THE IMPACT OF HIGH-SPEED RAIL ON LOCAL ECONOMIC ACTIVITIES: APPLICATION OF NIGHTTIME LIGHTS DATA

By

LEE, Hye Rin

THESIS

Submitted to

KDI School of Public Policy and Management

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF DEVELOPMENT POLICY

2020
THE IMPACT OF HIGH-SPEED RAIL ON LOCAL ECONOMIC ACTIVITIES: APPLICATION OF NIGHTTIME LIGHTS DATA

By

LEE, Hye Rin

THESIS

Submitted to
KDI School of Public Policy and Management
In Partial Fulfillment of the Requirements
For the Degree of
MASTER OF DEVELOPMENT POLICY

2020

Professor Baek, Ji Sun
THE IMPACT OF HIGH-SPEED RAIL ON LOCAL ECONOMIC ACTIVITIES: APPLICATION OF NIGHTTIME LIGHTS DATA

By

LEE, Hye Rin

THESIS

Submitted to
KDI School of Public Policy and Management
In Partial Fulfillment of the Requirements
For the Degree of
MASTER OF DEVELOPMENT POLICY

Committee in charge:
Professor Baek, Ji Sun, Supervisor
Professor Kim, Dongseok

Approval as of May, 2020
ABSTRACT

THE IMPACT OF HIGH-SPEED RAIL ON LOCAL ECONOMIC ACTIVITIES: APPLICATION OF NIGHTTIME LIGHTS DATA

By

LEE, Hyerin

China’s rapid and successful development of HSR network provides a sound empirical setting to analyze the effect of transport infrastructure on regional development. Did the HSR promote local economic activities in the areas near HSR stations? This study aims to empirically examine the changes in economic activities associated with the introduction of the Wuhan-Guangzhou HSR line on December 26th, 2009. For analysis, this study uses nighttime lights data from the U.S. National Oceanic and Atmospheric Administration (NOAA) to proxy for local economic activities and construct an 8-year panel dataset at the township level. Using a differences-in-differences approach, the study finds that on average, townships near HSR stations have seen an increase in economic activities by 9% four years after the launch of the HSR. A series of robustness checks confirms the main result.

Keywords: high-speed rail, economic activities, nighttime lights, differences-in-differences, China
ACKNOWLEDGEMENT

I would first like to express my deepest gratitude to my thesis advisor, Professor Jisun Baek, for her professional guidance and insightful comments through each stage of the process. I cannot imagine having come this far without the support of Professor Baek.

Apart from my advisor, I am grateful to all of my friends at KDI School of Public Policy and Management for the sleepless nights we studied together and for all the fun we had. I would always cherish the memories we made together. I owe a special thank you to my best friend, Songyi Shin, who kindly answered countless questions of mine regarding economics and made me laugh so much every day.

Most importantly, I wish to acknowledge the support and great love of my family: my grandfather, Jaeyoung Lee; my grandmother, Hwaja Kim; my father, Dongwoo Lee; my mother, Sookyung Han; my sister, Hyewon Lee. They kept me going by supporting through endless prayers and warm encouragements. This journey would not have been possible without their input.

Above all, I thank and praise the Lord God Almighty for being my strength when I am weak and for giving me constant love, grace and wisdom to persevere to the end.
# TABLE OF CONTENTS

1. Introduction .............................................................................................................. 1

2. Literature Review ........................................................................................................ 2

3. Background .................................................................................................................. 4
   3.1. China’s HSR Development ................................................................................. 4
   3.2. Wuhan-Guangzhou HSR .................................................................................. 6

4. Data and Summary Statistics ....................................................................................... 8
   4.1. Data .................................................................................................................... 8
   4.2. Summary Statistics ........................................................................................... 11

5. Empirical Strategy ...................................................................................................... 13
   5.1. Identifying Assumptions .................................................................................. 13
   5.2. Specifications .................................................................................................. 16

6. Empirical Results ....................................................................................................... 17
   6.1. Main Results .................................................................................................... 17
   6.2. Robustness Checks ......................................................................................... 18

7. Conclusion and Discussion ........................................................................................ 20

8. Reference .................................................................................................................... 21
LIST OF TABLES

Table 1. Descriptive Statistics ........................................................................................................ 12
Table 2. The Effect of High-Speed Rail on Local Economic Activities ............................. 18
Table 3. Robustness Checks .......................................................................................................... 19
LIST OF FIGURES

Figure 1. Map of “Eight Vertical and Eight Horizontal” HSR Network Plan .................. 5
Figure 2. Map of Wuhan-Guangzhou HSR Line .................................................... 7
Figure 3. Correlation between China GDP Growth (%) and Light Intensity Growth (%) ........ 9
Figure 4. Station Influence Area at the Prefecture Level ............................................. 9
Figure 5. China Railway Map .................................................................................... 14
Figure 6. Parallel Trends in Nighttime Light Intensity of Townships ............................... 16
1. Introduction

Transport infrastructure plays a critical role in promoting economic growth and development (Banister & Berechman, 2001, 2003; Boopen, 2006; Liu & Hu, 2010; Hong, Chu, & Wang, 2011). As an important transport innovation, high-speed rail (HSR) has become one of the most revolutionary transport technologies since the beginning of the 21st century, providing passengers with a reliable and safe intercity mass rapid transit service. Enormous carrying capacity, high operating speed, and connectivity are three main characteristics that define HSR, and these features substantially increase accessibility and enable a huge flow of capital, labor and goods to the areas around the HSR station. Hence, the introduction of a new HSR line significantly changes land use patterns (Xiao et al., 2006) in the areas near HSR stations, leading to changes in local economic activities.

