6225 words + 5 tables = **7475 word counts**

Regional and Local Economic Effect from Proximity of High-Speed Rail Stations in Japan: Difference-in-Differences and Propensity Score Matching Analysis

Jetpan Wetwitoo (Corresponding author)
Creative Research and Planning Co., Ltd.
6Fl, Mitomi Building, 1-20-22, Ebisu, Shibuya-ku, Tokyo 150-0013, Japan
Phone: +81-3-5791-1133; Fax: +81-3-5791-1144
E-mail: j_wetwitoo@crp.co.jp

Hironori Kato
Department of Civil Engineering, The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
Phone: +81-3-5841-7451; Fax: +81-3-5841-7496
E-mail: kato@civil.t.u-tokyo.ac.jp

ABSTRACT

This paper gives empirical evidence about the economic effect of proximity of high-speed rail (HSR) on regional/local production and labor productivity in the case of Japan. The effect on a regional scale and a local scale are analyzed based on a Difference-in-Differences (DID) method and Propensity Score Matching (PSM) method. The prefecture-level analysis investigates the effect on regional production in prefecture-level during 1981-2006 and the municipality-level analysis investigates the effect on the local tax revenue and the tax revenue per capita in municipality level, particularly focusing on new HSR extensions during 2010-2015. The results from both levels show statistically insignificant estimations of both DID effect and the Average Treatment Effect on the Treated (ATT) from PSM, which implies that, on average, no direct effect of the proximity to HSR service on the regional/local production nor productivity. One of the potential reasons, the effective range of HSR service to the local economy, are discussed. Based on the local characteristics where HSR station is located, the effective range of HSR service could vary from 10 to 30 km.

Keywords.

High-speed rail; proximity effect, Difference-in-Differences (DID); Propensity Score Matching (PSM); regional economic development; Japan;

INTRODUCTION

The high-speed rail (HSR) has been regarded as one of the main inter-city modes of transportation in many developed countries in Europe and Asia. Recently, the rapid development of extensive HSR network in China has caught an attention from many developing countries since it coincided with the rapid economic growth in China in the past few decades *(1)*. Not only the case of China, but the contributions of HSR to economic development in developed countries have been reported by many studies *(2, 3)*. Yin et al. *(1)* provides many great examples of how HSR affects the development from several perspectives.

The effect of HSR to the economy could be observed with many indicators, such as the reduction of travel time *(4)*, change in land use pattern *(5)*, as well as the employment effect *(6–8)* and migration effect *(9, 10)*. However, the production, determined by the gross domestic product (GDP), could be the most straightforward indicator to determine the contribution from HSR to the economy. The positive effect of HSR to production, typically measured with GDP, and the positive effect on productivity, typically measured with a labor productivity such as GDP per employment, have been reported from many countries. Generally, studies from European countries have reported the positive effect of HSR. For instance, in France, GDP per capita in the region with HSR service tends to be higher with the expense of lower GDP per capita in an adjacent region *(11)*. In Spain, the investment of HSR positively affects GDP and employment level through better accessibility *(12)*. A similar argument was also suggested in the case of UK HS2 project where HSR investment could boost the economy through an agglomeration impact *(13)*. The availability of HSR is also reported positively affecting regional GDP in South Korea *(10)*. However, empirical studies on Japanese HSR have displayed quite mixed results. Sasaki et al. *(14)* concluded HSR network expansion contributes to regional development in non-core areas to some extent, yet it strengthens the development in core areas in a long run. Bernard et al. *(15)* reported that firms located close to HSR service tend to perform better, especially those located within 30 km from HSR station. Wetwitoo and Kato *(16)* also found that the HSR effects on productivity are significant in the early stages of introduction where HSR network was still limited in the core areas of Japan.

This study highlights expected impacts particularly from the proximity of HSR stations on the regional economic productivity, or the proximity effect of HSR. The proximity effect means that the economic impacts in the zones that are located within the buffer of HSR stations. In many countries where HSR has been already introduced, local municipalities along the existing HSR lines often request an introduction of additional stations into their municipalities since they expect an improvement of accessibility for local people/business as well as its indirect impacts on the local economy through the better accessibility and/or existence of HSR station. One example in Japan is an introduction of a new station between Shin-Yokohama and Odawara along the Tokaido Shinkansen[1] line, which has been proposed by local municipalities since 1975, but it has not yet been realized. The main issue in proposing the introduction of new HSR stations is the financial burden of local governments for the construction of new infrastructure. Thus, the local/regional economic impacts should be one of the major concerns for many local societies to justify an expensive spending in the construction of new HSR station facilities. However, for example, Wetwitoo and Kato (*16*) showed that the HSR effects are found positive with respect to the network utilization and the agglomeration effect but the presence of HSR stations does not affect the productivity. Does the proximity of HSR stations really have no impact on economic productivity?

