sectors reliant on the transportation cost of moving people.

In non-Seoul regions (Panel b, Figure 6), similar patterns are observed, with significant declines in employment shares in transportation, construction, and manufacturing sectors. These sectors are traditionally reliant on the movement of goods. On the other hand, significant positive effects are noted in the banking, retail, medical services, and private education sectors, indicating an increase in employment shares in service-oriented industries[10].

Overall, Figure 6 underscores the transformative impact of HSR on local economies, highlighting a significant reallocation of employment shares towards service-oriented sectors. This shift is more pronounced in non-Seoul regions, suggesting that HSR plays a critical role in regional economic development, particularly in areas outside major urban centers.

Effect on Gender Shares Next, we investigate how each industry's gender share has been affected by HSR. Figure 7 illustrates the gender share effects of HSR by sector, presenting distinct outcomes for Seoul (Panel a) and non-Seoul (Panel b) regions. The graphs depict the coefficients for the female-to-male employment ratio within districts as the dependent variable, with significant coefficients at the 10% level marked in red.

In Seoul (Panel a), several sectors show statistically significant changes in gender share due to HSR. Notably, the transportation and maintenance sectors exhibit positive and significant coefficients, indicating an increase in the female-to-male employment ratio. This suggests that HSR has contributed to a higher relative demand for female labor in these sectors.

In non-Seoul regions (Panel b), significant positive effects are observed in primary education, general retail, and restaurant and bar sectors. These increases suggest that HSR has similarly boosted female employment shares outside Seoul, particularly in service-oriented sectors.

A key observation is the more pronounced gender share effects in non-Seoul regions compared to Seoul. This discrepancy might be due to differences in economic structures and the varying impact of HSR on local economies. In non-Seoul areas, sectors like primary education and retail, which traditionally have higher female employment, show stronger positive responses, potentially reflecting a shift towards service-based local economies.

Overall, the sectors showing significant employment share increases (i.e., banking, finance, and service industries) are also the ones which show higher female-to-male employment share in general as well as where female labor demand has notably increased with HSR. This reinforces the argument that HSR contributes to the increased female employment

[10] Figure A7 presents similar graphs with the Y variable representing the logarithm of sectoral employment.

opportunities, particularly in female-intensive sectors.

6.2 Labor Supply Channel

From a labor supply perspective, new amenities around train stations could have eased the burden of childcare and domestic labor for females, allowing them to participate in labor activities outside of their homes. To empirically test this argument, we estimate the impacts of HSR on the local amenity level. Table 3 shows the impact of HSR on the level of local amenity, quantified as the number of establishments per 10,000 residents in certain service industries following (Diamond, 2016)'s approach. The results show that after the installation of HSR stations, the number of establishments in major service sectors that are essential for daily activities, such as banking and finance (Column (3)), and retail (Column (7)), increased in both Seoul and non-Seoul areas.

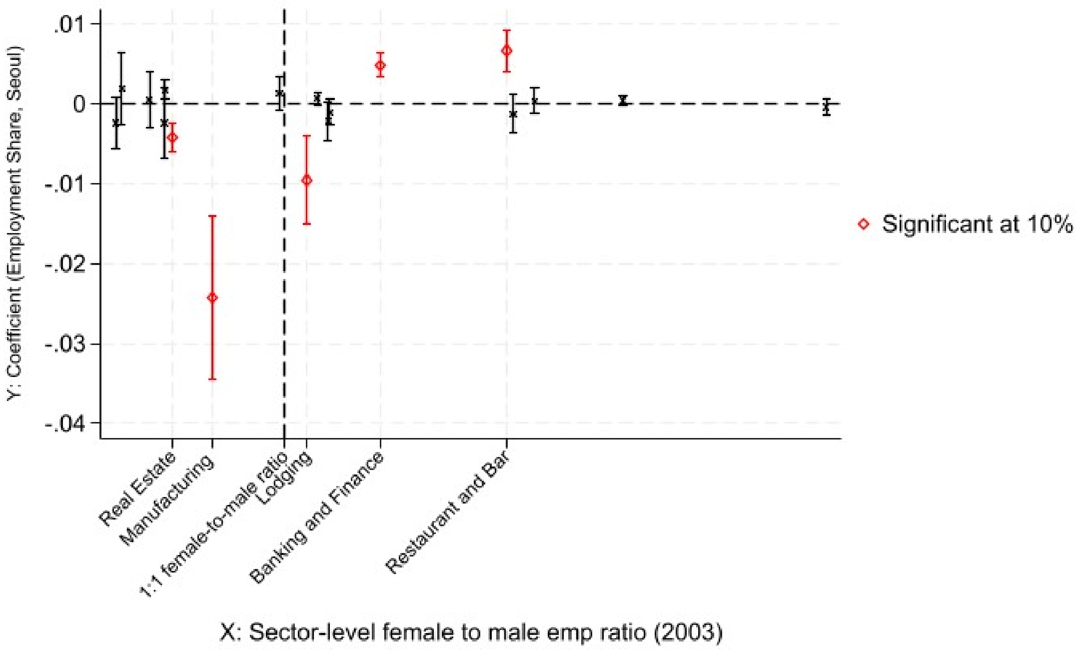
While both areas experienced growth in amenities, there are differences. Seoul's increase was more pronounced in sectors such as movie theaters, banking and finance, R&D, and medical services, whereas non-Seoul areas saw larger increases in private education, general retail, clothing, and hair salons. These differences reflect the varied needs and initial service availability between Seoul and non-Seoul areas, suggesting that HSR installations have a differential impact based on the demands of residents and non-residents brought by HSR.

Table 4 delves deeper into educational amenities by investigating childcare employment and private education employment per woman[11], which is more directly related to the residents, particularly women with children. Interestingly, preschool and elementary school employment per woman decreases in Seoul areas by 12.867 per 10,000 women (Column (1)), whereas the number increases by 7.204 per 10,000 in non-Seoul areas (Column (2)). Although insignificant, the private education employment per woman shows qualitatively similar results in Columns (3) and (4).

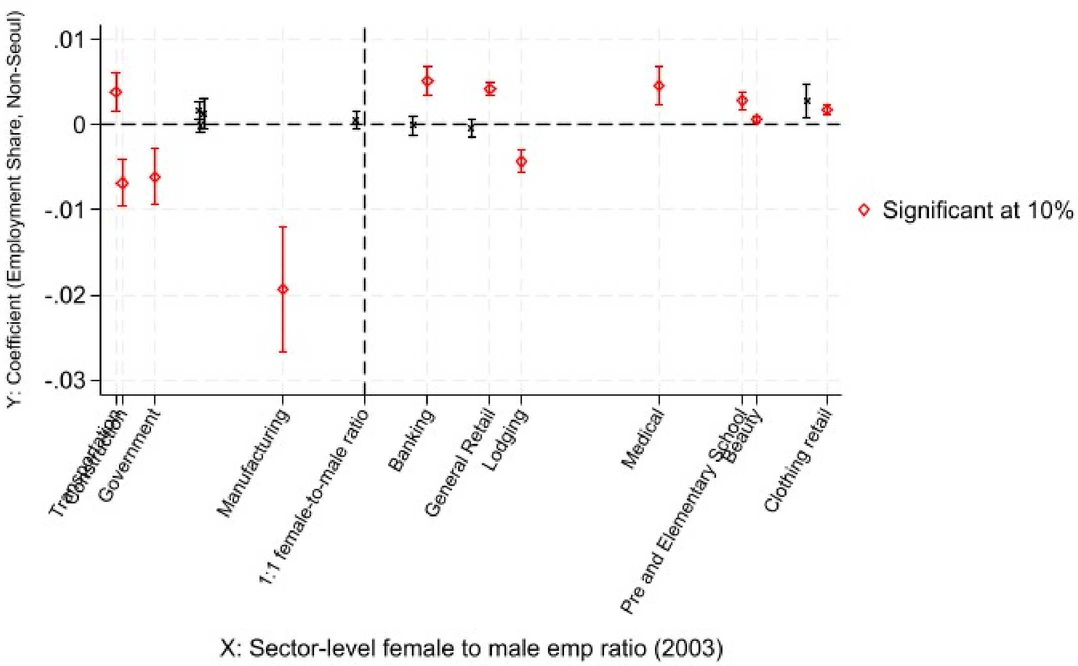
[11] Note that due to the harmonization process of industry codes (KSIC) 8th and 9th editions, we can only observe up to 3 digits of the industry code from the Census on Establishment, and some industries can only be observed when we merge them together. For example, childcare and elementary school employment are combined as P851 (KSIC 9th edition), and private education employment includes specialized schools (P854), private academic institutes (P855), other educational institutions (P856), libraries (R902), sports institutes (R911), and other entertainment venues such as amusement parks (R912).