Over the last two decades, China has experienced unprecedented development in HSR network. Following the rapid development of HSR lines, there has been a growing body of literature examining HSR impacts on Chinese cities. The HSR network increases accessibility (Cao et al., 2013; Luo, Xu, & Zhang, 2004; Shaw et al., 2014; Zhou, Yang, & Li, 2018), facilitates tourism industry (Wang et al., 2012) and affects housing prices (Chen & Haynes, 2015; Geng, Bao, & Liang, 2015). This body of literature offers useful insights into understanding the importance of HSR network on regional development in China in multiple aspects. However, empirical literature assessing the impact of each individual HSR line is still limited. The main contribution of this study is the construction of an 8-year panel dataset (2006-2013) from the nighttime lights data to measure local economic activities on a small spatial scale at a township level. Using night lights illumination from satellite observations, this study may be among the first research to identify the economic impact of the Wuhan-Guangzhou HSR line at the township level.

In this study, I aim to estimate the impact of the introduction of the Wuhan-Guangzhou HSR line on local economic activities, proxied by nighttime light intensity. Employing a differences-in-differences (DID) strategy, the study compares townships in the treatment group that are located within 10 kilometers (km) of the nearest HSR station and those in the control group that are located further.
than 10km, four years before and after the launch of the HSR line. The results show that economic activities in townships adjacent to HSR stations grew about 9% more after the introduction of the HSR.

The remainder of this study is organized as follows: Section 2 reviews the relevant literature, which will help to understand the existing literature on the impacts of HSR and the usage of nighttime lights data; Section 3 provides an overview of China’s history of HSR development and in particular, the Wuhan-Guangzhou HSR line; Section 4 discusses data and summary statistics; Section 5 presents the empirical strategy of the study; Section 6 shows the results of the DID estimator and robustness checks; and Section 7 draws conclusions and discusses policy implications.

2. Literature Review

Transport infrastructure is the key driving force of economic growth by reducing the time cost of travel and improving accessibility. Better accessibility affects the economic development of the regions connected with transport infrastructure in multiple aspects, for example, productivity and industry growth (Melo, Graham, & Brage-Ardaão, 2013; Nadiri & Mamuneas, 1996), total output growth (Ahlfeldt & Feddersen, 2018), spatial allocation of economic activity (Chandra & Thompson, 2000), trade facilitation (Duranton, Morrow, & Turner, 2014; Stone & Strutt, 2010), farmland value (Donaldson & Hornbeck, 2016), the real estate market and housing price (Du & Mulley, 2007; Henneberry, 1998; Zheng & Kahn, 2013), employment (Nelson et al., 2013), urban peripheral patterns (Baum-Snow, 2007; Qin, 2017), density and mobility (Cervero & Kang, 2011; Zhu & Diao, 2016), and tourism development (Masson & Petiot, 2009; Wang et al., 2012). The effects of transport infrastructure are not limited to the regions connected with a transport link, but the negative or positive spillovers (Moreno & López-Bazo, 2007; Yu et al., 2013) span other regions, changing the spatial structure (Baum-Snow, 2007) and economic, social and political dynamics in the neighboring regions.

Existing literature investigating the direction of transportation infrastructure spillovers - positive or negative - has opposing views. A study by Wang, Zhang, and Duan (2019) finds that the introduction of the HSR network creates unequal accessibility, with HSR cities winning from the shortened distance and better quality of service and cities with conventional rails relatively
disadvantaged. Also, Kim, Lee, and Park (2013) utilizes a neoclassical growth model confirming large-scale infrastructure projects such as HSR cause negative regional economic growth due to the Straw Effect\(^1\). Although there is a considerable amount of research claiming negative spillover effects of transport infrastructure, literature exploring the positive spillover effects of transport infrastructure dominates (Chen & Haynes, 2017; Jiang & Kim, 2016), which mainly focuses on how a transport link enhances the mobility of key production factors such as labor, knowledge and capital (Datta, 2012) and drives regional economic growth.

