The Regional Economics Applications Laboratory (REAL) is a unit in the University of Illinois focusing on the development and use of analytical models for urban and regional economic development. The purpose of the Discussion Papers is to circulate intermediate and final results of this research among readers within and outside REAL. The opinions and conclusions expressed in the papers are those of the authors and do not necessarily represent those of the University of Illinois. All requests and comments should be directed to Sandy Dall’erba, Director.

The Impact of High-Speed Railway on Urban Growth: A Multi-Scale Spatiotemporal Approach

Haozhi Pan, Peng Chen, Sandy Dall’erba and Yina Zhang

REAL 18-T-4

May, 2018
The Impact of High-Speed Railway on Urban Growth: A Multi-Scale Spatiotemporal Approach

Abstract: The literature offers conflicting evidence on the impact of high-speed railway (HSR) operations on urban growth. This paper argues that the source of the differences lies in the variety of spatial scales under study. In order to address this conundrum, this paper studies the impact on urban growth of the start-of-operation of two HSR lines built in the Yangtze River Delta, China. Based on a high spatial and temporal granularity dataset and a non-parametric spatiotemporal model, the findings suggest that HSR operations have a significant positive impact locally (<5 km) but not at the county level. In addition, this paper finds that the local spillovers of urban growth decay drastically 3 to 6 months after the beginning of the operations. The findings support the development of economic activities linked to passenger transportation at or near the train stations.

Key Words: High-speed railway, Urban growth, Night-time light, Spatiotemporal model, Multi-scale, Yangtze River Delta region;

1. Introduction

Many high-speed railway (HSR) projects are built in many parts of East Asia and Europe. HSR projects improve the development of cities through better mobility and accessibility, agglomeration of economic activities, new projects associated with the HSR construction, and better input-output linkages (Banister and Berechman, 2002; Blum et al., 1997; Bruinsma et al., 2007; Chen and Hall, 2011a; Garmendia et al., 2012, 2012; Hirota, 1985; Levinson, 2012; Monzón et al., 2013; Ortega et al., 2012a; Van den Berg and Pol, 1998). However, several studies suggest also that a large part of HSR-related urban growth is not created, but rather shifted by HSR through a redistribution of wealth and opportunities across regions (Bruinsma et al., 2008; Hall, 2009; Hay, 1993; Levinson, 2012; Preston and Wall, 2008; Sands, 1993; Vickerman, 2015). As a result, the literature on HSR has not reached an agreement on how these projects bring economic prosperity and urban growth to localities that host HSR stations (Ortega et al., 2012; Sánchez-Mateos and Givoni, 2012; Ureña et al., 2009).
Part of the reason is that HSR’s effects vary by scale and over time (Ortega et al., 2012; Sánchez-Mateos and Givoni, 2012; Ureña et al., 2009) and the literature has not come to a consensus on which scale or time period is appropriate. For instance, several studies have emerged on 1) the areas in proximity to the HSR stations (Andersson et al., 2010; Garmendia et al., 2012; Monzón et al., 2013; Ortega et al., 2012b; Ureña et al., 2009; Van den Berg and Pol, 1998; Willigers, 2003), 2) the county or city scale (Banister and Berechman, 2002; Blum et al., 1997; Levinson, 2012; Ureña et al., 2009), 3) “intermediate” counties or cities (counties that are adjacent to or passed by HSR routes but do not have HSR stations themselves) (Hall, 2009; Levinson, 2012; Preston and Wall, 2008; Sasaki et al., 1997; Ureña et al., 2009; Vickerman, 2015), and 4) the regional scale (Chen and Hall, 2011b; Hirota, 1985; Nakamura and Ueda, 1989; Sands, 1993).

Furthermore, the studies above do not necessarily distinguish the effect of HSR on land use changes in the vicinity of a station from the effect on the economic development of the region over time. For example, the urban growth near HSR terminals and major nodes can be the result of the integrated layout of hotels, office, ecological park, shopping center, convention, and residential apartments, or so-called HOPSCA, a neologism used in the real estate advertisement industry in China. On the other hand, growth in larger regions (municipality, county, or metropolitan area) can result from an increase in productivity due to economies of scale in production, distribution and consumption from a better accessibility to a wider market base for diverse inputs (raw materials or labors) and diverse outputs (intermediate and finished goods). Therefore, it is necessary to conduct a multi-scale spatiotemporal study in order to better understand the true impact of HSR’s role and its evolution in both time and space. To our knowledge, this approach has never been applied to HSR lines and urban growth.