Figure 6: The Employment Share Effects of HSR by Sector

(a) Seoul



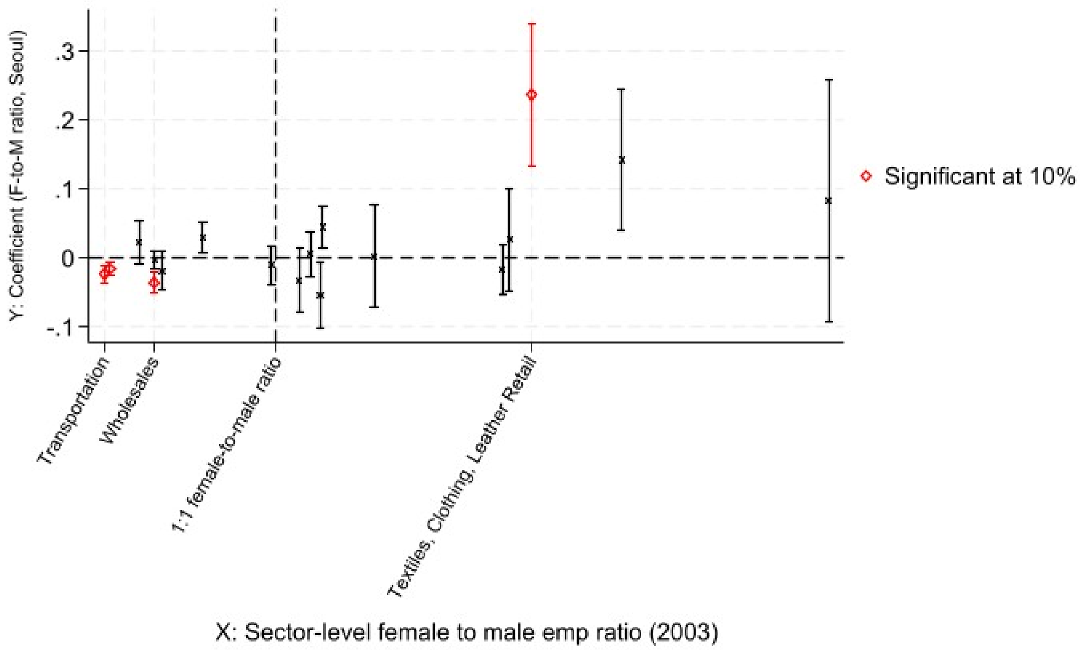
(b) Non-Seoul



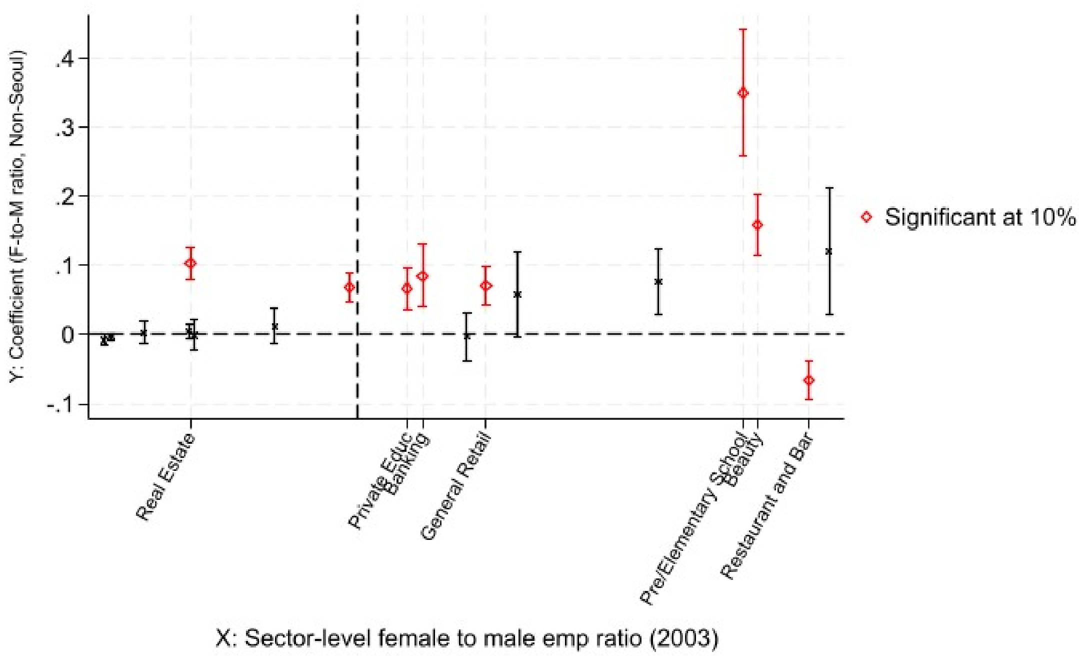
Notes: The graphs plot the Average Treatment Effect for the Treated (ATT) using the method described by (Callaway and Sant'Anna, 2021) for each sector. The Y-axis represents the coefficient results from the subsample analysis of each industry, where the Y variable is the employment share of each industry. The X-axis shows the sector-level female-to-male employment ratio, calculated from the Census of Establishments in 2003. Red bars and diamond dots indicate regression results that are statistically significant at the 10% level. The coefficients and the statistics used in the graphs are also presented in Appendix A3.

Figure 7: The Gender Share Effects of HSR by Sector

(a) Seoul



(b) Non-Seoul



Notes: The graphs plot the Average Treatment Effect for the Treated (ATT) using the method described by (Callaway and Sant'Anna, 2021) for each sector. The Y-axis represents the coefficient results from the subsample analysis of each industry, where the Y variable is the female-to-male share. The X-axis shows the sector-level female-to-male employment ratio, calculated from the Census of Establishments in 2003. Red bars and diamond dots indicate regression results that are statistically significant at the 10% level. The coefficients and the statistics used in the graphs are also presented in Appendix A3.

Table 3: The Impact of HSR on Endogeneous Amenities

	(1) Movie	(2) F&B	(3) FIN	(4) R&D	(5) School	(6) Private Educ	(7) Gen Retail	(8) Food, Bev Retail	(9) Clothing	(10) Med	(11) Beauty
Panel A. Seoul											
ATT	0.337**	-1.406	1.337***	0.585***	0.043	-0.004	2.559***	0.488	0.257	0.999**	0.486
	(0.145)	(3.622)	(0.366)	(0.154)	(0.266)	(1.102)	(0.676)	(0.669)	(2.466)	(0.289)	(0.550)
Mean(2003)											
Panel B. Non-Seoul											
ATT	-0.098	-5.369***	0.708**	0.142*	0.192	1.687**	2.984***	-0.763	2.896***	0.101	0.965**
	(0.073)	(1.739)	(0.325)	(0.076)	(0.156)	(0.718)	(0.543)	(1.125)	(1.042)	(0.214)	(0.388)
Mean(2003)											

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The coefficients are the Average Treatment Effect for the Treated (ATT) using the (Callaway and Sant'Anna, 2021) estimator, with 'never-treated' and 'not-yet-treated' districts as control groups. Each column represents the endogenous amenity of different sectors, defined as the number of establishments divided by 10,000 population for each industry sub-sample, following (Diamond, 2016). Panel A is for Seoul, defined as districts within the Seoul metropolitan area of Incheon, Seoul, or Gyeonggi provinces, and Panel B is the sub-sample for Non-Seoul areas, encompassing all other districts except a few islands. Column (1) is for the movie industry, (2) is for restaurants and bars, (3) is for banking, finance, insurance, and pension, (4) is for research and development, (5) is for elementary, secondary, and high school, (6) is for private education, (7) is for general retail, (8) is for food and beverage retail, (9) is for medical, (10) is for clothing, textile, and leather stores, and (11) is for hair salons and spas. (For detailed industry characteristics and definitions, refer to Appendix A2.) Standard errors are in parentheses, clustered at the district level.