Research on the effects of HSR also vary. The primary mechanism by which the HSR could induce positive effects on local economies is by increasing accessibility that promotes development of region-specific industries, which increases the gross regional domestic product (GRDP). Chen and Silva (2013) review methods and models to analyze the link between HSR and regional growth and find that distance reduction between regions increases productivity, lowers production costs, increases employment and investment, accelerating local economic activities. Another important effect of HSR in local economies is an upgrade in regional industrial structure (Lu, Yu, & Han, 2016). The inauguration of an HSR network increases the number of tourists and business travelers, indirectly affecting the structure of urban industries. Lu et al. (2016) conducts a case study of Wuhan city to examine the impact of HSR passenger types on local economies, and finds a significant increase in activities particularly in the service industry. HSR also changes spatial patterns of regional tourism industry (Liang, 2010; Wang et al., 2012; Wang, Niu, & Qian, 2018; Zhou & Li, 2018). Wang et al. (2012) concludes that a mass HSR network construction in China transforms tourist markets and allocates urban tourism centers. On the contrary, research analyzing the negative effects of HSR network on regional economies mainly reflects straw effects, taking the form of growing regional inequality (Lu et al., 2014) and holdup in upgrading industrial structure (Wang & Li, 2014). In terms

---

\(^1\) The term ‘Straw Effect’ refers to the relocation of economic activities shifting from small- and medium- to larger cities due to the improved connectivity and reduced transport costs from the opening of transport systems such as HSR. The term was coined in Japan in the 1960s when Tokyo “sucked” core economic functions of other areas since the opening of the Tokaido Shinkansen in 1964.
of transportation, the introduction of HSR network raises competition concerns among different modes of transportation. This is because HSR enjoys comparative advantages with its huge capacity, rapid speed and high levels of safety than its counterparts such as air and conventional rail transport. Yao (2010) argues that the HSR industry is superior to the airline industry in terms of operating cost and shows the evidence of declining airline passengers after the launch of HSR lines.

These studies tend to focus its analysis on a relatively large spatial scale due to data constraints on a small spatial scale. This study adds on to the existing work by constructing an 8-year panel dataset using nighttime lights data to measure local economic activity at the township level. Recently, nighttime lights data has been widely used in urban (Li, He, & Wang, 2018), poverty (Noor et al., 2008), and regional economic research as a proxy for economic activities at the subnational level (Henderson, Storeygard, and Weil, 2012; Elvidge et al., 2009; Ghosh et al., 2010; Mellander et al., 2015). Henderson et al. (2012) supports using nighttime light intensity as a measure of economic activity, showing a strong relationship between growth in lights and GDP at the national level in their research. This study aims to find out the effects of the Wuhan-Guangzhou HSR line on local economic activities at the township level. The nature of the study design requires data on economic indicators at the township level. However, since the data at the township-level is limited, this study acknowledges the recent contribution of Henderson et al. (2012) and uses night light intensity to proxy for local economic activities.

3. Background

3.1. China’s HSR development

The history of HSR network development in China is short but successful. The development at the national scale began in 2004 when the State Council announced the Mid-Long-Term Railway Network Plan (MLTRP) in an attempt to upgrade passenger service and provide an additional capacity for an overloaded network (State Council of the People’s Republic of China, 2005). A national plan for HSR networks was first proposed in 2004, due to the rapidly increasing freight volume that burdens the railway capacity and the low speed of the existing railway limiting its competitiveness in passenger transport (Lawrence, Bullock, & Liu, 2019). To combat the problems, the Chinese government set the
goal to build 12,000 km of HSR lines and build a total of 100,000 km of railway network by the year 2020 that runs at a speed of 200 km per hour (km/h) and faster (Amos, Bullock, & Sondhi, 2010). The heart of China’s expansion into HSR is a “Four Vertical and Four Horizontal” network with four lines running north-south and the others crossing east-west, mainly consisting of passenger-dedicated lines (PDLs). In 2008, the target was revised upward to a total of 120,000 km railroads by 2020, with HSR lines accounting for 16,000 km (National Development and Reform Commission, 2008).

As of 2015, the initial goal for the year 2020 was met when the “Four Vertical and Four Horizontal” railway network was near completion, and total mileage of the railway and HSR surpassed the planned figures (Li, 2017). In response to the achievement, China upgraded the goal in 2015, announcing the plan to double the existing “Four Vertical and Four Horizontal” network to “Eight Vertical and Eight Horizontal” railway network that runs 175,000 km, including 38,000 km of HSR by 2025 (State Council of the People’s Republic of China, 2016). Figure 1 shows the proposed national grid of “Eight Vertical and Eight Horizontal” network, which forms the backbone of the HSR network in China.

![FIGURE 1 — Map of “Eight Vertical and Eight Horizontal” HSR Network Plan](http://www.gov.cn/xinwen/2016-07/20/content_5093165.htm)

*Notes:* The map shows the “Eight Vertical and Eight Horizontal” HSR network plan proposed in the 2016 Mid-Long-Term Railway Network Plan. The map is retrieved from the State Council of the People’s Republic of China website: [http://www.gov.cn/xinwen/2016-07/20/content_5093165.htm](http://www.gov.cn/xinwen/2016-07/20/content_5093165.htm)
At the end of 2019, China’s HSR network extended even further and reached 35,000 km in total length, spanning two-thirds of the HSR tracks around the world (China Railways, 2019). With the longest HSR network in the world, China is now pushing towards even more ambitious goals to expand its HSR network worldwide, strengthening China’s connectivity with the world (Lim et al., 2016). This rapid expansion of the HSR network has fostered substantial changes in the economic geography of China by notably improving accessibility among cities (Cao et al., 2013), and hence, assessment on the impact of HSR on regional economic development is important in guiding future decisions on long-term and short-term investment.