This paper processes urban nighttime light-level data in the Yangtze River Delta region to identify the growth pattern of land use from 2008 to 2014. To evaluate how the start-of-operation of HSR lines influences urban growth in the areas contiguous to a station, this paper compares light-level growth before and after the start-of-operation of two HSR lines as well as the light-level growth at different proximity levels to HSR stations. The paper starts with a set of descriptive statistics and a cross-sectional regression method to examine the role of the operation of HSR lines and accessibility improvement on various spatial scales (500x500 meters, 1x1 km, ...
Furthermore, this paper investigates the dynamics of such spatiotemporal spillover effects of urban growth over time so that it can offer a “3-dimensional boundary of impacts”. In order to proceed, this study relies on a non-parametric model built on the spatiotemporal covariance matrices of urban light-level growth on multiple spatial and temporal scales (1 to 12 months).

Ultimately, this paper verifies two hypotheses that are related to the capacity of HSR to spur urban development at two distinct spatial scales. The first hypothesis is that an HSR service results in an agglomeration of commercial and business activities around HSR terminals which is reflected in the presence of urban growth within a short distance. The second hypothesis is that the operation of HSR promotes the economic development of the region through gains in accessibility and spatial spillovers of growth emanating from the activities near the station. If this phenomenon is present, one should notice that urban growth takes place beyond the close proximity of the stations. Various distance cut-offs will be tested in this paper to clarify what distance is considered in the “proximity” to a station or not.

The rest of this paper is organized as follows. Section 2 provides a literature review on the relationship between HSR lines and urban growth and it also goes over the methodologies traditionally used when modeling spatiotemporal spillovers. The study area, the data, and the research method are introduced in Section 3. Section 4 presents the results and Section 5 concludes the paper with some additional remarks.

2. Literature Review

Urban land-use growth is directly generated by the expansion of economic activities (commercial land-uses) and the construction of residences (residential land-uses) (Deal et al., 2018). Night-time light levels are related but not equal to land-use growth (Shi et al., 2014) as the former encompasses other human activities such as energy consumption (Letu et al., 2010) and economic activities such as late-night retails and manufacturing (Li et al., 2013).

The construction of HSR lines is a factor that stimulates urban growth through multiple mechanisms. First, the construction of HSR lines and stations may result in changes in land-use
and restructuring of the urban space over time. For example, the urban growth near HSR terminals and major nodes can be the result of the development of commercial establishments and amenities. Land-use reconfiguration is often found in the urban areas located near railway stations (Monzón et al., 2013; Ureña et al., 2009). A research on several early HSR lines in Europe by Van den Berg and Pol (1998) shows that new urban projects and high quality services are attracted by HSR lines and stations. Willigers and Van Wee (2011) find out that the location choices of corporate offices are significantly affected by HSR lines. Some HSR projects are also associated with urban renewal projects in Europe (Bruinsma et al., 2007). However, the relationship is not always clear. For instance, the work of Ribalaygua (2005) shows that a series of commercial and housing developments associated with HSR projects in several urban sub-centers of Spain have had very limited successes.

When it comes to growth in larger regions (municipality, county, or metropolitan area), the effect can derive from an increase in productivity due to a better access to a wider market base for diverse inputs (raw materials or labor) and diverse outputs (intermediate and finished goods), as well as economies of scale in production distribution and consumption (Li et al., 2014). Levinson (2012) argue that HSR lines can connect cities and transform metropolitan systems into megalopolitan systems. Garmendia et al. (2012) point out that the most relevant contribution of HSR systems is the shortened travel time between HSR cities. Ortega et al. (2012) find that an increase in accessibility at the national and corridor level and a decrease in inequality of income across regions along the corridor resulted from the construction of the Spanish Galician HSR. For the area under study in this paper (the Yangtze River Delta), Chen and Haynes' (2017) results suggest that the middle reach of the Yangtze River Delta experienced a greater increase in accessibility than other HSR lines in China and that it has led to a greater increase in agglomeration of various industrial sectors there than in other parts of the Yangtze River Delta. Their results are in tune with those of Shao et al. (2017) on the same sample.

However, the accessibility and economic agglomeration effects of HSR lines on larger regions are challenged because these effects are found to have limited spatial spread. Sánchez-Mateos and Givoni (2012) analyze the HSR lines’ accessibility gain in the UK and show that the spatial spread of that gain is limited to the major urban conurbations in England. Hall (2009), Preston
and Wall (2008), and Sasaki et al. (1997) claim that new HSR lines have spatial impacts in urban cores but not in the periphery cities. Preston and Wall (2008), Ureña et al. (2009), and Vickerman (2015) point out to the lack of growth in intermediate cities (cities that are adjacent to or passed by HSR routes but do not have HSR stations) around HSR lines in Europe. They claim that wealth and opportunities are shifted to the cities that have HSR stations. The “redistribution” effect implies that HSR lines do not create enough opportunities and wealth far beyond the station location.