Table 4: **The Impact of HSR on Education and Childcare Amenities**

	(1)	(2)	(3)	(4)
	Pre,Elementary School Emp per Women		Private Educ Emp per Women	
	Seoul	Non-Seoul	Seoul	Non-Seoul
Treat	-12.868***	7.204**	-10.488	7.726
	(3.866)	(3.768)	(20.516)	(5.939)
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Mean (2003)	98.06	124.02	288.58	211.92
Observations	1,166	2,144	1,166	2,144

Notes: *** p<0.01, ** p<0.05, * p<0.1. The coefficients are the Average Treatment Effect for the Treated (ATT) using the (Callaway and Sant'Anna, 2021) estimator, with 'never-treated' and 'not-yet-treated' districts as control groups. The dependent variables in Columns (1) and (2) are the number of workers in pre-school and elementary school (KSIC 9th code P851) per 10,000 female population, while Columns (3) and (4) represent the number of workers in private education institutions (KSIC 9th codes P854, P855, P856, R902, R912, R911) per 10,000 female population. (For detailed industry characteristics and definitions, refer to Appendix A2.) Standard errors are in parentheses, clustered at the district level.

Combining the results in Table 3 and Table 4, the findings imply that HSR increases the levels of amenities relative to residents in Seoul and the non-Seoul areas. However, it appears that Seoul's amenities are geared toward non-residents' foot traffic, whereas non-Seoul's amenities are geared towards residents, particularly women with children. This could potentially contribute to the reduction in women's labor force participation costs, particularly non-Seoul areas. More research is needed to fully understand these dynamics.

7 Conclusion

This study investigates the impact of High-Speed Rail (HSR) on the redistribution of economic activities and gender disparities in population and labor market outcomes. Employing a staggered Difference-in-Differences (DID) approach as the primary analytical method, we demonstrate that HSR construction in South Korea prompted shifts in population and employment from Seoul metropolitan areas (core) to the rest of the country (non-core) areas. Women, in particular, benefit disproportionately from HSR expansion with the increasing labor demand from the local service sector, a field where women are predominantly employed. Furthermore, the increase in endogenous local amenities due to HSR potentially eased female childcare and domestic burdens, promoting labor supply.

Through our research, we contribute to the existing literature on transportation infrastructure and gender disparities in labor market outcomes, an area that has been relatively underexplored. Our findings advance the discussion by examining how advanced transportation technologies affect both the demand and supply sides of the gender-segmented labor market. This study underscores the importance of understanding the varied impacts of infrastructure investments in different demographic groups, especially in the context of significant spending on transportation infrastructure.

Despite contributions to the literature, this study has several limitations and areas for future work. An immediate and most important task is to explore the choices of residential and workplace locations of men and women before and after HSR using Census data to understand how HSR changes these choices (Venator, 2022; Fan and Zou, 2021; Costa and Kahn, 2000). Additionally, there are many other channels through which HSR might affect women's labor supply decisions, such as accessibility to grandparents (Anstreicher and Venator, 2024), which this study does not fully explore. Future research could delve deeper into these issues, providing more definitive evidence and broader implications from a comprehensive general equilibrium perspective.

References

Ahlfeldt, Gabriel M and Arne Feddersen (2018). “From periphery to core: measuring agglomeration effects using high-speed rail”. In: *Journal of Economic Geography* 18.2, pp. 355–390.

Aksoy, Cevat Giray, Ozkan Berkay, and Julia Philip (2019). “Robots and the Gender Pay Gap: Evidence from Europe”. In: *2nd IZA/CREA Workshop: Exploring the Future of Work*. Vol. 5. 6.

Anstreicher, Garrett and Joanna Venator (2024). “To Grandmother’s House We Go: Informal Childcare and Female Labor Mobility”. In.

Bartik, Timothy J (1991). “Who benefits from state and local economic development policies?” In.

Baum-Snow, Nathaniel (2007). “Did highways cause suburbanization?” In: *The Quarterly Journal of Economics*.

Baum-Snow, Nathaniel et al. (2017). “Roads, railroads, and decentralization of Chinese cities”. In: *Review of Economics and Statistics*.

Black, Dan A, Natalia Kolesnikova, and Lowell J Taylor (2014). “Why do so few women work in New York (and so many in Minneapolis)? Labor supply of married women across US cities”. In: *Journal of Urban Economics* 79, pp. 59–71.

Blau, Francine D and Lawrence M Kahn (2017). “The gender wage gap: Extent, trends, and explanations”. In: *Journal of Economic Literature* 55.3, pp. 789–865.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2022). “Revisiting Event Study Designs: Robust and Efficient Estimation”. In: *Available at SSRN 2826228*.

Callaway, Brantly and Pedro HC Sant’Anna (2021). “Difference-in-differences with multiple time periods”. In: *Journal of econometrics* 225.2, pp. 200–230.

Charnoz, Pauline, Claire Lelarge, and Corentin Trevien (2018). “Communication costs and the internal organisation of multi-plant businesses: evidence from the impact of the French high-speed rail”. In: *The Economic Journal* 128.610, pp. 949–994.

Chauvin, Juan Pablo (2017). “Gender-Segmented Labor Markets and the Effects of Local Demand Shocks”. In: *Job Market Paper, Harvard University*.

Costa, Dora L and Matthew E Kahn (2000). “Power couples: changes in the locational choice of the college educated, 1940–1990”. In: *The Quarterly Journal of Economics* 115.4, pp. 1287–1315.

Diamond, Rebecca (2016). “The determinants and welfare implications of US workers’ diverging location choices by skill: 1980-2000”. In: *American Economic Review* 106.3, pp. 479–524.

Dinkelman, Taryn (2011). “The effects of rural electrification on employment: New evidence from South Africa”. In: *American Economic Review* 101.7, pp. 3078–3108.

Donaldson, Dave (2018). “Railroads of the Raj: Estimating the impact of transportation infrastructure”. In: *American Economic Review* 108.4-5, pp. 899–934.

Dong, Xiaofang, Siqi Zheng, and Matthew E Kahn (2020). “The role of transportation speed in facilitating high skilled teamwork across cities”. In: *Journal of Urban Economics* 115, p. 103212.

Duranton, Gilles, Peter M Morrow, and Matthew A Turner (2014). “Roads and Trade: Evidence from the US”. In: *Review of Economic Studies* 81.2, pp. 681–724.

Faber, Benjamin (2014). “Trade integration, market size, and industrialization: evidence from China’s National Trunk Highway System”. In: *Review of Economic Studies* 81.3, pp. 1046–1070.

Fan, Jingting and Ben Zou (2021). “The dual local markets: Family, jobs, and the spatial distribution of skills”. In: *SSRN Electronic Journal,(November)*.

Farré, Lıdia, Jordi Jofre-Monseny, and Juan Torrecillas (2020). “Commuting Time and the Gender Gap in Labor Market Participation”. In.

Garcia-López, Miquel-Àngel, Adelheid Holl, and Elisabet Viladecans-Marsal (2015). “Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities”. In: *Journal of Urban Economics* 85, pp. 52–67.

Goldin, Claudia and Lawrence F Katz (2016). "A most egalitarian profession: pharmacy and the evolution of a family-friendly occupation". In: *Journal of Labor Economics* 34.3, pp. 705–746.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu (2005). "Engines of liberation". In: *The Review of Economic Studies* 72.1, pp. 109–133.

Greenwood, Jeremy et al. (2016). "Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation". In: *American Economic Journal: Macroeconomics* 8.1, pp. 1–41.

Heuermann, Daniel F and Johannes F Schmieder (2019). "The effect of infrastructure on worker mobility: evidence from high-speed rail expansion in Germany". In: *Journal of economic geography* 19.2, pp. 335–372.

Kawabata, Mizuki and Yukiko Abe (2018). "Intra-metropolitan spatial patterns of female labor force participation and commute times in Tokyo". In: *Regional Science and Urban Economics* 68, pp. 291–303.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet (2020). "Gender differences in job search: Trading off commute against wage". In.

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

Liu, Sitian and Yichen Su (2020). "The Geography of Jobs and the Gender Wage Gap". In: *Available at SSRN*.

Maurer-Fazio, Margaret et al. (2011). "Childcare, eldercare, and labor force participation of married women in urban China, 1982–2000". In: *Journal of Human Resources* 46.2, pp. 261–294.

Morten, Melanie and Jaqueline Oliveira (2016). *Paving the way to development: Costly migration and labor market integration*. Tech. rep. National Bureau of Economic Research.