3.2. Wuhan-Guangzhou HSR

China’s Beijing-Guangzhou HSR line is the world’s longest HSR line in operation with a total length of 2,298 km (Xu, 2018). The HSR line runs with CRH380-AL bullet trains designed for a maximum operating speed of 350 km/h, and consists of four sections: Beijing-Shijiazhuang, Shijiazhuang-Wuhan, Wuhan-Guangzhou, and Guangzhou-Hong Kong.

This study puts special focus on the Wuhan-Guangzhou section of the current Beijing-Guangzhou HSR line, which is one of the earliest HSR lines in China that connects Wuhan, the capital city of Hubei province famous for rich labor resources, and Guangzhou, the capital city of Guangdong province known as the “world’s factory” due to its robust manufacturing sector. Construction of the Wuhan-Guangzhou HSR line commenced in 2005 and the operation began on December 26th, 2009. The total length of the line is 968km with a maximum operating speed of 350km/h (Xu, 2018). Figure 2 illustrates the route of the HSR line with the station influence area depicted in yellow, which will be discussed in detail in Section 4.1.
FIGURE 2 —
Map of Wuhan-Guangzhou HSR Line

Notes: The map shows the Wuhan-Guangzhou HSR Line with its 15 stations labelled aside. The station influence area is color-coded in yellow, which are all the townships under 14 prefecture-level regions the HSR line passes. The map is made by the author using QGIS.

With the opening of the HSR line, the travel time from Wuhan to Guangzhou has been shortened from at least 8 hours with the conventional rail transport to about 3 hours (Chen & Haynes, 2015). The transport system affects land use patterns (Eboli, Forciniti, & Mazzulla, 2012; Shaw & Xin, 2003), and therefore, the reduced travel cost and improved overall accessibility are likely to result in a more vibrant pattern of economic activities. This study is designed to use nighttime light intensity as a proxy for economic activities to explore whether areas around the HSR stations show a significant increase in economic activities, and understand the average size of the HSR impact.
4. Data and Summary Statistics

4.1. Data

The main source of data for this study is the nighttime lights satellite data provided by the U.S. National Oceanic and Atmospheric Administration (NOAA)\(^2\). This satellite imagery covers annual maps of nighttime lights of the world, collected by the Operational Linescan System (OLS) sensors installed on the Defense Meteorological Satellite Program (DMSP) satellites. Following most of the economic literature (Besley & Reynal-Querol, 2014; Hodler & Raschky, 2014; Michalopoulos & Papaioannou, 2014; Noor et al., 2008), this study uses the Stable Lights series from among the Version 4 DMSP-OLS Nighttime Lights Time Series (1994-2013). The Stable Lights series is an image dataset of nighttime light intensities of the earth’s surface averaged annually. In this series, observations affected by cloud coverage or ephemeral lights or background noise are replaced with the value of zero (Baugh et al., 2010). Each map is a raster data composed of pixels of 30 x 30 arc-seconds grid that corresponds to approximately one square kilometer. For each pixel, nighttime light intensity is reported in digital numbers (DN) ranging from 0 to 63, with higher values implying more intense light.

One potential caveat of using the Stable Lights series is that any DN values over 63 are truncated. With the values between 0 and the maximum of 63, the Stable Lights maps do not pick up differences between bright urban centers and their periphery. However, the fraction of pixels recording DN 63 for this study is less than 4%, and does not pose any major concern for the analysis.

The nighttime lights data is used to proxy for local economic activities, since every social and economic activity at night requires light. Adding on to the existing literature on the strong positive association between growth in nighttime lights and GDP (Henderson et al., 2012; Hodler and Raschky, 2014), this study also documents a similarly positive correlation at the level of subnational administrative regions in China in Figure 3, and thus, presumes nighttime light intensity well reflects the real local economic activities.

---

FIGURE 3 —
Correlation between China GDP Growth (%) and Light Intensity Growth (%)

*Notes:* The figure plots the correlation between GDP growth rate of China and nighttime light intensity growth rate for the period of analysis (2006-2013). Growth rates are the averaged value of annual growth rates. The numbers are computed by the author using data from China Premium Database (CEIC Data) and the U.S. National Oceanic Atmospheric Administration (NOAA).

FIGURE 4 —
Station Influence Area at the Prefecture Level

*Notes:* The figure illustrates the areas this study focuses on. The prefecture-level cities of the sample include Changsha, Chenzhou, Foshan, Guangzhou, Hengyang, Jingzhou, Qingyuan, Shaoguan, Shenzhen, Wuhan, Xiantan, Xianning, Yueyang, Zhuzhou listed in alphabetical order. The study selected a sample of 1,829 townships by selecting all the townships under 14 prefecture-level regions the HSR line passes. The map is made by the author using QGIS.
To create a township-level panel data using nighttime lights data, a geographical information system (GIS) software is used to construct an 8-year panel dataset with 14,632 observations at the township level. As illustrated in Figure 4, a sample selection of 1,829 townships is done by selecting all the townships under 14 prefecture-level regions the Wuhan-Guangzhou HSR line passes. However, due to data constraints on control variables, out of the 14,632 observations, 7,444 observations without missing control variables are used as the base sample for the analysis. For any fear of systemic missing values that may bias the regression coefficients, a robustness test is carried out in Section 6.2 using only the observations with missing control variables. If the coefficient estimate of the robustness test appears consistently positive and significant, then it can be concluded that the missing data does not pose a significant threat to the main estimation.