When it comes to HSR, Ortega et al. (2012) argues that the choice of the spatial level of analysis (international, national, regional or local) has strong consequences on the conclusions drawn, as one would expect from the well-known modifiable areal unit problem (Fotheringham and Wong, 1991). The accessibility analysis on the British and Spanish HSR lines by Sánchez-Mateos and Givoni (2012) concludes that a wide geographic area must be examined to understand true geographic spread of HSR lines’ impacts.

The spatial boundary of HSR lines’ impacts is determined by how the investment and accessibility improvement from HSR stations spill over to larger regions spatially and temporally. While the literature on the spatial spillover effects of public transportation infrastructure is fairly large, very few contributions focus on HSR. Gutiérrez et al. (2010) evaluate the spatial spillovers of transport infrastructure plan scenarios in Spain by using accessibility indicators as spatial weight matrices. Yu et al. (2013) use a spatial Durbin model and connectivity characteristics to evaluate spatial spillover effects of China’s transport infrastructure investments. Parent and LeSage (2010) set forth a space-time filter (spatial dynamic panel model) to study the highway induced travel demand. The above-mentioned spatial (temporal) spillover research finds positive spatial spillovers for transport infrastructure nationally or regionally for at least certain periods of time. Li et al. (2007) and Choi et al. (2013) present a non-parametric spatiotemporal covariance matrices sampling method. It has the potential to be implemented on high space-time granularity datasets and the results are simple to interpret. For these reasons, this study develops the non-parametric spatiotemporal modeling approach and applies it to the analysis of spatiotemporal spillover effects emanating from HSR stations.
3. Research Design

3.1 Study Area

This analysis focuses on the Yangtze River Delta, one of the main city-region in China. A city-region refers to the spatial extent of economic interaction over a wide geographic region that is well beyond the jurisdictional administration of any single city (Scott, 2001). Such extended metropolitan fields of interaction are closely linked by high-speed transport corridors stretching over several hundred kilometers from the core metropolitan centers.

The study area consists of 4 Provincial units (Shanghai, Zhejiang, Jiangsu, and Anhui) and 319 counties in the Yangtze River Delta region. Two major HSR lines, the Huhang-Huangning line (HH10 hereafter) and the Ninghang line (NH13 hereafter) started to operate in the region. HH10 has 30 stations in Shanghai, Zhejiang, and the Jiangsu Province, and NH13 has 11 stations in Zhejiang and the Jiangsu Province. The north part of the Jiangsu Province, the southern part of the Zhejiang Province, and the eastern part of the Anhui province are the areas in the Yangtze Delta that receive the least accessibility gain from HH10 and NH13 operations (distance to the stations are over 20km). In this study, they are identified as “untreated” spatial reference units.

In this research, 28 counties are identified as HH10 counties, and 8 counties are identified as NH13 counties. HH10 and NH13 form a triangle that circles 11 counties that do not have direct access to HSR stations but have easy access to adjacent counties for HSR services. These counties are identified as “intermediate” counties in this research. Counties outside this circle are not defined as “intermediate counties” because they are neither crossed nor encompassed by the two HSR lines in this research. The boundaries of counties in the study region and locations of HSR stations are shown in Figure 1.
Figure 1. Study area, county boundaries and HSR stations. Island territories are omitted because they are irrelevant to this research.

3.2 Data

3.2.1 Nighttime light data

In this study, urban growth is measured by the nighttime light level. Urban land-use growth is directly generated by the expansion of economic activities (commercial land-uses) and the
construction of residences (residential land-uses) (Deal et al., 2018). Night-time light levels are related but not equal to land-use growth (Shi et al., 2014). Nighttime light images derived from satellites can detect the artificial lights from settlements and other human activities at night, thereby providing spatially explicit and timely measurements of demographic and economic related urban expansion. Because of this salient advantage, nighttime light data have been used in a myriad of applications ranging from mapping urban sprawl (Elvidge et al., 2007; Henderson et al., 2003; Imhoff et al., 1997; Liu et al., 2012; Shi et al., 2014; Zhang and Seto, 2011), estimating electricity consumption (Letu et al., 2010) to monitoring impervious surfaces (Shao and Liu, 2014). Our datasets of the nighttime light level come from two sources. The first source is the Defense Meteorological Satellite Program’s Operational Linescan System (DMSP-OLS, DMSP hereafter). The DMSP provides the yearly measurements over 1992 to 2013 with a spatial resolution of 1x1km. This study uses this information to focus on the impact of the start-of-operation of the HH10 line. The second source is the Visible Infrared Imaging Radiometer Suite (VIIRS) day–night band (DNB) carried by the Suomi National Polar Orbiting Partnership (NPP) satellite (VIIRS hereafter). The VIIRS provides monthly data that ranges from April 2012 to present with a spatial resolution of 500x500-meters. The latter data is used to focus on the start-of-operation in NH13 HSR. Additional information about the data clearing and quality assurance process as well as about the advantages and disadvantages of studying urban growth with DMSP and VIIRS data will also be outlined in Section 3.3.1.