Qin, Yu (2017). "'No county left behind?'The distributional impact of high-speed rail upgrades in China". In: *Journal of Economic Geography* 17.3, pp. 489–520.

Redding, Stephen J and Matthew A Turner (2014). "Transportation Costs and the Spatial Organization of Economic Activity". In: *CEP Discussion Paper* No. 1277.

Rosenthal, Stuart S and William C Strange (2012). "Female entrepreneurship, agglomeration, and a new spatial mismatch". In: *Review of Economics and Statistics* 94.3, pp. 764–788.

Severen, Christopher (2018). "Commuting, labor, and housing market effects of mass transportation: Welfare and identification". In.

Sun, Liyang and Sarah Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". In: *Journal of Econometrics* 225.2. Themed Issue: Treatment Effect 1, pp. 175–199.

The World Bank (2024). *World Development Indicators: Consumer Price Index (2010 = 100)*. Retrieved from https://databank.worldbank.org/source/world-development-indicators.

Tsivanidis, Nick (2018). "The aggregate and distributional effects of urban transit infrastructure: Evidence from bogotá's transmilenio". In: *Job Market Paper*.

Venator, Joanna (2022). *Dual-earner migration decisions, earnings, and unemployment insurance*. Boston College.

Vidart, Daniela (2020). "Human Capital, Female Employment, and Electricity: Evidence from the Early 20th Century United States". In: *Unpublished Working Paper, UCSD*.

Zheng, Siqi and Matthew E Kahn (2013). "China's bullet trains facilitate market integration and mitigate the cost of megacity growth". In: *Proceedings of the National Academy of Sciences* 110.14, E1248–E1253.

A Appendix: Data Appendix

A.1 The Driving Time from District to KTX Station

Our primary geospatial data source, the Korea Transport DataBase (KTDB), is a comprehensive transport database created and operated by the Korea Transport Institute (KOTI). In addition to geospatial information for transport analyses, the database offers a broad spectrum of transport modal, fiscal, and passenger data and statistics. Among its available information, we primarily employed the Geographic Information System (GIS) database to construct travel time variables and geographic distance buffers. The GIS database contains geographic and attribute information of transport objects, including roads, public transits, railways, stations, and other transport facilities. Using the geo-coordinates and attributes of KTX stations, along with an open-source map providing real-time road and speed information, we calculated the measures for identifying treatments.

To determine treatment districts, we conducted a network analysis of Origin-Destination (OD) travel time from district centroids to the nearest KTX station for each year. In defining the district centroids as origin points, the current version of the paper uses the midpoint of each district polygon as a straightforward centroid measure. Many districts, especially in Seoul and major provincial cities, have multiple population centers that are distant and scattered. Therefore, using population-weighted centroids might not be a credible method in the South Korean context. Another alternative is to select a major point of interest, such as a department store or town hall, as the city center indicator. We will conduct robustness checks using several major objects to alter the centroids.

To mark the KTX stations as destination points, we combined the KTDB station layers with the station opening dates obtained from KORAIL. The earliest available year for KTDB GIS data without requiring a geocoordinate system conversion was 2009, which includes transport information from 2008. We chose this option to reduce analysis errors arising from geocoordinate system discrepancies. Consequently, we had to manually construct the KTX station map for years before 2008 using the later-year layers. Changes in station openings and closures in 2004 (the railway's opening year), 2007, and 2008 necessitated the construction of new maps for these years.

From the centroids to the nearest KTX stations, we calculated driving time using network analysis toolkits. After testing several available options, we chose a method that does not require manual input of road network and quality attributes. Manual input of geographic data often causes network analysis failures due to slight mismatches among objects, incorrect attribute information, or slight distortions from geocoordinate system discrepancies. Indeed, the test run of the network analyses showed couple failures due to object geocoordinate mismatches. Therefore, we concluded that relying on a third-party network map, operated by professional geographers and network specialists, is the best option for our needs.

To calculate varying travel times by year, we attempted to obtain historical maps from various third-party services, including Google Maps and Naver Maps. However, we could not find sources that offer network map information from more than two decades ago. Most data sources provide real-time maps or geospatial information from less than five years ago, not from the early 2000s. Therefore, we calculated the OD driving time matrix with the current third-party network map information and used it for our treatment cutoffs.

To define the treatment group based on Japanese old rail stations from a century ago, we concluded that the current road network and driving time calculation are irrelevant. Therefore, the study calculated distance buffers from these old rail stations and chose a 50 km buffer as our treatment threshold. In constructing these buffers, we excluded the old rail stations located in the northwest of the country, as they are logically irrelevant for path-dependent station construction and are naturally far from the current station areas.

A.2 The Harmonization of Districts

During our analysis period from 2000 to 2016, there were many administrative changes. For example, Sejong was elevated to a level-2 administrative unit, becoming "Sejong Special Self-Governing City." There were also several major changes in the level-3 units of districts. To address these changes, we harmonized the districts into their most consolidated form. For example, with the consolidation of Chungju and Chungwon cities, we used the unified version of Chungju City for our 2000 through 2016 classifications. Ansan, Goyang, Yongin, Pohang, and some cities and districts were similar cases. We judged that merging pieces into one pie is a more objective identification than slicing it with an arbitrary ruler.

In such harmonization efforts, we found that the most dramatic reform was made in 2010. The significant change includes the consolidation of the three large cities of Changwon, Masan, and Jinhae into Greater Changwon City, marking a major shift in the level-3 administrative units in South Korea. Therefore, the study used the 2010 administrative units as a harmonization guideline. After submerging districts into the most consolidated version, and removing island districts of Jeju, Ongjin, Seoguipo, Sinan, and Ulleung, which are not accessible via ground transport, our final analysis districts reduced to 228.

Different data sources use different administrative codes. In South Korea, there are two versions of administrative codes provided by KOSTAT and MOIS.[12] The KOSTAT code is used purely for statistical purposes, while the MOIS code links administrative units to legal units, which are practically identical. Most of our data uses the KOSIS version, but the IMS data uses the MOIS version.[13] Therefore, we harmonized the MOIS code to the KOSIS version using the crosswalk file published by the Seoul Institute.[14]

A.3 Census on Establishments

The Census on Establishments is an annual business survey that gathers information from all 3.3 million establishments in South Korea. Beginning in 1994, this employer census has provided detailed business information for each establishment, including five-digit industry codes, geographic location, whether the establishment is part of a larger enterprise, and the number of full-time, part-time, and temporary employees by gender. The granularity of this information has enabled us to calculate several business performance indicators. The current version of the paper reports estimation results based on the number of firms and the number of female, male, and total workers by district, year, and industry. In future versions, we plan to add more indicators such as the number of new establishments and the number of female

[12] The two codes are different from the zip codes managed by Korea Post.

[13] KLIPS used their own classification, so we used their district names to merge the data.

[14] The published crosswalk is updated until 2013, so we first harmonized the data to the 2010 KOSIS codes and then converted the IMS MOIS code to the KOSIS version.

and male new hires by these establishments.

A few limitations stem from its large-scale survey implementation. First, the new firm indicator requires careful examination of the confidential raw survey responses. If a business was open after its spring survey cycle and exited before the spring next year, the survey was collected from an employer who opened another new business at the same address. Related, early years data reports that some business opening years are one year later than the data collection year.[15] Also, around five percent of the data contains only the enterprise demographics without worker counts. Therefore, we had to address the missing observations by imputing previous and next year's means. Despite this, the census is the most representative and credible data source for understanding sectoral, yearly, and worker compositions by districts. Therefore, we conducted our sectoral mechanism analysis using the Census on Establishments.

A.4 Population and Housing Census

The Population and Housing Census is a population census that collects demographic, residential, and usual travel information of all 50 million Koreans and foreign residents in South Korea. The survey cycle is every five years. The data contains individual-level information including demographics (i.e. age, gender, and education level), socio-economic status (i.e. employment status), and usual geographic and travel information (i.e. residential location, workplace location, commuting patterns, and the residential location of the respondent five years before the survey).

Our study uses two percent sample data from 2000, 2005, 2010, and 2015. We extracted the current residential location, residential location in the previous round (five years before), commuting location, commuting modality, and demographic characteristics such as educational attainment, gender, and marital status of the sample individuals. It is the most granular data on citizens and non-citizens in South Korea. However, the major limitation of the census is that it is only available every five years.