Although several estimation techniques have been utilized in past studies, many of them failed to clearly present the relevant outcomes by comparing the situation before an introduction of HSR/HSR stations

[1] The word for "Japanese High-Speed Rail"

with that after. One of the simple, but powerful approaches to solve this issue is to apply the estimation under the framework of Difference-in-Differences (DID) and Propensity Score Matching (PSM). DID has been widely applied in social sciences, including transportation-related studies in urban scale *(17)*, in regional scale *(18)* as well as the effect of HSR introduction in Japan *(19, 20)*, Spain *(21)*, Uzbekistan *(22)*, and China *(23)*. Similarly, PSM originally introduced in social sciences studies, and later applied in transportation-related studies *(24–26)*. PSM was also applied by Komikado and Kato *(27)* to investigate the effect of HSR on regional innovation. These two matching methods will be utilized in this paper to investigate the effect of the proximity of HSR stations to regional/local production and productivity in the case of Japan. The long history of HSR service, combined with the continuous network expansion in Japan could be one of the unique characteristics to the robustness in the results where further investigation would be appealing to understand the effect of HSR to the economy.

The paper is organized as follows. Next section discusses a concept behind both matching technique and reviews the empirical evidence where DID and PSM have been utilized in the studies related to HSR. Next, the empirical data used in this study and definitions of treatment and control groups in DID and PSM analysis are shown. The empirical analysis section where DID and PSM are both applied includes the analysis in prefecture-level during 1981-2006 and the analysis in municipality-level, focusing on the regions with recent HSR extensions during 2010-2015. The conclusions section summarizes the findings from the empirical analyses and further issues to be discussed.

METHODOLOGY

Difference-In-Differences Design

This study assumes a proximity to HSR as one of the major policies to promote regional development in Japan. DID is applied to evaluate the effect of introducing HSR stations by comparing two situations: before the introduction and after the introduction. DID is an econometric technique that estimates the policy effect from the difference between a treatment group (regions/municipalities in which HSR stations have been newly introduced) and a control group (other regions/municipalities), based on the changes from before a policy shock (the introduction of HSR stations) and after the shock. Under the DID framework, the parallel trend assumption must be held. In other words, without the consideration of the HSR proximity effect, "time effect," which captures other effects that occur between before and after consideration period, in the treatment group must be the same as that in the control group. In order to achieve more accurate estimation, additional control variables should be introduced to ensure the parallel trend assumption. In contrary, "Treatment effect" captures the possible difference between the treatment group and the control group. Finally, "DID effect" captures the possible difference, which deviates the outcome from the parallel trend assumption. It is assumed that the concerned policy (introduction of HSR stations, in our case) which implemented during this period creates the DID effect.

Although the DID framework expresses the concept under the two time-period analysis, it is also possible to extend it to DID with multiple years. In this case, some studies introduced year control variables instead of the time effect or even other studies introduced both. However, the estimation with multiple years under the DID framework could suffer from an autocorrelation problem which leads to overestimation of the significant level *(28)*. Thus, it is suggested that if the sample size were large enough, the analysis in two time-period, before and after, would yield a more consistent result with some cost of a reduction of insignificance *(28)*. Therefore, our study will strictly follow the analysis in a two-time-frame approach.

To apply the DID method, a general model of our analysis is formulated as:

$$Y_{it} = \beta_0 + \beta_1 c_i + \beta_2 t_t + \beta_3 DID_{it} + \boldsymbol{\theta\delta_{it}} + \varepsilon_{it} \quad (1)$$

where

Y_{it} = Dependent variable in region *i* at year *t*;
c_i = Treatment effect. 0 if *i* belongs to the control group and 1 if *i* belongs to the treatment group;
t_t = Time effect. 0 if *t* is before HSR introduction and 1 if *t* is after HSR introduction;
DID_{it} = DID effect, which equals to $c_i \times t_t$;
$\boldsymbol{\delta_{it}}$ = A vector of control variables in region *i* at year *t*;
$\beta, \boldsymbol{\theta}$ = Unknown coefficients; and
ε_{it} = Error component.