3.2.2 Railway passenger volume data and other socioeconomic data

Railway passenger volume data and socioeconomic data are obtained from the public Fudan University social science database (Zhang, 2017). This study uses passenger volume for the month of November (as available in the database) for both HH10 and NH13 from the start of HSR operations to the year 2014.
3.3 Methods

3.3.1 Method to process the DMSP and VIIRS data

Both temporal and spatial inherent limitations that could reduce the correlation between urban growth, land-use, socioeconomic activity and the nighttime light data need to be dealt with in order to investigate the urban dynamics accurately. Because DMSP uses multiple sensors and may encounter weather and cloud conditions, it is irrelevant to use raw data to compare the light levels across different years (Liu et al., 2012). Moreover, the DMSP dataset suffers from a saturation problem at the spatial level, as the nominal resolution of 1 km is resampled from the 2.7 km native resolution (Doll et al., 2000) and the DMSP sensor is only able to record light up to a certain purported value (i.e. the value of the average digital number ranges from 0 to 63 with zero being the absence of light) (Zhang et al., 2013). As a result, the intensity of light is not measured past a certain threshold value; hence the light emitted at the rural-urban fringe areas appear as the same level as that emitted in the urban core region.

To minimize the inconsistency problem of nighttime light data over years, this study employs several inter-period correction methods which are adapted from Liu et al. (2012). They are summarized in Equations 1 to 4 below.

First, a forward (Equation 1) and backward (Equation 2) correction helps ensure that light levels monotonically increase from initial to the ending year:

For t in range [2008, 2012)

\[
N'_{i,t} = \begin{cases} 
N_{i,t+1} & \text{if } N_{i,t+1} < N_{i,t} \\ 
N_{i,t} & \text{if } N_{i,t+1} \geq N_{i,t} 
\end{cases}
\]  

(1)

For t in range [2012, 2008)

\[
N''_{i,t} = \begin{cases} 
N'_{i,t-1} & \text{if } N'_{i,t} < N_{i,t-1} \\ 
N'_{i,t} & \text{if } N'_{i,t} \geq N_{i,t} 
\end{cases}
\]  

(2)

Then the difference of nighttime light levels are taken and standardized to mean 0 for each year:
\[ D_{i,t} = N''_{i,t} - N''_{i,t-1} \]  \hspace{1cm} (3)

\[ \bar{D}_{i,t} = D_{i,t} - \bar{D}_{i,t} \]  \hspace{1cm} (4)

where \( N_{i,t} \) is the nighttime light level of cell \( i \) at time \( t \) and \( D_{i,t} \) is the light level increase of cell \( i \) from time \( t \) to time \( t + 1 \). The same processing method is taken for the VIIRS data. In that case, \( t \) evolves on a monthly basis from April 2012 to August 2014. To mitigate the pixel saturation effect of the DMSP dataset, this paper removes all the cells with a saturated light level of 64 at any time \( t \). Figure 2 displays the result of DMSP and VIIRS data after processing.
Figure 2. Sample processing results of DMSP (saturation-removed) and VIIRS data

3.3.2 Measuring HH10 accessibility

This study calculates an accessibility index as a variable for the latter panel data regression model. The accessibility index used in equation (5) relies on both the passenger volume of the
HSR stations and the distance between each cell and the HSR stations. For cell $i$ at year $t$, accessibility of $a_{i,t}$ to the HH10 stations is calculated using Equation 5.

$$a_{i,t} = \sum_{s \in S} \mathbb{1}_{d_{i,s} < \tilde{d}} \frac{\log(v_{s,t})}{d_{i,s}}$$  \hspace{1cm} (5)$$

where $S$ is the set of all HH10 HSR stations; $d_{i,s}$ is the distance between station $s$ and cell $i$, $\tilde{d}$ is the cutoff distance (20km is chosen in this study as it is about the distance that a passenger in an intermediate county (defined in Section 3.1) needs to travel to reach the nearest station); $v_{s,t}$ is the passenger volume (the sum of on-boarding and de-boarding) at station $s$ and year $t$. The inverse distance model is similar to a gravity model.