Although the study covers only about 4% of the country’s entire area, the selected regions play a vital role in China’s economy, given that the regions include the second (Guangzhou), fifth (Wuhan), and tenth (Changsha) largest cities in the country and their location in central China serve as an important transport and industry hub. For the analysis, the study excluded the townships that indicate specific buildings or lake or forests, and randomly excluded those overlapping. Also, ‘Banshichu (办事处)’ that refer to places where certain businesses or services are provided such as the Yong’an West Street Community Clinics (永安办事处) are not included in the sample.

To identify treatment status, this study employs the distance of a township from the nearest HSR station. To compute the distance measure, the locations of HSR stations were first geocoded using their GPS coordinates from Google Maps, and the center point of each township was recorded by determining the centroid of each township boundary. Then, the linear distance of a township from the nearest HSR station is calculated in kilometers (km) from the GPS coordinates of the nearest HSR station to the centroid of a township. Finally, two treatment variables are determined by the distance to the nearest HSR station: the first binary treatment variable used for the main analysis takes the value of 1 if the distance is less than 10 km, and 0 otherwise; and the second binary treatment variable extends
the treatment zone to 15km to test the robustness of the main estimation, taking the value of 1 if the
township is located less than 15m, and 0 otherwise.

Due to data limitations at the township level, data on control variables are collected at the
county level, assuming townships within a county roughly follow the same trends of the county. Data
on GDP, retail sales of consumer goods, fixed asset investment, government revenue and expenditure,
the number of industrial enterprises, population, the number of households, average wage are taken
from the China Premium Database (CEIC Data), and data on the county land area is obtained from the
China City Statistical Yearbook.

4.2. Summary Statistics

Panel A of Table 1 presents summary statistics for townships in the treatment group, and Panel
B provides information for the remaining townships in the control group before and after the
introduction of the Wuhan-Guangzhou HSR line. The first three columns describe the mean, standard
deviations, the number of observations for the pre-period (2006-2009), and the other columns are for
the post-period (2010-2013) of the analysis. There is a total of 14,632 observations, with 1,160 in the
treatment group and 13,472 in the control group, and the panel data are strongly balanced with a
consistent number of observations for both the pre- and post-period.
### TABLE 1 — DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td><strong>Panel A. Treatment Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nighttime light intensity (0-63)</td>
<td>35.27</td>
<td>19.86</td>
</tr>
<tr>
<td>Retail sales of consumer goods (in mil)</td>
<td>8.08</td>
<td>0.64</td>
</tr>
<tr>
<td>Government revenue (in mil)</td>
<td>5.97</td>
<td>0.85</td>
</tr>
<tr>
<td>Government expenditure (in mil)</td>
<td>6.98</td>
<td>0.57</td>
</tr>
<tr>
<td>Fixed asset investment (in mil)</td>
<td>8.37</td>
<td>0.9</td>
</tr>
<tr>
<td>Number of industrial enterprises</td>
<td>4.98</td>
<td>0.6</td>
</tr>
<tr>
<td>Average wage</td>
<td>9.81</td>
<td>0.37</td>
</tr>
<tr>
<td>GDP (in mil)</td>
<td>9.34</td>
<td>0.68</td>
</tr>
<tr>
<td>Population (in thous)</td>
<td>13.43</td>
<td>0.36</td>
</tr>
<tr>
<td>Number of households (in thous)</td>
<td>5.3</td>
<td>0.35</td>
</tr>
<tr>
<td>Area (km²)</td>
<td>7.55</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>Panel B. Control Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nighttime light intensity (0-63)</td>
<td>14.31</td>
<td>21.26</td>
</tr>
<tr>
<td>Retail sales of consumer goods (in mil)</td>
<td>7.73</td>
<td>0.78</td>
</tr>
<tr>
<td>Government revenue (in mil)</td>
<td>5.58</td>
<td>0.91</td>
</tr>
<tr>
<td>Government expenditure (in mil)</td>
<td>6.69</td>
<td>0.62</td>
</tr>
<tr>
<td>Fixed asset investment (in mil)</td>
<td>7.96</td>
<td>0.77</td>
</tr>
<tr>
<td>Number of industrial enterprises</td>
<td>4.64</td>
<td>0.95</td>
</tr>
<tr>
<td>Average wage</td>
<td>9.83</td>
<td>0.3</td>
</tr>
<tr>
<td>GDP (in mil)</td>
<td>8.91</td>
<td>0.78</td>
</tr>
<tr>
<td>Population (in thous)</td>
<td>13.34</td>
<td>0.55</td>
</tr>
<tr>
<td>Number of households (in thous)</td>
<td>5.2</td>
<td>0.56</td>
</tr>
<tr>
<td>Area (km²)</td>
<td>7.66</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Sources: Author’s calculation using data from the U.S. National Oceanic Atmospheric Administration (NOAA), China Premium Database (CEIC Data), and China City Statistical Yearbook

Notes: Townships located less than 10km from the nearest HSR station are in the treatment group, while those located more than 10km are in the control group. All monetary units are measured in China’s renminbi (¥).