[15] On average, two to three percent of the sample shows the tendency from the year 2000 to the year 2010.

A.5 Korean Labor and Income Panel Study, KLIPS

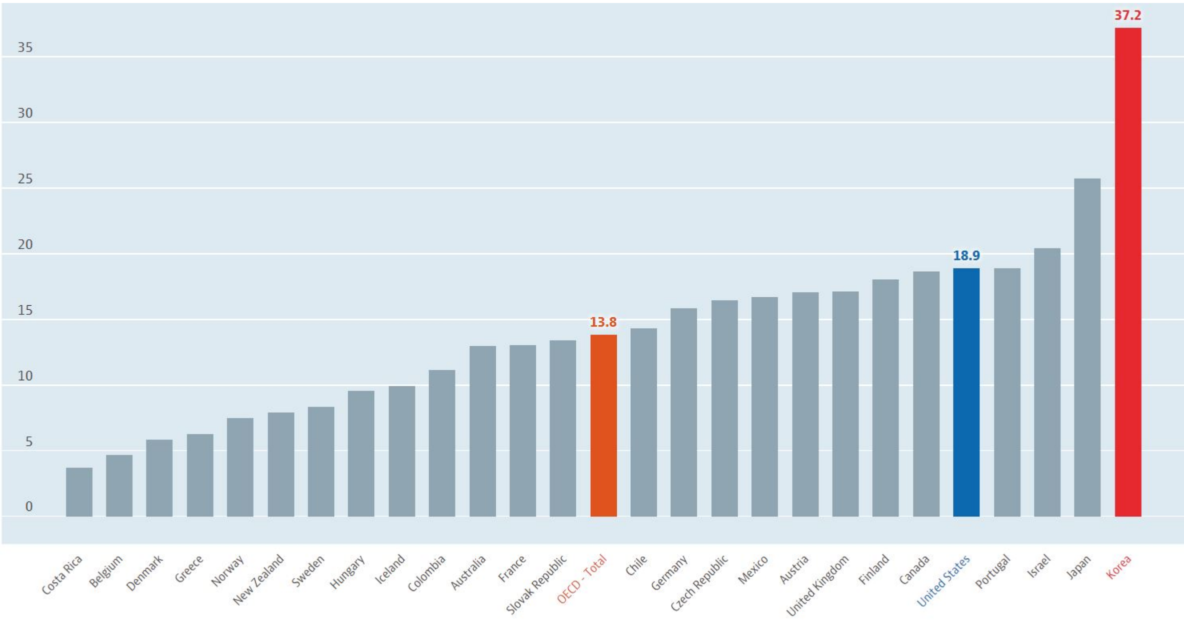
The Korean Labor and Income Panel Study (KLIPS) is a comprehensive longitudinal survey of the labor market and income activities in Korea. Since its inception in 1998, the Korea Labor Institute (KLI) has been collecting and managing the data. The initial sampling in 1998 extracted 5,000 households and approximately 10,000 to 15,000 affiliated members from the 1995 Population and Housing Census housing units, selecting samples only from urban districts. To address geographical limitations, a second batch of 1,400 households, primarily from rural districts, was added in 2009. Each year, KLIPS publishes two versions of data: household data, which contains detailed household information, including the demographics of members in all age and child-related expenses (e.g., educational expenses for both public and private institutions), and individual-level data, which provides extensive information on employment and assets for the sample above age 15, with nearly 550 variables covering various aspects of employment and earnings. In principle, these data sets offer the opportunity for the most rigorous measurement of wage and other earnings metrics in South Korea. In fact, there is bare alternative for wage data in South Korea.

Processing KLIPS turned out to be the most challenging part of our data harmonization. Initially, the sample did not cover rural districts until 2009, which limited our ability to construct sufficient non-Seoul observations until 2008. The omission of rural districts is a socioeconomically and geographically non-random decision, leading to sample selection issues in non-Seoul analyses. More importantly, a significant number of observations were dropped from the sample due to missing wage or weight information, further compromising its representativeness. The small sample size of only 15,000 workers required us to impose weights when constructing wage variables. However, observations with both wage and weight information constituted only one-third of the entire sample, around 5,000 annually. This reduced sample size not only increased the number of missing district-year strata but also undermined the credibility of weighting and district-level wage averaging.

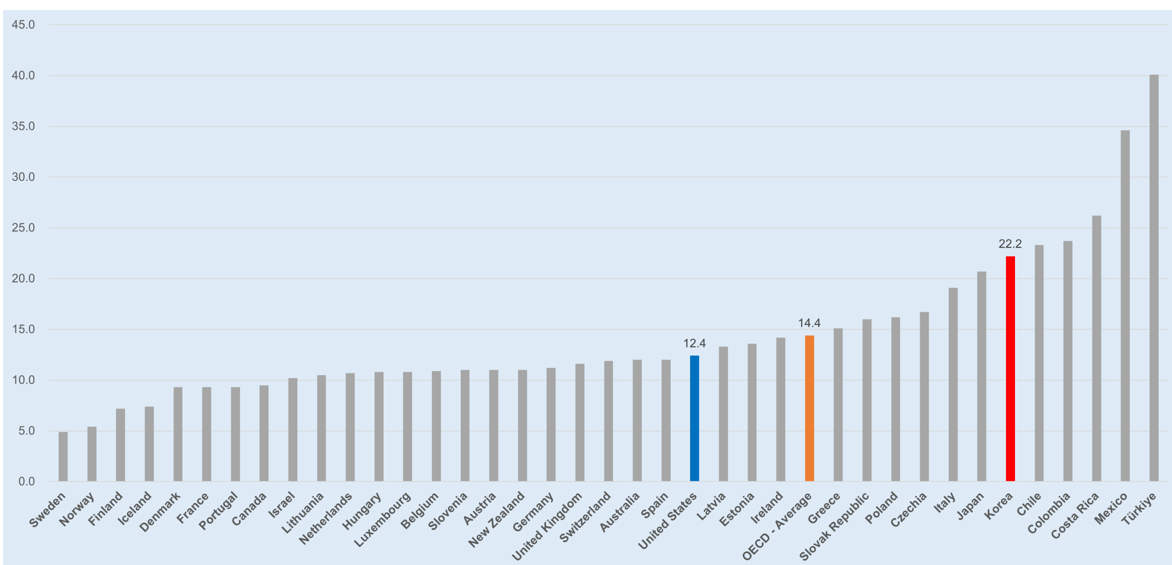
Given the importance of wage estimation in labor analysis and the lack of viable alternatives, we constructed weighted average real wage at the district level using KLIPS. Although the small sample size became our main concern, the data showed its merit with a decades-long panel structure, which offers both cross-sectional and longitudinal weights for time-consistent estimations. Using these two weights, we constructed both cross-sectional weighted and longitudinal weighted average real wages at the district level. The two variables show consistent results. Although the statistical power is weak for cross-sectional analysis, this is likely due to the small sample size rather than bias in the wage measure. We did real price adjustments for wages using the 2010 Consumer Price Index (CPI) from the World Development Indicators (WDI) The World Bank (2024).

Figure A1: Cross-Country Gender Gap in the Labor Market

(a) Gender Wage Gap (Percentage Points)



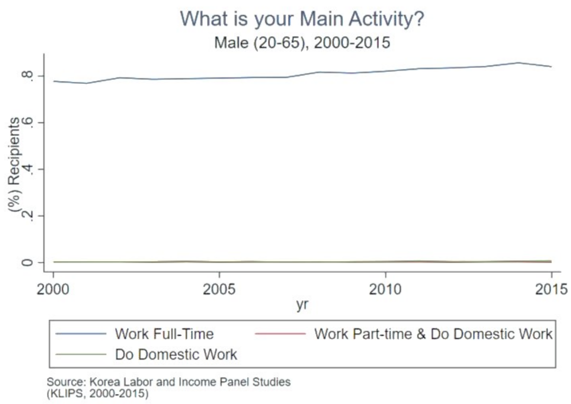
(b) Labor Force Participation Gap (Percentage Points)



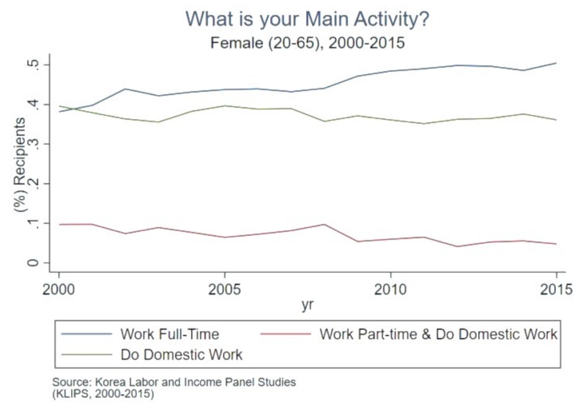
Notes: The gender wage gap in panel (a) is the difference between the median earnings of females and males. Here, the workers only include full-time employees and self-employed individuals. The labor force participation gap in panel (b) is the difference in the labor participation rate of males and females in paid work, including full-time workers. Source: OECD (2020).