3.3.3 Panel data regression modeling (DMSP light-level and HH10)

This paper investigates the impact of HH10 on urban growth over 2008-2012. First, a set of statistics are used to describe how aggregated light in counties with HH10 stations differ from light levels in the rest of the counties before and after the start-of-operation. Next, a panel data model of light’s yearly growth at the 1x1km cell level is constructed. The model is specified in Equation 6.

$$y_{i,t} = \beta_1 a_{i,t} + \mathbf{x}'_{i,t} \gamma_0 + \mathbf{\beta}_i + \delta_t + \delta_0 w_t + \delta_1 w_t a_{i,t} + u_{i,t}$$

$$u_{i,t} \sim N(0, \sigma^2)$$  \hspace{1cm} (6)$$

where $y_{i,t}$ is the light level increase at cell $i$ and year $t$ and $a_{i,t}$ is the HSR station accessibility treatment (specified in Equation 5) with coefficient $\beta_1$; $\mathbf{x}'_{i,t}$ is the matrix that includes per capita GDP growth between $t - 1$ and $t$ of the city where $i$ is located as well as urban-rural income difference ratio with coefficient array $\gamma_0$. The per capita GDP represents the development level of each county at each year while the urban-rural income difference is the main proxy for local disparities of development. $\mathbf{\beta}_i$ is a county-level fixed effect; $\delta_t$ is a time fixed effect; $\delta_1 w_t a_{i,t}$ is an interaction term between year fixed effect $w_t$ and accessibility treatment $a_{i,t}$. The interaction
term captures both the spatial and temporal “availability” of HSR to one cell at location \( i \) and time \( t \). Finally, \( u_{i,t} \) is an independent Gaussian error term with mean 0 and standard deviation \( \sigma \).

Equations 6 compares the growth rate of light level across treated cells and two benchmark groups: 1) treated cells but in the years prior to HSR and 2) cells that are too far-away (20 km) from any HSR stations during the post HSR years. Thus, the results can offer solid evidence on whether accessibility to HH10 stations on a 1x1km spatial level really contributes to higher urban growth rates.

The study tests for the possible presence of an omitted variable bias. The assumption is based on the fact that the HSR may be built in some cities to accommodate accessibility to certain urban renewal projects. If that is the case, the current model would fail to include urban renewal projects that yet have a correlation with both HSR operation and urban growth, and it would incorrectly allocate the urban growth effect from renewal projects to the HSR lines. In order to avoid this issue, this study uses an instrumental variable approach. More precisely, this study uses the percentage of railway passenger volume in all types of transport passengers as an instrument for accessibility. It is intuitive that the percentage of railway transportation does not relate to urban growth in one specific year, while it is highly correlated with the presence of HSR stations. Hausman-Wu and Wald F-test are applied in this study to test whether HSR accessibility is exogenous to light-level growth and whether the instrument used has a significant impact to the model results respectively. The Hausman-Wu test rejects the null hypothesis of exogeneity while the null hypothesis of weak instrument is rejected by the Wald F-test, which means that the choice of our instrument is warranted and a two-stage-least-square approach is required.

3.3.4 Nonparametric spatiotemporal autocorrelation modeling (VIIRS light-level and NH13)

High temporal (monthly) and spatial (500x500 meters) granularity of the VIIRS data allows this study to investigate the distribution of the spatiotemporal spillovers that follow the start of operation of the HSR. More precisely, this paper first verifies whether the impact of HSR on urban growth varies across spatial scales. The spatial scales included are 500x500 meter (the finest scales that contain a HSR station site), 1x1 km (a typical HSR station district planning
scale—such as Hongqiao Development Zone, Shanghai), 5x5 km (a typical size of municipalities in the region) and county level.

This study uses the estimation method for the spatiotemporal covariance matrix of Li et al. (2007) and Choi et al. (2013) to model spillovers at any point in time and across any location:

$$\tilde{C}(h, u) = \frac{1}{|S(h)||T_n|} \sum_{s \in S(h)} \sum_{t=1}^{n-u} \bar{Z}(s, t) \bar{Z}(s + h, t + u)$$  \hspace{1cm} (7)

where $\tilde{C}(h, u)$ is the sample auto-correlation of the spatiotemporal process with spatial distance $h$ and time distance $u$; $S(h)$ and $T_n$ are the numbers of observations within the space-time distance; $\bar{Z}(s, t)$ is the mean of all observations with spatial location $s$ and temporal location $t$.