The model is estimated with the Ordinary Least Square (OLS) method. When the coefficient of the treatment effect (β_1) is estimated positive, the treatment group has a higher effect than the control group even before the HSR introduction. This could represent a selection bias, which arises from the fact the region with higher economic performance could be selected as a member of the treatment group. When the coefficient of the time effect (β_2) is estimated positive, the trend of economic growth is positive under the parallel trend assumption. Finally, when the coefficient of the DID effect (β_3) is estimated positive, we may conclude that the introduction of HSR stations positively affect the economic performance.

Propensity Score Matching

PSM is another econometric technique that estimates the policy effect from the difference between a treatment group. However, the time before and after implementation, and the assumption of the parallel trend is not considered. Instead, PSM estimates the score as the probability of being a treatment group from the rest of the explanatory variables based on the logit model. The probability function to estimate propensity score can be expressed as:

$$f(X_i) = P(T_i = 1|X_i) \quad (2)$$

where

$f(\cdot)$ = A propensity score function;
T_i = Treatment effect. 0 if zone *i* belongs to the control group and 1 if *i* belongs to the treatment group; and
X_i = A vector of control variables in zone *i*;

Once the score is estimated, PSM matches the samples between treatment group and control group with the similar score. In practice, several methods and criteria have been applied to the matching process but in this study, we selected the nearest matching method with a 1-to-1 matching restriction which is the most popular method used in PSM. In addition, we do not set any matching caliper, as our treated samples are quite limited so we want to match the treated samples as much as possible.

In the PSM analysis section, we provide two methods to estimate the HSR proximity effect on the economy. First, we follow the typical PSM approach by estimate the Average Treatment Effect on the Treated (ATT) which can be calculated as the difference of the concerning response variable between the matched pair. This ATT could be treated as an equivalent to DID effect if the time effect is already absorbed by other control variables. Second, we consider the possibility of the endogeneity between HSR effect and other control variable. In another word, the possibility that HSR could affect not only the economy but also the other control variable such as input levels and industrial structure as well. Therefore, we use only the matched data to re-estimate DID effect following Eq.(1). Later, we will call this method "matched DID."

Table 1 compiles the literature related to HSR that analyzed its impacts using the DID and PSM method. As for DID, most of the studies revealed a statistical significance in the DID parameter which implies that HSR creates positive economic impact in the region with new HSR service. However, it should be noted that all studies with DID analysis listed in Table 1 used a panel data with multiple years of data to analyze the DID effect. In contrary, in order to provide the clear distinction between before and after the implementation of HSR, this study considers the economic condition only before the introduction of HSR stations and after that. Similarly, PSM analysis from other studies usually considers the treatment effect within the single-time frame. However, two-time frame data is also applied to PSM in order to make a clear comparison between the same dataset in this study. A two-time frame data is also applied to PSM in order to make a clear comparison between the same dataset whereas PSM from other studies usually considers the treatment effect within one-time frame.

Table 1. Examples of studies related to HSR effect analyzed under DID and PSM framework

Study	Study Area	Time	Dependent Variable(s)	Independent Variables	DID / PSM significance
DID					
Hernandez & Jimenez *(21)*	Spain	2001-2010	Revenues/capita, Public debt/capita, Fiscal gap/capita,Yearly property tax	treatment effect + time effect + DID + other control variables (population, unemployment, pop.density, area, time-control, etc.)	Significance (conditionally)
Yoshino & Abidhadjaev *(22)*	Uzbekistan	2005-2012	GDP growth rate	DID + year control + pref. control + other control variables (government expenditure/private investment on each sector)	Significance
Yoshino & Abidhadjaev *(19)*	Japan	1982-2013	Tax Revenue	DID + year control + pref. control + no. of tax payer	Significance
Li & Xu *(20)*	Japan	1980-2003	Population	DID + year control + municipality control + highway investment	Significance (negative)
Jia et al. *(23)*	China	2000-2013	GDP GDP/capita	treatment effect + time effect + DID + other control variables (HSR frequency, investment, local expenditure, industrial structure, road density, etc.)	No significance
PSM					
Komikado and Kato (*27*)	Japan	1981-2006	Patent applications	GDP, GDP per capita, Firm density, Employment density, Number of University, etc.	Significance is found with a small caliper