Figure A2: Gender Differences in Time Use across Years

(a) Men



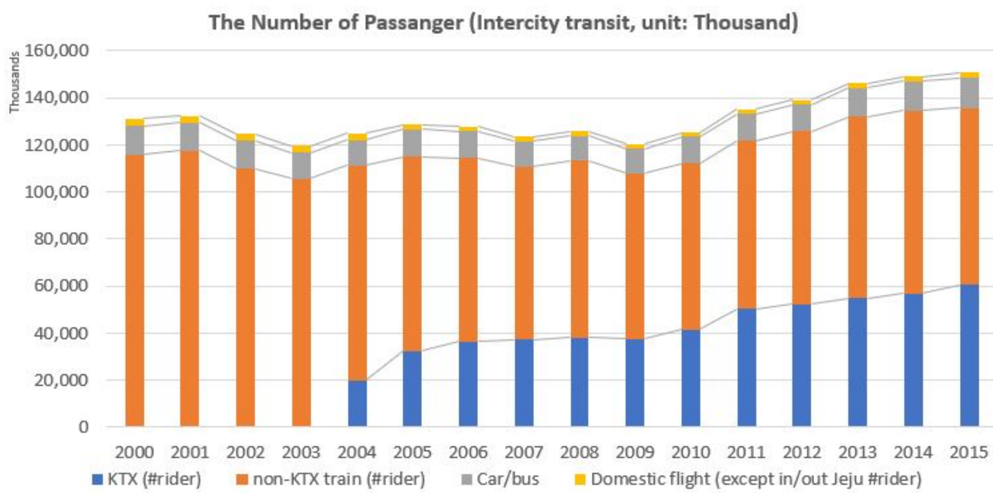
(b) Women



Notes: The graphs show the annual share of survey answers to the question, 'What is your main activity last month?' from the Korea Labor and Income Panel Studies (KLIPS), for men (Panel (a)) and women (Panel (b)). The answer options are: worked full time (blue bar), worked part time (red bar), and did domestic work (green bar). 'Other' is omitted from the graph.

Figure A3: Annual Ridership across Different Transportation Modes

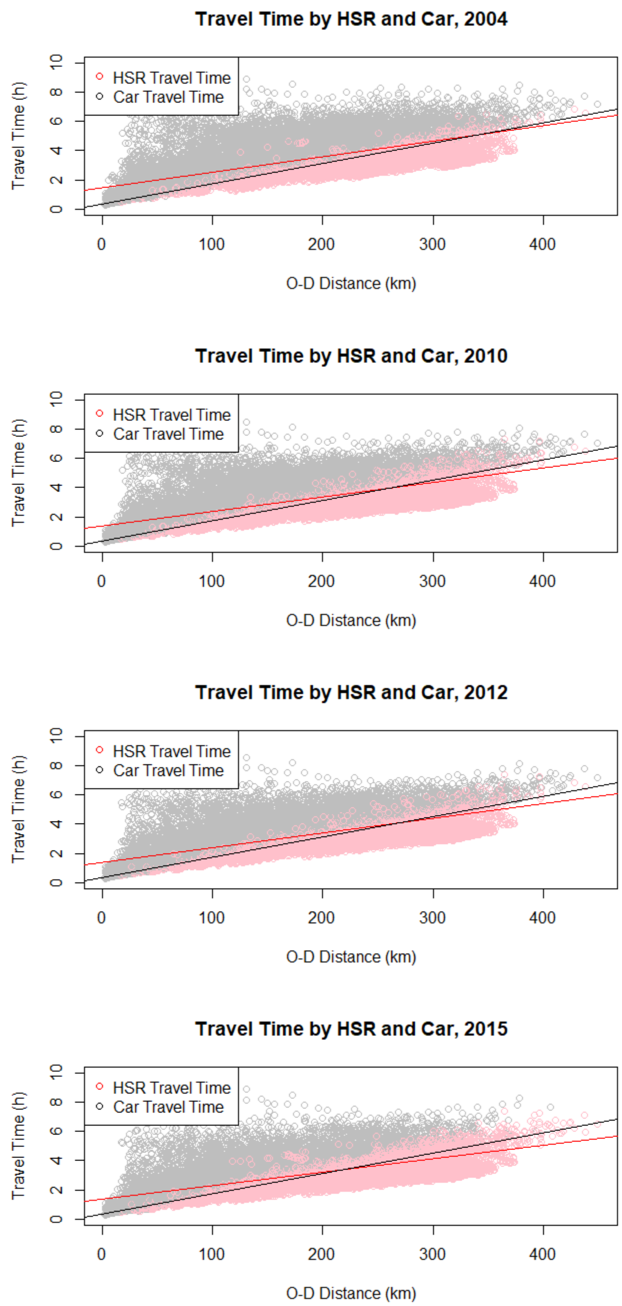
(a) Passenger Volume



(b) Total Passenger Travel Distance (KM)

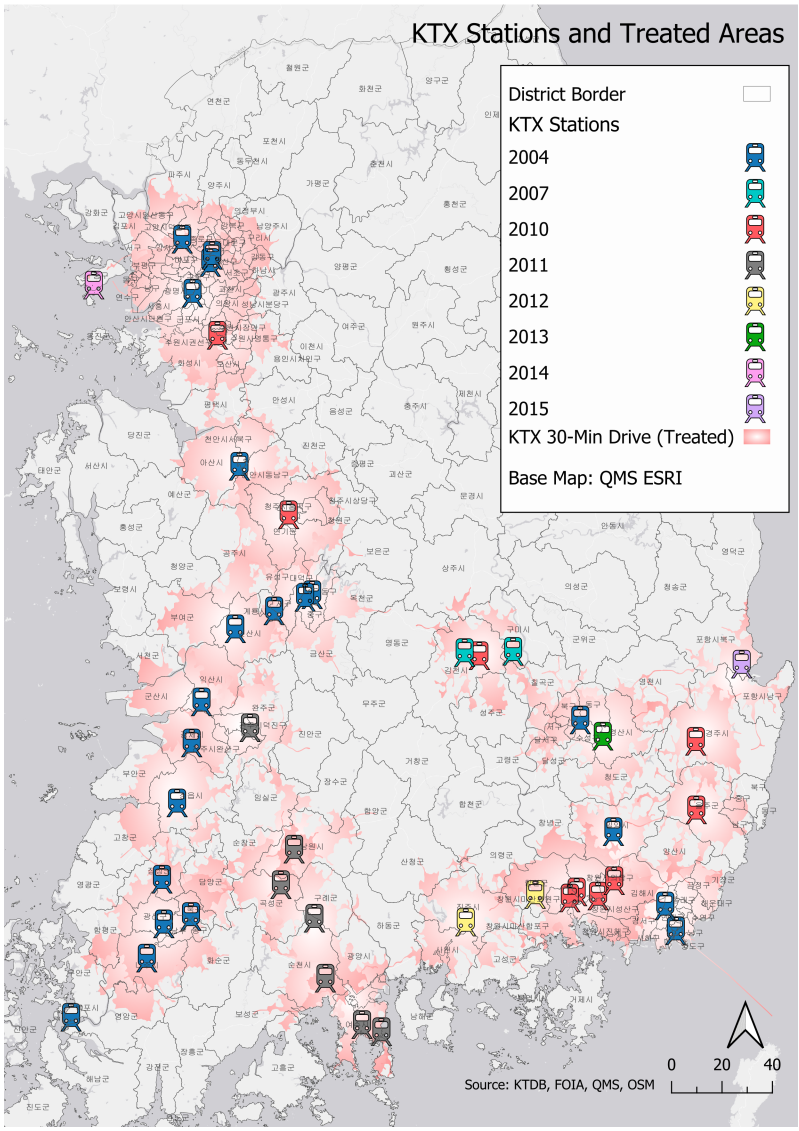
Notes: Source: Korea Transportation DataBase (KTDB).