In the pre-period, the means of nighttime light intensity show, in general, the local economic activities of townships in the treatment group are somewhat more than twice as vibrant than those in the control group. Other economic indicators - the retail sales of consumer goods, government revenue and expenditure, fixed asset investment, the number of industrial enterprises, average wage, and GDP - imply not much significant difference between the groups, with the treatment group marking slightly
higher numbers in most of the economic indices. The variables reporting regional characteristics – the number of households, population, and land area – show that townships in both groups in the pre-period are overall very similar.

In the post-period, the mean values tend to have increased in both groups, but with larger increase for those in the treatment group. The mean of nighttime light intensity of the treatment group jumped from 35.27 to 42.34 after the opening of the HSR line, while that in the control group also showed an increase from 14.31 to 17.2. For the mean values of economic indicators, values in the treatment group appear higher in numbers in all the indices. The regional-characteristics variables did not change much before and after the treatment in 2010.

5. Empirical Strategy

To identify the effects of the introduction of the Wuhan-Guangzhou HSR line on local economic activities, this study employs a differences-in-differences (DID) strategy and compares trends of nighttime light intensity between the treatment and control groups before and after the introduction of the HSR line in 2010. The treatment group consists of townships located within 10km from the nearest HSR station, whereas the control group consists of townships located more than 10km from the station. To address the endogeneity concerns, China’s double-track railway network is used to provide strong evidence for the exogenous placement of the HSR line. The parallel trend assumption for the DID estimator is also satisfied, confirming the trends in the outcome variable would have been parallel in the treatment and control group if there had been no treatment.

5.1. Identifying Assumptions

The first main issue to address in estimating the economic effects of infrastructure is the concerns regarding the endogenous placement of HSR stations (Beyzatlar, Karacal, & Yetkiner, 2014; Graham & Van Dender, 2011). Do areas around the infrastructure – in this case, the HSR station - show economic growth because of the infrastructure, or is infrastructure intentionally built in areas showing
the highest potential for economic growth? In the case of the latter, a simple regression estimating the impact of the infrastructure is likely to be overstated.

Estimating the effects of China’s HSR lines also calls for solving the endogeneity issue. Due to huge construction costs involved in building HSR lines, cities with the highest economic potential may have been chosen to produce sufficient ridership and attain higher investment returns. This concern is supported by the summary statistics presented in Table 1 in Section 4.2. The townships near the HSR stations were on average larger and richer compared with those further away from the HSR stations.

FIGURE 5 — China Railway Map

Notes: The map shows how the routes of Wuhan-Guangzhou HSR line and the conventional double-track line overlap. The red line indicates HSR line and connects Wuhan (Wuchang) with Guangzhou, passing the major cities of Changsha, Hengyang, and Shaoguan. The green line is the conventional double-track line completed in 1978 and follows a very similar route with that of the HSR line, connecting the major cities. The map was last updated in January, 2020. The figure is a screenshot of an online map plotting the passenger railway transportation system of China. Retrieved from http://cnrail.geogv.org/enus/?useMapboxGl=false
To address the potential endogeneity concern, the double-track railway network in China can act as evidence to support the exogenous placement of the Wuhan-Guangzhou HSR line. Figure 5 depicts how the routes of Wuhan-Guangzhou HSR line and the conventional double-track line overlap. The green line is the double-track railway network, which began its construction in the late nineteenth century and was completed in 1978. The validity of the double-track railway network lies in the fact that the network started and completed its construction before China’s economic reform in 1978, the year that marks a significant milestone of China’s transition from a centrally planned to a market economy. Prior to the economic reform in 1978, the double-track railway network primarily served as freight transportation, shipping raw materials and finished products rather than a source that promotes regional economic activities (Xue, Schmid, & Smith, 2002). If the HSR were to be built in areas with high economic potential, there has been ample time for economic activities to relocate following the dynamic changes in spatial patterns of China’s economic growth since 1978. Hence, the Wuhan-Guangzhou HSR line shown in red color in Figure 4 was built exogenously overlaid onto the existing railway network with a goal to enhance regional connectivity and accessibility rather than to connect cities with high demand for trade or migration. In other words, the double-track railway network is unlikely to be correlated with any other determinants of the growth in local economic activities, while it is highly likely to affect the placement of HSR stations.