This study reports the sampled spatiotemporal autocorrelation for different spatial (500x500-meter, 1x1-km, 5x5-km, county, intermediate county level) and temporal (1, 3, 6, 12 months) distances. In addition, this study reports both local (HSR station proximity) as well as global (all region) autocorrelations, as well as the degree of autocorrelation in all the study months and post-operation months of the NH13 stations.

Table 1 presents the definitions and descriptive statistics of various variables for both HH10 and NH13 models.

### Table 1. Definitions and descriptive statistics of the variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH10 (Year 2008 to 2012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light-Level Increase</td>
<td>Normalized light-level increase for every spatial cell by each month</td>
<td>0</td>
<td>2.395</td>
<td>-2.0</td>
<td>36.0</td>
</tr>
<tr>
<td>GDP</td>
<td>Per Capita Annual GDP in million RMB</td>
<td>0.054</td>
<td>0.025</td>
<td>0.117</td>
<td>0.017</td>
</tr>
<tr>
<td>Income_Diff</td>
<td>Urban-rural income difference ratio</td>
<td>2.265</td>
<td>0.353</td>
<td>9,500</td>
<td>22,234</td>
</tr>
</tbody>
</table>
Accessibility Monthly HSR passenger volume weighted by cell distance to the station 8.12e-4 1.86e-3 0 9.85e-3
Rail_Percent Share of all passengers that use railway as means of transport 0.043 0.065 0 0.460
Year_FE Fixed-effect for the time stamp of the data
City_FE Fixed effect for the city of the data

<table>
<thead>
<tr>
<th>Light-Level Increase</th>
<th>NH13 (Year 2012 to 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

4. Results

4.1 Growth of nighttime light-level at the county level

4.1.1 HH10 and non-HH10 counties

Figure 3 shows the county-level trend for the HH10 counties and non-HH10 counties before and after the start of operations (from 2008 to 2012). Each red line represents one HH10 county. The other lines are reference (non-HH10) counties. The wide green bar indicates the month of July 2010 when the operation of HH10 started. Visual validation indicates that counties with HH10 do not display a higher level of growth after the start-of-operation. The reasons why figure 3 also highlights NH13 counties are that: 1) it indicates more clearly that HH10 counties have less light growth after the HH10’s operation; and 2) it shows that NH13 counties experience greater light growth in the same period. The strong growth of NH13 counties is unlikely from the effect of HH10’s operation. A more plausible explanation is that the investments in urban infrastructure before the start of operation of NH13 are the source of the stimulus. This issue will be discussed further in Section 4.1.2.
In the two years and seven months before HH10 starts its operation, the average light level increase among HH10 counties was 0.145 (27 observations) while the non-HH10 counties experienced an increase of 0.258 (292 observations). Since the beginning of HH10 operation, HH10 counties experienced an average increase of 0.113 (81 observations) and non-HH10 counties experienced an average increase of 0.367 (876 observations). The t-test results indicate that the differences between HH10 and the rest of the counties are significant at the 5% level both before and after the beginning of HH10 operation. Furthermore, the region’s light level increase attributable to HH10 counties was 4.92% before the operation while it was 2.76% after
the operation. These descriptive results show that counties with HH10 stations do not experience faster urban growth than the rest of the counties.

The panel data regression results are reported in Table 2. The results show that the measure of HH10 station accessibility has a positive effect on nighttime urban light level growth at the 1x1km spatial resolution in all models. Another noticeable finding is that places with lower per capita GDP experience a greater increase in urban light-level. This result confirms the neoclassical growth theory (Barro and Sala-i-Martin, 1992) that predicts that poorer economies grow faster than the richer ones. The fact that places with lower per capita GDP and higher accessibility to HSR experience faster growth means that HH10 HSR successfully helps the poorer economies in the Yangtze River Delta Region close the gap with the richer places.

Table 2. Cross-sectional regression results for DMSP nighttime light level increase

<table>
<thead>
<tr>
<th></th>
<th>Light-level Increase</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Dependent variable:</strong></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>OLS</strong></td>
<td><strong>2SLS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-4.459*** -5.007***</td>
<td>-9.849***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.479) (0.479)</td>
<td>(0.495)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income_Diff</td>
<td>-0.736*** -0.749***</td>
<td>-1.856***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.029) (0.029)</td>
<td>(0.030)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accessibility</td>
<td>41.941*** 33.427*** 40.468*** 31.608*** 570.783***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.523) (2.501) (2.518) (2.495) (2.230)</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>City_FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year_FE:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accessibility</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Constant</td>
<td>2.098*** 0.571*** 1.853*** 0.266*** 4.599***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.068) (0.031) (0.063) (0.013) (0.071)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>966,620</td>
<td>966,620</td>
<td>966,620</td>
<td>966,620</td>
<td>966,620</td>
</tr>
<tr>
<td>R²</td>
<td>0.211</td>
<td>0.211</td>
<td>0.211</td>
<td>0.211</td>
<td>0.156</td>
</tr>
<tr>
<td>Residual Std.</td>
<td>2.127</td>
<td>2.127</td>
<td>2.127</td>
<td>2.128</td>
<td>0.156</td>
</tr>
</tbody>
</table>
4.1.2 NH13 and non-NH13 Counties