Figure A4: District to District Travel Time by KTX and Car



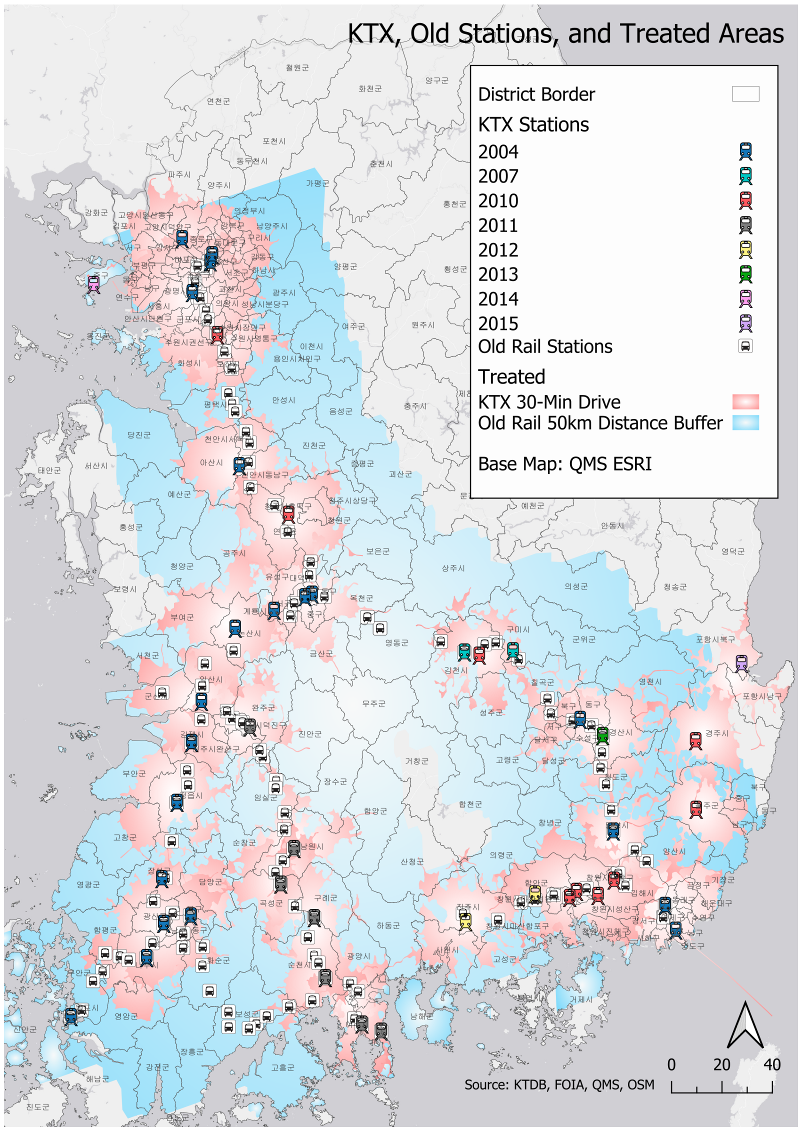
Notes: The red line represents travel time by KTX whereas the black line represents travel time by car. The X-axis denotes the geographic distance between OD districts. Grey dots indicate OD pairs where driving time is faster than HSR travel time, while pink dots indicate OD pairs where HSR travel time is faster. There is an increasing trend in OD pairs where HSR travel time dominates, especially over shorter distances. Source: KTDB.

Figure A5: KTX and Treated Areas



Notes: A district is treated if its centroid is within 30 minutes from the closest KTX station by car. Source: KTDB.

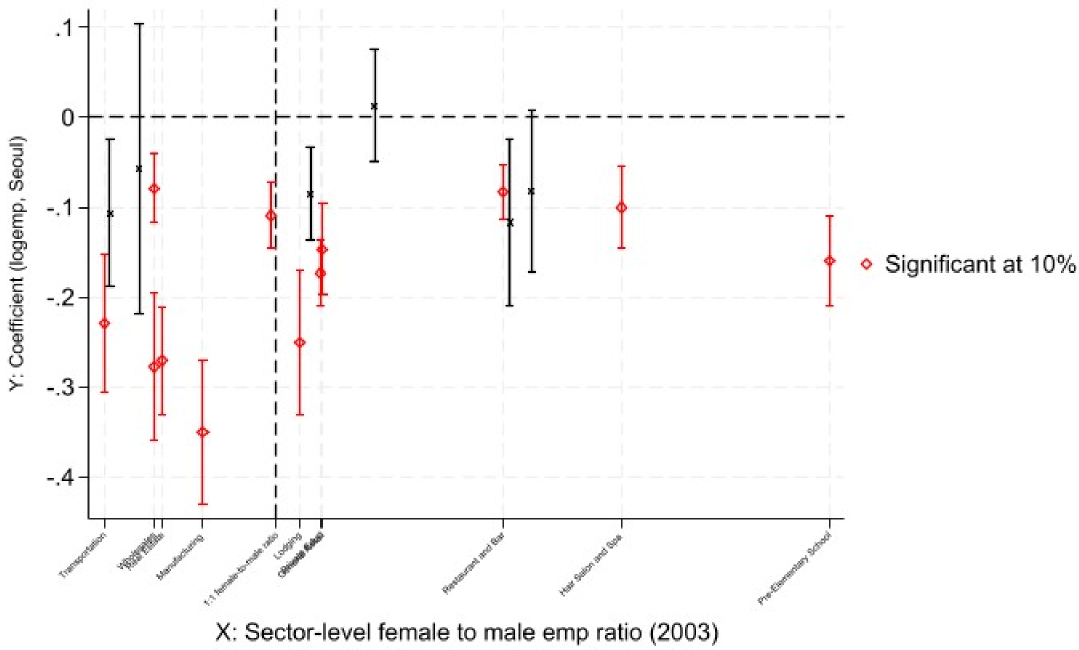
Figure A6: Overlap of KTX Treated Districts with Old Railroad Buffers



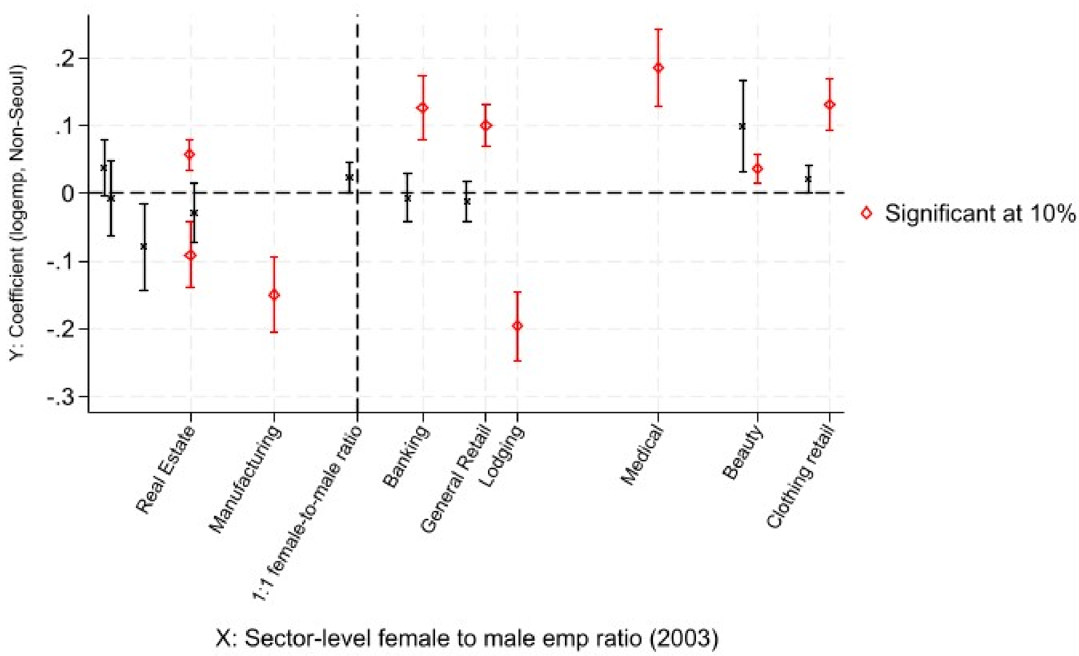
Notes: The map depicts the treated districts and together marking 50km buffer (brown shades) from old railroad stations.
Source: Korea Transportation DataBase (KTDB)

Figure A7: The Employment Effects of HSR by Sector

(a) Seoul



(b) Non-Seoul



Notes: The graphs plot the Average Treatment Effect for the Treated (ATT) using the method described by (Callaway and Sant'Anna, 2021) for each sector. The Y-axis represents the coefficient results from the subsample analysis of each industry, where the Y variable is the logarithm of employment. The X-axis shows the sector-level female-to-male employment ratio, calculated from the Census of Establishments in 2003. Red bars and diamond dots indicate regression results that are statistically significant at the 10% level. The coefficients and the statistics used in the graphs are also presented in Appendix A3.