Another assumption critical in applying the DID strategy is to test whether the difference between the treatment and control group is parallel over time, in the absence of treatment. That is, any unobserved characteristics affecting the nighttime light intensity do not vary over time, allowing townships in the control group to provide a baseline for determining what would have happened for those in the treatment group if the HSR line had not been launched.
FIGURE 6 — Parallel Trends in Nighttime Light Intensity of Townships

Notes: The figure shows the trends of averaged nighttime light intensity of townships in the treatment and control group between 2006 and 2013. The years are normalized to zero. Townships located less than 10km from the nearest HSR station are in the treatment group, while those located more than 10km are in the control group. The dotted vertical line represents the introduction of the HSR line on December 26, 2009. The numbers are computed by the author using data from the U.S. National Oceanic Atmospheric Administration (NOAA).

A graphical analysis is carried out to empirically probe the credibility of the parallel trend assumption. Figure 6 provides evidence of parallel trends, plotting the trends of nighttime light intensity of townships in the treatment and control group for the 8-year period (2006-2013). In the pre-period, townships moved almost in parallel in both the treatment and control group. After the introduction of the HSR, the slope of the treatment group increases much more sharply than that of the control group, and this sharper slope is the evidence of the HSR launch effect.

5.2. Specifications

The main objective is to find out the impact of the introduction of the HSR line on local economic activities. Therefore, the differences-in-differences strategy compares nighttime light intensity of the townships within 10 km from the nearest HSR station to that of the townships further away. The study therefore estimates the following equation:

$$Light_{ict} = \beta_0 + \beta_1 HSR_{ic} \cdot Post_t + X'_{ict} \theta + \tau_t + \eta_i + \epsilon_{ict} \quad (1)$$
where $Light_{ict}$ is the average night light intensity of township $i$ in county $c$ in year $t$; $HSR_{ic}$ is a treatment status dummy variable, indicating whether the township $i$ in county $c$ is located within 10 km of the nearest HSR station or not; $Post_t$ is a dummy variable stating the period after the year 2010 when the HSR began its operation; the interaction term of $HSR_{ic}$ and $Post_t$ captures the treatment effect of HSR; $X'_{ict}$ is a vector of control variables included to control for potential time-varying characteristics in the composition of townships; $\tau_t$ is year fixed effects; and $\eta_{ic}$ is township fixed effects. A panel dataset enables the study to include $\tau_t$ and $\eta_{ic}$ to control for the unobserved variables that are time-invariant and region-invariant and to compare economic trends, ensuring that the identification comes from the launch of the HSR.

The coefficient of interest is $\beta_1$ which measures the impact of the HSR on local economies based on the changes in nighttime light intensity. If positive impacts of the HSR dominate, then we expect $\beta_1 > 0$. If negative effects dominate, then we expect $\beta_1 < 0$. If the two opposing effects cancel out each other, or if there are no effects, then $\beta_1 = 0$ will be shown in the results.

6. Empirical Results

6.1. Main Results

Table 2 presents the results of estimating equation (1), the effects of HSR on local economic activities in townships that are located near the HSR stations. Each coefficient represents an estimate of $\beta_1$ from separate regressions, and all of them are statistically significant at the 1% level. The point estimates range from 3.15 to 3.6, with the coefficient estimate obtained with the richest set of controls reporting 3.15 in column (4). Using the mean night light intensity of townships near HSR stations in the pre-period, 3.15 increase in night light intensity implies that there was about 9% increase in economic activity from the launch of the HSR.
### TABLE 2 — THE EFFECT OF HIGH-SPEED RAIL INTRODUCTION ON LOCAL ECONOMIC ACTIVITIES

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;10km x post</td>
<td>3.60***</td>
<td>3.39***</td>
<td>3.22***</td>
<td>3.15***</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
<td>(0.94)</td>
<td>(0.89)</td>
<td>(0.84)</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Regional characteristics controls</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Economic-related controls</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.239</td>
<td>0.246</td>
<td>0.261</td>
<td>0.264</td>
</tr>
</tbody>
</table>

Notes: The table reports regression coefficients from 4 separate regressions from the 8-year panel data (2006-2013), with a township-level sample of 7,444. The dependent variable in all regressions is nighttime light intensity ranging from 0 to 63. Townships located less than 10km from the nearest HSR station are in the treatment group, while those located more than 10km are in the control group. All regressions include year and township fixed effects. Column 1 examines how nighttime light intensity of townships within 10 km respond to the HSR introduction in 2010, relative to the control group. To interpret the result, the coefficient estimate is divided by the mean night light intensity of townships in the treatment group in pre-period, and multiplied by 100 to calculate percentages change. Column 2 and 3 respectively includes regional characteristics controls and economic-related controls, and Column 4 reports coefficients with all specifications. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Column (1) reports that the estimated impact of the HSR without any control variables is 3.6. In column (2), the regression includes regional characteristics control variables such as the number of households, population, land area to control for factors that may influence the night light intensity, which adjusted the coefficient estimate down to 3.39. Column (3) shows the estimates with control variables indicating time-varying economic characteristics of each county – including the retail sales of consumer goods, government revenue and expenditure, fixed asset investment, the number of industrial enterprises, average wage, and GDP -, which reduced the coefficient estimate to 3.22. The last column suggests the main coefficient estimate of the study, showing that the HSR increased nighttime light intensity of the areas near the station by 3.15.