Table 3 reports the increase in light level at various spatial scales before and after NH13’s start of operation. The figures indicate that growth in the proximity of NH13 stations (within a 0.5, 1 and 5km radius) is significantly higher than in other spatial units (t-test p-value below 5%); and it is also higher before the operation of NH13. This evidence implies that investing in the construction of NH13 stations and in planning railway station special economic district around the stations spurs urban growth. At the county level, however, the NH13 counties lagged behind other counties in terms of urban growth but started to experience positive growth after the HSR. We do not find any statistical evidence of this effect for the HH10 lines. These preliminary remarks indicate that the majority of urban growth effects from NH13 stations are in the proximity zones (<5km). We do find that the stations display a positive impact on urban growth at the county level but their impact is not statistically significant.

Table 3. Average VIIRS nighttime light level growth at different spatial resolutions.

<table>
<thead>
<tr>
<th></th>
<th>0.5km</th>
<th>1km</th>
<th>5km</th>
<th>Intermediate Counties</th>
<th>NH13 Counties</th>
<th>Other Counties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before NH13</td>
<td>0.253</td>
<td>0.202</td>
<td>0.106</td>
<td>0.009</td>
<td>-0.002</td>
<td>0.003</td>
</tr>
<tr>
<td>After NH13</td>
<td>0.125</td>
<td>0.115</td>
<td>0.069</td>
<td>0.016</td>
<td>0.211</td>
<td>0.194</td>
</tr>
</tbody>
</table>

4.2 Nonparametric spatiotemporal autocorrelation

This study samples the spatiotemporal covariance matrices for different space-time distance combinations to understand the spillover patterns that take place between NH13 stations and their surrounding urban growth. A sample global covariance matrix for a 1km spatial distance is shown in figure 4. For every pair of cells located 1km away from each other, the results show the sampled global and local temporal autocorrelation for the time intervals of 0, 1, 3, 6 and 12 months. Sampling one covariance matrix (for example, 1km distance over 1 month) has the
computational costs of $O(n^2)$ where $n$ is the total light-level cell (500x500 meters) in the region with $n = 3,975,048$. A total of 44 (11 spatial scales and 4 temporal scales) covariance matrices are sampled. The huge computational challenge is addressed through parallelization of the processes.

**Figure 4.** Sample spatiotemporal global covariance matrix for a 1km spatial distance. The darker blue color indicates higher spatiotemporal autocorrelations. Diagonal lines indicate that the data with the same spatial gaps and temporal gaps (for example, all cell pairs located 1km from each
other and separated by 3 months to each other) are used to calculate $\bar{Z}(s, t)\bar{Z}(s + h, t + u)$, as shown in Equation 7.

The results of the spatiotemporal autocorrelation matrices are shown in Figures 5 and 6. The 8 matrices allow us to compare the spatiotemporal spillovers among the whole study area, within the NH13 counties, within less than 5km proximity zones of NH13 stations, and within the intermediate counties. Temporally, the figures are focused either on the whole study period or on the time period after the start-of-operation of NH13. This study aims to find out the qualitative differences between spatiotemporal spillovers patterns on different spatial and temporal focuses. Figure 5 shows the spatiotemporal autocorrelation of nighttime light level increase for the whole region and for the area in proximity of HSR stations by month (0 to 12 month) and by spatial units (0 to 5 km) before and after the NH13 operations. Note that the 0 km proximity zones of HSR stations are not sampled due to too few observations. One pattern that becomes obvious is that spatiotemporal autocorrelation decays dramatically at 5 km after 3 months and in any distance after 6 months. It is also clear that a pattern of seasonality emerges after a 12 month period as the level of autocorrelation goes back up. This is the yearly seasonality effect that can be expected in many real-life time-series dataset (i.e., sale volumes of seasonal goods and agricultural productions). Spatiotemporal autocorrelation levels are generally smaller after the NH13’s operation (figure 5b and 5d). It is also notable that spillovers at less than 1 km from the stations are high at any point in time.
Figure 5. Summary of spatiotemporal autocorrelation covariance matrices. Figure 5(A) captures the degree of spatiotemporal autocorrelation for the whole region (entire period); figure 5(B) focuses on the whole region but after NH13 operation only; figure 5(C) focuses on the areas at proximity of NH13 and for the entire period while figure 5(D) focuses on the areas at proximity of NH13 but after the NH13 operation only.
Figure 6. Summary of spatiotemporal autocorrelation covariance matrices continued. Figure 6(A) captures the degree of spatiotemporal autocorrelation for the counties with NH13 stations (entire period); Figure 6(B) focuses on counties with NH13 but after NH13 operation only; Figure 6(C) captures the degree of spatiotemporal autocorrelation for intermediate counties of NH13 (entire period); Figure 6(D) focuses on intermediate counties but after NH13 operation only.