DATA

In order to examine the effect of the proximity of HSR stations on economic production and labor productivity in Japan, we divide the analysis into two parts by utilizing the dataset measured in two different geographical levels: prefecture level and municipality level. All of the data used in this study are acquired from e-Stat provided by The Statistics Bureau of Japan. The prefecture-level analysis uses the dataset at the prefecture level in Japan, the first administrative level whose population size is equivalent to around the half of that in states of United States. It assumes the situation in 1981 as the year before the shock and that in 2006 as the year after the shock. The municipality-level analysis uses the dataset in the municipality level in Japan, the second administrative level similar to a county in the U.S. In the municipality-level analysis, we exploit the rich dataset in later years to investigate effect from the new HSR stations in the smaller scope between 2010 and 2015.

One of the most important issues in the DID analysis is to define the treatment and control groups. In our context, the treatment group should consist of prefectures/municipalities in which HSR stations had been newly introduced during our time frame. Then, how should we define the control group? For example, the control group could include the regions that have already had HSR services in the before year together with other regions that have no HSR services at both before and after years, although the megacities were discarded from the analysis *(19, 21)*. As we are aware of the long-term effect of HSR, we have decided to divide the control group into two subgroups. The control group 1 consists of the regions where no HSR service was available at both before and after years, while the control group 2 consists of the regions where HSR service was available at both before and after years.

In the prefecture-level analysis, we estimate two DID models, two PSM models and two matched DID models using the control groups 1 and 2 respectively. Certainly, the control group 1 should be more highlighted since the objective of our study is to clarify the economic effect of the proximity of HSR stations when its service becomes available. The model estimation with the control group 2 aims to ensure whether or not the introduction of HSR stations has caused a single-time boost in the regional economies right after HSR station service becomes available since the control group 2 is also assumed not to be affected by new HSR stations. Furthermore, the test with control group 2 also intends to examine a potential correlation between site selection of HSR stations and size of the regional economy because regions with higher economic activities tend to be selected as the sites of HSR stations before others. The prefecture-level analysis covers the period from 1981-2006, which could be the best period to investigate the effect of HSR station investment in the mid-to-late HSR extension plan in Japan. These HSR investments were intended to promote the regional economy rather relieve the traffic congestion, which had been the main objective for early HSR development in Japan. Thus, in the prefecture-level analysis, it is assumed that the treatment group contains the prefectures where the HSR services were available during 1981 and 2006; the control group 1 contains the prefectures where no HSR station was available before 2006; and the control group 2 contains the prefectures where the HSR services had been already available before 1981.

In the municipality-level analysis, we estimate regression models with the treatment group and the control group 1 only, using the municipality-level dataset; but we shift our scope to the period of 2010-2015. The municipality-level analysis focuses on the municipalities in seven prefectures where the latest extensions of the Kyushu Shinkansen in 2011 and Hokuriku Shinkansen in 2015 were implemented. Other municipalities outside the study area are excluded from the analysis. The treatment group is assumed to contain the municipalities within 5 km buffer from new HSR stations *(21)* along the HSR lines extended during the period from 2010-2015, and the control group 1 assumes other municipalities beyond 5 km buffer from HSR station.

EMPIRICAL ANALYSES

The prefecture-level analysis investigates the economic impact of the proximities of HSR stations in regional scale from 1981-2006. The network extensions during this period are intended to promote regional development in the Eastern and the Western parts of Japan. From the general model in Eq. (1), we restructure the to-be estimated DID model as a form of production function as follows:

$$Y_{it} = \beta_0 + \beta_1 c_i + \beta_2 t_t + \beta_3 DID_{it} + \theta_1 K_{it} + \theta_2 L_{it} + \varepsilon_{it} \quad (3)$$

where

Y_{it} = ln Gross Regional Product (GRP) in prefecture *i* at year *t*;
c_i = Treatment effect. 0 if *i* belongs to the control group and 1 if *i* belongs to the treatment group;
t_t = Time effect. 0 if *t* is 1981 and 1 if *t* is 2006;
DID_{it} = DID effect, which equals to $c_i \times t_t$;
K_{it} = ln Capital input in prefecture *i* at year *t*; and
L_{it} = ln Labor input in prefecture *i* at year *t.*

The municipality-level analysis investigates the economic impact of two new HSR lines and stations in municipality scale from 2010-2015. As richer socio-economic data are available in municipality level during this period, it enables us to control the economic growth so that the effect from DID components could be estimated as close to the parallel trend assumption as possible. This also benefits PSM as well since better matching could be achieved from more control variable. It should be noted that, as the result of urbanization and depopulation in Japan, there are many mergers among neighboring municipalities, thus we aggregate data of those merged municipalities based on the municipality list after 2014[2].