Table A1: History of the KTX Station Expansion

Date	Line	From/To	Project Type	Station
2004 (+21)	Gyeongbu	Seoul/Daegu*	New line	Seoul, Hangshin, Gwangmeng, Cheonan-Asan, Daejeon, Dong-Daegu
		Daegu/Busan**	New line	Gupo, Milyang, Busan
	Honam**	Daejeon/Mokpo	New line	Yongsan, Seo-Daejeon, Gyeryong, Nonsan, Iksan, Gimje, Jeongeup, Jangsung, Gwangju, Gwangju-songjeong, Naju, Mokpo
2007 (+2)	Gyeongbu	Daegu/Busan	Add-station	Gumi, Gimcheon
2010 (+9, -2)	Gyeongbu*	Seoul/Daegu	Add-station	Suwon, Osong
		Daegu/Busan	Add/Close-station	Shin-Gyeongju, Ulsan, Gumi-Gimcheon, (Gumi), (Gimcheon)
			Improve speed	Busan
	Gyeongjeon**	Milyang/Masan	New line	Jinyoung, Changwon-Joongang, Changwon, Masan
2011 (+7)	Jeonla**	Ilsan/Yeosu	New line	Jeonju, Namwon, Guryegu, Sooncheon, Yeocheon, Yeosu-expo, Goksung
2012 (+2)	Gyeongjeon**	Masan/Jingu	Add-station	Haman, Jinju
2013 (+1)	Gyeongbu	Daegu/Busan	Add-station	Gyeongsan
2014 (+2)	Int'l Airport*	Incheon Airport/Seoul	New line	Gumam, Incheon International Airport
2015 (+1, -3)	Donghae*	Busan/Pohang	New line	Pohang
	Multi lines	-	Add/Close-station	Gongju, (Gimje), (Jangsung), (Haman)

Notes: *High Speed Railroad is defined to be maximum speed greater than $305 km/h$. **Electrified conventional railway directly connected with HSR ($V_{max} < 180 km/h$). Source: KORAIL.

Table A2: Definition of Industry

(1) Industry Name	(2) F-to-M ratio Seoul	(3) F-to-M ratio Non Seoul	(4) Seoul Emp.	(5) Seoul Share	(6) Non Seoul Emp.	(7) Non Seoul Share	(8) Included KSIC9 (3 digit)
Transportation and Logistics	0.120	0.111	465,346	0.080	423,316	0.070	all H
Construction, Utility	0.149	0.134	308,205	0.053	335,643	0.055	all D and F
Government	0.297	0.251	130,292	0.022	190,816	0.031	O841-O845
Wholesales	0.374	0.427	544,876	0.094	544,876	0.090	G451-G453, G461-G468
Other Services	0.376	0.408	186,875	0.032	223,893	0.037	S941, S942, S949, S951, S952, S953
Real Estate and Rental Service	0.416	0.413	239,009	0.041	139,470	0.023	all L but L694
Manufacturing	0.624	0.709	1,386,472	0.239	1,386,472	0.228	all C but C191, E383
Other Retail	0.974	0.973	357,417	0.062	421,095	0.069	G473, G475, G476, G477, G478, G479
Lodging	1.117	1.564	55,661	0.010	80,423	0.013	I551, I559
Food, Bev, Tobacco Retail	1.173	1.386	81,990	0.014	118,622	0.020	G472
Other Private Education	1.226	1.177	333,068	0.057	326,881	0.054	P854, P855, P856, R902, R911, R912
General Retail	1.236	1.453	167,172	0.029	185,463	0.031	G471
Banking, Finance, and Insurance	1.500	1.229	286,050	0.049	236,073	0.039	all K and L694
Restaurant and Bar	2.157	2.587	733,845	0.126	852,016	0.140	I561, I562
Medical	2.192	2.059	212,860	0.037	221,073	0.036	Q861, Q862, Q863, Q869
Textiles, Clothing Retail	2.303	2.660	112,989	0.019	126,384	0.021	G474
Hair Salon and Spa	2.764	2.407	100,799	0.017	113,831	0.019	S961
Pre and Elementary School	3.825	2.356	107,898	0.019	143,831	0.024	P851

The table lists the industry names used in our analysis and the 3-digit level industry codes included in each category, covering about 90% of the annual employment data in the Census on Establishments. The categorization is based on the industry codebook of the Census on Establishments. The definition of each industry is derived from the crosswalk between KSIC 8th (2000-2005) and KSIC 9th (2006-2015) at the three-digit industry level. We have dropped industries that have 8th and 9th industry codes that cannot be consistently matched at the 3-digit level. The list of industry names is sorted based on column (2).

Table A3: Industry-specific Regression Results for Seoul and Non-Seoul

(1)	(2)	(3)	(4)	(5)
	Employment Share		Gender Share	
Industry Name	Seoul	Non Seoul	Seoul	Non Seoul
Manufacturing	-0.024**	-0.019***	0.031	0.012
	(0.010)	(0.007)	(0.022)	(0.026)
Construction, Utility	0.002	-0.007**	-0.015*	-0.003
	(0.004)	(0.003)	(0.009)	(0.005)
Wholesales	-0.002	0.001	-0.035**	-0.001
	(0.004)	(0.002)	(0.014)	(0.022)
General Retail	-0.001	0.004***	0.045	0.071**
	(0.002)	(0.001)	(0.031)	(0.028)
Food, Bev, Tobacco Retail	0.001	-0.000	0.006	-0.003
	(0.001)	(0.001)	(0.032)	(0.035)
Textiles, Clothing, Leather Retail	0.000	0.002***	0.237**	0.121
	(0.002)	(0.001)	(0.103)	(0.091)
Other Retail	0.001	0.000	-0.010	0.068***
	(0.002)	(0.001)	(0.028)	(0.021)
Transportation and Logistics	-0.002	0.004*	-0.023*	-0.008
	(0.003)	(0.002)	(0.012)	(0.008)
Lodging	-0.010*	-0.004***	-0.032	0.058
	(0.006)	(0.001)	(0.046)	(0.061)
Restaurant and Bar	0.007**	0.003	-0.017	-0.066**
	(0.003)	(0.002)	(0.036)	(0.028)
Banking, Finance, and Insurance	0.005***	0.005***	0.003	0.085**
	(0.001)	(0.002)	(0.075)	(0.045)
Real Estate and Rental Service	-0.004**	-0.000	-0.018	0.103***
	(0.002)	(0.001)	(0.028)	(0.024)
Government	0.000	-0.006*	0.023	0.003
	(0.004)	(0.003)	(0.031)	(0.016)
Other Private Education	-0.002	-0.000	-0.054	0.067**
	(0.002)	(0.001)	(0.047)	(0.030)
Pre and Elementary School	-0.000	0.003**	0.083	2.356***
	(0.001)	(0.001)	(0.174)	(0.35)
Medical	-0.001	0.005**	0.027	0.076
	(0.002)	(0.002)	(0.074)	(0.047)
Other Services	0.002	0.002	-0.002	0.004
	(0.001)	(0.001)	(0.013)	(0.011)
Hair Salon and Spa	0.000	0.001*	0.143	0.159***
	(0.001)	(0.000)	(0.102)	(0.045)

Notes: The data is sourced from the Census on Establishments. Columns (2), (3), and (4) list each industry's 2003 statistics for the number of firms, employment, and the female-to-male employment ratio, respectively. Columns (5) and (6) present regression results showing coefficients for Seoul and non-Seoul regions, respectively, using specifications in equation (1).