Figure 6 displays the same results as figure 5 but at the county level (NH13 counties vs. intermediate counties). Even if the spatial scale has changed, the results confirm the previous ones in that the spatiotemporal spillovers decay dramatically at 5 km after 3 months and at any distance after 6 months. Intermediate counties display on average a higher level of spatiotemporal autocorrelation than the NH13 counties, which further illustrates that the HSR
stations’ operation does not have any significant impact on urban growth patterns for the counties that contain HSR stations.

5. Conclusion

The results on nighttime light-level and urban growth validate this paper’s hypothesis that investing in the construction of HSR station spurs urban growth at the local level. Nighttime light-level growth is significant at the 1x1km level for the HH10 HSR stations and at the 500x500m level for NH13 HSR. The areas located less than 5 km from the NH13 stations show significant growth even before the start-of-operation of NH13, which implies that investing in the construction of HSR stations promoted urban growth at the local level. This study also finds out that spatiotemporal spillovers of economic activities are most significant within 3 months and in a radius of less than 5 km. This study hypothesizes that it is a decrease in construction-related activities near or at the station after the start-of-operation that results in lesser spillovers of urban growth.

On the other hand, an increase in economic efficiency and accessibility of HSR does not show any significant impact on urban growth at the county level, which contradicts our second hypothesis. Indeed, this study finds that HH10 station counties have lower nighttime light-level increase both before and after the operation of HH10 compared to non-HH10 counties. When it comes to the NH13 station counties, they do not have a higher nighttime level growth than non-NH13 counties before and after the NH13’s operation either.

Two major policy implications are derived from the conclusions above. First, the construction of and the investment in HSR station have significant impacts on urban growth at and near the stations (less than 5km). This result is in line with similar arguments from previous studies (Monzón et al. 2013; Ortega et al., 2012; Ureña et al., 2009; Van den Berg and Pol, 1998). As such, HSR stations are more than just a transportation hub for the local economic development. The investment and economic activities brought in with the construction of the stations can generate growth in a radius of up to about 5 km. Thus, it is important to strategically locate the HSR stations as they act as potential local growth drivers. One logical way to extend the growth benefits to a larger area is to support a station-specific economic district planning to fully capture
the HSR stations’ economic aggregation effects. Second, the results indicate that HSR has limited effects in helping local (for example, county level) economic prosperity. This result is in tune with the findings of Willigers (2003) and Ribalaygua (2005) about the lack of impact of HSR on regional-level business and firm growth. Moreover, a recent HH10 and NH13 passenger exit survey by Zhang (2017) indicates that the median commuting time after de-boarding the HSR trains is over 30 minutes. Since the large majority of such trips is for business purposes (44%), the contribution to urban growth in close proximity to the station is very limited. As such, policymakers should consider HSR as a lesser efficient tool for economic development than more traditional strategies such as industry clustering, products differentiation and investment in education (Malinberg and Maskell, 2002; Rodríguez-Pose, 2013).

From a technical viewpoint, this paper demonstrates the applicability of the space-time covariance matrix sampling method proposed by Choi et al.’s (2013) and Li et al.’s (2007) to very large volume datasets. While other econometric methods, such as the dynamic spatial autoregressive model (Debarsy et al., 2012; Parent and LeSage, 2010, 2012), can be applied to space-time panel data, the latter is not scalable to a fine-resolution large dataset like light-level data used in this study. This spatiotemporal covariance matrix approach can be applied to many other large datasets. However, a word of caution is needed as in the current research we used more than 100,000 core-hours on supercomputing resources (20 cores CPU with 256GB RAM) that are not necessarily available to many researchers. We therefore recommend parallel computation instead of more traditional methods. The earlier is somewhat easy to set for case studies like ours because different intervals of space-time autocorrelation can be computed independently.

Reference