Yet, another question arises: "what is the effective range of HSR to the local economy?" Bernard et al. *(15)* utilized the firms' data in Kyushu island, Japan and suggested 30 km could be the effective range of HSR service to firms' performance. However, an area with a radius of 30 km roughly is, in fact, nearly equal to that of the prefecture in Japan, which we have already investigated in the prefecture-level analysis. Then, we assume that the treatment group contains the municipalities within 5 km buffer from new HSR stations, following Hernandez & Jimenez *(21)*. In the municipality-level analysis, we use tax revenue as the dependent variable instead of GRP since municipality-level GRP is not available across the nation. This analysis further analyzes the labor productivity in addition to the production, using tax revenue per capita. The dependent variables are the tax revenue and the tax revenue per taxpayer. The potential control variables are selected, following Hernandez & Jimenez *(21)* and Jia et al. *(23)*; then the model is specified as follows:

$$Y_{it} = \beta_0 + \beta_1 c_i + \beta_2 t_t + \beta_3 DID_{it} + \theta_1 P_{it} + \theta_2 A_{it} + \theta_3 s_{it} + \theta_4 w_{it} + \theta_5 d_{it} + \theta_6 ISE_{it} + \theta_7 ITE_{it} + \varepsilon_{it} \quad (4)$$

where
Y_{it} = ln tax revenue, or ln tax revenue per number of taxpayer in municipality *i* at year *t*;
c_i = Treatment effect. 0 if *i* belongs to the control group and 1 if *i* belongs to the treatment group;
t_t = Time effect. 0 if *t* is 2010 and 1 if *t* is 2015;
DID_{it} = DID effect, which equals to $c_i \times t_t$;
P_{it} = ln number of taxpayer in municipality *i* at year *t*;
A_{it} = ln area of municipality *i* at year *t*;
s_{it} = male/female ratio of municipality *i* at year *t*;

[2] No municipality merging was conducted after 2014

w_{it} = working[3] population/total population ratio of municipality *i* at year *t*;
d_{it} = daytime population/nighttime population ratio of municipality *i* at year *t*;
ISE_{it} = share of workers in secondary industry in municipality *i* at year *t*; and
ITE_{it} = share of workers in tertiary industry in municipality *i* at year *t*.

Table 2 summarizes the estimation results of the DID model in the prefecture-level analysis. This shows that the estimates of the time effect are significantly positive in both Models 1-1a and 1-2a, while the estimate of the treatment effect is positive in Model 1-1a but that is negative in Model 1-2a. It also shows that the estimate of the treatment effect is significantly positive in Model 1-1b but insignificant in Model 1-2b and that both Models 1-1b and 1-2b have the significantly negative time effect. First, the results that the treatment effect is positive for the control group 1 but negative or insignificant for the control group 2 suggest a selection bias that the regions with higher GRP tend to have HSR service earlier. Next, the results that the time effect is positive without controlling capital and labor inputs but negative with controlling them suggest that the capital and labor inputs have grown significantly during the period but the overall technology or productivity has declined. This reduction in technological productivity may reflect the reduction of productivity of Japanese service firms during this period due to the lack of investment and increase of the share of part-time workers *(29)*. Finally, the results of all models unveil that our concerned parameter, DID, does not reach an acceptable significance level.

As for the municipality-level analysis, first, when the control variables are not introduced, the treatment effect is found significantly positive on tax revenue as shown in Models 2-1k and 2-1h, but not on tax revenue per taxpayer as shown in Models 2-3k and 2-3h. However, when the control variables are introduced, the treatment effects are insignificant in both HSR lines as shown in Models 2-2k, 2-2h, 2-4k, and 2-4h. They mean that the control variables can well absorb the difference between the treatment group and the control group, suggesting that the municipalities where the HSR lines were introduced had not been biasedly selected in these extension cases. Next, the results show that the time effect is consistently insignificant when the control variables are not introduced, which obviously reflects the economic stagnation during 2010-2015 especially in the non-core areas of Japan such as Kyushu and Hokuriku regions. However, as the shrinkage of the local economy has been controlled in Model 2-2 and Model 2-4, the time effect becomes significant. Finally, the results show that again, DID effect is not significant in any model.