### 6.2. Robustness Checks

In this section, the study performs a series of robustness checks. One concern with this estimation is that the missing values of control variables that halved the original observations may be missing systematically, leading to a biased estimate. I test this concern by excluding the observations
that have missing values and running a separate regression with only those with full control variables. Column (1) of Table 3 shows the results with 7,188 observations; the size of HSR effect shrinks slightly, but the new estimate is statistically the same as the previous estimate at the 1% level.

**TABLE 3 — ROBUSTNESS CHECKS**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Townships with missing control variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larger treatment zone (&lt;15km)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add linear time trends</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add linear &amp; quadratic time trends</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo diff-in-diffs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base sample (2006-2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo sample (2004-2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10km x post</td>
<td>2.53***</td>
<td>2.60***</td>
<td>2.47***</td>
<td>-2.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.57)</td>
<td>(0.53)</td>
<td>(2.91)</td>
<td></td>
</tr>
<tr>
<td>&lt;15km x post</td>
<td>1.81***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>County FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Regional characteristics controls</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Economic-related controls</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Linear time trends</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Quadratic time trends</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.476</td>
<td>0.264</td>
<td>0.264</td>
<td>0.265</td>
<td>0.128</td>
</tr>
<tr>
<td>Observations</td>
<td>7,188</td>
<td>7,444</td>
<td>7,444</td>
<td>7,444</td>
<td>5,559</td>
</tr>
</tbody>
</table>

Notes: The table checks robustness of the main result with 5 different regressions. Each coefficient in columns 1-4 are from the base sample (2006-2013), and the coefficient in column 5 is from a subsample with an extended pre-period (2004-2009) for the pseudo analysis. The dependent variable in all regressions is nighttime light intensity ranging from 0 to 63. Townships located less than 10km from the nearest HSR station are in the treatment group, while those located more than 10km are in the control group, except for column 2 where treatment boundary of the HSR is enlarged to 15km. All regressions include year and township fixed effects. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Additionally, to test whether the size of the impact decreases with the distance, the study conducts a robustness test by enlarging the buffer zone of the treatment boundary from 10km to 15km. Column (2) of Table 3 shows the results including the same specifications as in equation (1). With a
larger treatment zone, the result is still positive, statistically significant at the 1% level; the size of HSR impact reduced to 1.81 from 3.15, as expected.

As another check, I added interaction terms between \( HSR_{tc} \), a dummy variable indicating 1 if the township is located within 10km to the HSR station, and a linear time trend to the model. These interactions capture any pre-existing trends across treatment and control group. As shown in column (3), adding the linear time trend again lowers the estimates to 2.6, but statistically significant at the 1% level. Column (4) adds quadratic time trend in addition to the linear time trend to allow for both non-straight and straight trend line. The result is not much different from the result in column (3) that only included linear time trend.

Lastly, for a further test on time trends, column (5) of Table 3 presents the results from the regression that uses a different sample of the years (2004-2009) before the launch of the HSR with a treatment year arbitrarily assigned. The year 2004 was chosen since the data on most of the control variables is not available before 2004, and the treatment year of 2006 was arbitrarily determined to have three-year period set after the treatment year, as it was in the main analysis. With the pseudo sample, I ran a regression on the same specifications as written in the equation (1). If the results are not picking up the time trends, one would expect the opposite result for the estimates from the pseudo sample. As expected, the regression in column (5) has insignificant and negative coefficient. These robustness checks reinforce the main finding of the study that the impact of the introduction of HSR was significant, increasing 9% of local economic activities.

7. Conclusion and Discussion

Investment in transportation infrastructure could be an important driver of economic growth, and China’s rapid and successful development of HSR network provides a sound empirical setting to analyze the effect of transport infrastructure on local economic activities. In this study, we estimated the effect of the Wuhan-Guangzhou HSR line on local economic activities using a unique panel dataset at the township level. With a DID estimator, this study finds that the introduction of the HSR led to a
9% increase in local economic activities on average in townships near HSR stations four years after the launch of the HSR service.

This study adds on to the existing literature by providing empirical evidence to support the theoretical hypothesis that rail network promotes regional development. The findings of this research could be helpful for policy makers in understanding the role of transport network in changing economic geography and in making rational allocation decisions of public expenditure. In the meantime, the findings of this paper do not fully explain the mechanism behind the increase in economic activities. It would be valuable in the future to collaborate the results of this study and identify whether the increase in economic activity comes from a shift of activity, so-called the Straw Effect, or from the generation of a new activity in the region. Furthermore, since the available DMSP-OLS nighttime lights data is only from 1992 to 2013 for the current moment, the study had to focus on four years of post-period. Hence, future studies may attempt to revisit the topic and examine the effects in longer post-periods when more recent DMSP-OLS nighttime lights dataset is updated in the near future.

8. Reference


Masson, S., & Petiot, R. (2009). Can the high speed rail reinforce tourism attractiveness? The case of the high speed rail between Perpignan (France) and Barcelona (Spain). Technovation, 29(9), 611-617.