[3] Working population is defined as population in the age between 15-64 years old

Table 2. Estimation results from DID model

Prefecture Level

	CONTROL GROUP 1						CONTROL GROUP 2					
	Model 1-1a			Model 1-1b			Model 1-2a			Model 1-2b		
	Estimate	t-stat	sig	Estimate	t-stat	sig	Estimate	t-stat	sig	Estimate	t-stat	sig
CONST.	28.678	260.205	***	0.077	0.249		29.965	159.065	***	0.216	0.313	
C	0.483	2.493	*	0.045	2.977	**	-0.803	-2.887	*	0.02	0.778	
T	0.534	3.426	**	-0.159	-5.947	***	0.558	2.095	**	-0.107	-2.178	*
DID	0.049	0.178		-0.008	-0.393		0.025	0.063		-0.016	-0.478	
K				0.365	8.79	***				0.282	4.003	***
L				0.638	16.486	***				0.721	13.257	***
ADJ. R^2	0.303			0.996			0.315			0.995		
N	68						48					

Municipality Level

Y	**TAX REVENUE**											
	Kyushu						*Hokuriku*					
	Model 2-1k			Model 2-2k			Model 2-1h			Model 2-2h		
	Estimate	t-stat	sig	Estimate	t-stat	sig	Estimate	t-stat	sig	Estimate	t-stat	sig
CONST.	16.940	117.253	***	6.661	63.199	***	16.612	84.504	***	7.283	68.509	***
C	1.297	3.138	**	0.011	0.624		1.613	2.595	*	0.021	1.163	
T	0.034	0.166		0.053	4.494	***	0.002	0.008		0.028	2.870	**
DID	0.001	0.001		0.000	0.016		0.008	0.009		-0.006	-0.265	
P				1.009	128.987	***				1.034	194.161	***
A				-0.007	-1.043					-0.035	-6.442	***
S				-0.001	-0.759					0.001	1.115	
W				0.017	7.472	***				0.002	1.151	
D				0.001	2.868	**				0.001	1.887	.
ISE				-0.002	-2.072	*				0.002	2.761	**
ITE				0.002	2.527	*				0.001	1.476	
ADJ. R^2	0.086			0.999			0.070			0.999		
N	180						140					
Y	**TAX REVENUE PER PAYER**											
	Kyushu						*Hokuriku*					
	Model 2-3k			Model 2-4k			Model 2-3h			Model 2-4h		
	Estimate	t-stat	sig	Estimate	t-stat	sig	Estimate	t-stat	sig	Estimate	t-stat	sig
CONST.	7.847	629.353	***	6.643	63.710	***	7.846	809.916	***	7.150	60.118	***
C	0.043	1.205		0.015	0.926		0.055	1.788	.	0.029	1.436	
T	0.006	0.322		0.059	5.545	***	0.023	1.653		0.052	5.105	***
DID	0.005	0.105		0.000	-0.006		-0.005	-0.111		-0.006	-0.208	
A				-0.001	-0.292					-0.014	-2.859	**
S				-0.002	-1.285					-0.001	-0.534	
W				0.019	11.680	***				0.009	7.844	***
D				0.001	2.721	**				0.002	2.702	**
ISE				-0.001	-1.924	.				0.002	3.661	***
ITE				0.002	3.119	**				0.001	1.868	.
ADJ. R^2	0.002			0.796			0.040			0.613		
N	180						140					

Note: "***": $p < 0.001$; "**": $p < 0.01$; "*": $p < 0.05$; "." $p < 0.1$.

Next, we present the analysis from the PSM model. The prefecture-level analysis in control group 1 is based on Model 1-1b and control group 2 is based on Model 1-2b. We also assume that the time effect is already been absorbed by capital and labor inputs. From the general model in Eq. (2), we restructure the to-be estimated PSM model as follows:

$$f(X_{it}) = P(T_{it} = 1 | K_{it}, L_{it}) \quad (5)$$

As for the municipality-level, the analysis for tax revenue is based on Model 2-2k for Kyushu region and Model 2-2h for Hokuriku region, while the analysis for tax revenue per capita is based on Model 2-4k for Kyushu region and Model 2-4h for Hokuriku region. We also assume that the time effect is already been absorbed by control variables. Following Eq. (5), we restructure the to-be estimated PSM model as:

$$f(X_{it}) = P(T_{it} = 1 | P_{it}, A_{it}, s_{it}, w_{it}, d_{it}, ISE_{it}, ITE_{it}) \quad (6)$$

Table 3 summarizes the estimation results of ATT estimation from the PSM model in the prefecture-level analysis. The result of the PSM model shows a positive significant result at 95% confidence interval in control group 1. It suggests that there is an additional premium to GRP once new HSR and its stations have been established in the region. However, the result suggests no difference between the treated region and the region where HSR has already been established (control group 2). Originally, in the data section, we assumed that there is no HSR effect control group 2. However, according to the insignificant estimate from control group 2 analysis, this assumption could be wrong. There might be a long-term HSR effect in control group 2 so the result shows no difference of ATT between treatment group and control group 2. In contrary, all of the results from the municipality-level analysis show the insignificant result. This could further confirm the result from DID analysis that there is no significant difference between cities with new HSR service and those without in tax revenue and tax revenue per capita.

Table 3. Estimation results of ATT from PSM model

Prefecture Level

	ESTIMATE	T-STAT	SIG
CONTROL GROUP 1	0.210	1.969	*
CONTROL GROUP 2	-0.381	-1.578	

Municipality Level

	ESTIMATE	T-STAT	SIG
	Tax revenue		
KYUSHU	0.553	0.880	
HOKURIKU	0.372	0.355	
	Tax revenue per payer		
KYUSHU	0.046	0.901	
HOKURIKU	0.010	0.208	

Note: "***": $p < 0.001$; "**": $p < 0.01$; "*": $p < 0.05$.

Next, we present the results from the matched DID model in Table 4. As we utilize the matched data, similarly to the PSM model, the prefecture-level analysis in control group 1 is based on Model 1-1b and control group 2 is based on Model 1-2b. Similarly, in the municipality-level, the analysis for tax revenue is based on Model 2-2k for Kyushu region and Model 2-2h for Hokuriku region, while the analysis for tax revenue per capita is based on Model 2-4k for Kyushu region and Model 2-4h for Hokuriku region. By considering only the matched data, we expected the significant improvement from matched DID. However, the results shown in Table 4 are quite identical to the result shown in Table 2. This similarity could be explained by several reasons, such as the similarity between matched and unmatched data, or the limited sample size used in this analysis. Consequently, we still could not find the significant effect from new HSR lines/stations at both levels.

Table 4. Estimation results from matched DID model

Prefecture Level

	CONTROL GROUP 1			**CONTROL GROUP 2**		
	Estimate	t-stat	sig	Estimate	t-stat	sig
CONST.	0.421	1.107		0.561	0.881	
C	0.043	2.502		0.012	0.525	
T	-0.150	-4.655	***	-0.115	-2.622	
DID	-0.013	-0.509		-0.013	-0.420	
K	0.369	7.594	***	0.305	4.845	***
L	0.621	13.468	***	0.685	13.496	***
ADJ. R^2	0.996			0.994		

Municipality Level

	TAX REVENUE						**TAX REVENUE PER PAYER**					
	Kyushu			*Hokuriku*			*Kyushu*			*Hokuriku*		
	Estimate	t-stat	sig	Estimate	t-stat	sig	Estimate	t-stat	sig	Estimate	t-stat	sig
CONST.	7.018	15.579	***	6.569	27.535	***	6.854	29.088	***	5.941	12.911	***
C	-0.001	-0.072		0.007	0.457		0.035	2.792	**	0.004	0.225	
T	0.070	2.461		0.068	3.365	**	0.061	3.618	**	0.073	2.582	*
DID	-0.004	-0.152		-0.014	-0.632		-0.007	-0.384		0.002	0.067	
P	1.019	61.452	***	1.036	67.395	***						
A	-0.007	-0.590		-0.021	-1.734		0.003	0.476		0.005	0.363	
S	-0.004	-0.746	***	0.003	0.827		-0.004	-1.191		0.004	0.800	
W	0.014	3.269	**	0.015	4.103	***	0.017	6.252	***	0.019	5.278	***
D	0.001	0.749		0.000	-0.173		0.002	2.550	*	0.004	1.884	
ISE	0.001	0.748		-0.003	-1.913		-0.001	-0.695		0.000	-0.226	
ITE	0.002	0.791		-0.001	-0.666		0.002	2.159	*	0.001	0.352	
ADJ. R^2	1.000			0.999			0.933			0.890		

Note: "***": $p < 0.001$; "**": $p < 0.01$; "*": $p < 0.05$.

HSR Effective Range

Results from both prefecture- and municipality-level suggested the insignificance of the regional/local economic effect from the proximity to HSR stations. We considered the potential reason behind these results and found out that one of the common assumptions in this study is the effective range of a 5km radius from HSR station. This assumption could be potentially wrong; one of the possible reasons to explain the insignificant result likely comes from the spillover effects from the treatment group to control group. Not only the region with HSR service, but benefit could also extend to the adjacent regions if people in those regions are able to utilize the HSR service. Especially in Japan where the feeder systems such as local trains and buses are well-established, users can transfer from the feeder system to the HSR network seamlessly. Therefore, it might be difficult to distinguish the economic effect from the proximity of HSR station as a treatment and control group under DID framework as HSR effect may be observed in the control group as well.

Table 5 shows the DID estimates from municipality-level analysis with the assumption of different treatment groups ranging from 5 km radius to 30 km radius. Most of the estimates shown in Table 5 do not give a significant result. However, in the estimation of tax revenue per payer, we found a significant result at 30 km radius in Kyushu region and at 10 km radius in Hokuriku region. These results could support our assumption above that the effect of HSR service could extend to the adjacent regions as well if the accessibility to HSR station is well established. They also obviously reflect the fact that HSR in Kyushu region mainly runs through the plain area so the effective range of HSR service could be higher up to 30 km. While in Hokuriku region, the accessibility to HSR station is limited by the wide range of mountains so the effective range of HSR service could be limited at 10 km. Not only the geographical factor, but the feeder system to HSR stations may also play an important role to boost the spillover effect of HSR service although more analysis is needed to clarify the factors affecting the effective range of HSR.

Table 5. Estimation results of DID effect with a different radius

Y	**TAX REVENUE**					
RADIUS (KM)	*Kyushu (Model 3-2k)*			*Hokuriku (Model 3-2h)*		
	Estimate	t-stat	sig	Estimate	t-stat	sig
5	0.000	0.016		-0.006	-0.265	
10	0.007	0.620		0.019	1.458	
20	0.009	0.793		-0.016	-1.425	
30	0.024	1.829	.	-0.016	-1.381	
Y	**TAX REVENUE PER PAYER**					
RADIUS (KM)	*Kyushu (Model 3-4k)*			*Hokuriku (Model 3-4h)*		
	Estimate	t-stat	sig	Estimate	t-stat	sig
5	0.000	-0.006		-0.006	-0.208	
10	0.010	0.896		0.033	2.206	*
20	0.013	1.096		-0.014	-1.045	
30	0.025	1.998	*	-0.018	-1.312	

Note: "***": p < 0.001; "**": p < 0.01; "*": p < 0.05; "." p < 0.1.

CONCLUSION

This study investigated empirically the effect from the proximity to HSR stations to the regional/local economies, using the case of Japan's HSR development during 1981-2006 and 2010-2015. The results of empirical analyses from Japanese HSR are summarized as:

(1) In prefecture level, the treatment effect on production (GRP) is significantly positive but the DID effect is insignificant;
(2) In prefecture level, a significant difference between treatment group and control group 1 (No HSR) is suggested from ATT estimation;
(3) In municipality level, the treatment effect on production (tax revenue) is significantly positive but the DID effect is insignificant, and both the treatment effect and the DID effect on productivity (tax revenue per taxpayer) are insignificant.
(4) In municipality level, ATT estimations are not significant for both estimation of production and productivity
(5) Depends on the region, DID effect is found positively significance in various effective ranges.

The insignificant DID effect and ATT estimate can be interpreted that, after the new introduction of HSR, on average, there is no significant economic boost in the regions where the HSR stations were introduced, compared to the economic performance in the regions where HSR stations were not introduced and compared to those where HSR stations have been already in service. Furthermore, this concurs with the results shown by Jia et al. *(23)*, which reported an insignificant DID effect from the case study of HSR network in China. This result is also supported by the insignificant result from PSM model analysis as well. However, in the effective range analysis, significant results are found at a different range in each region. The significant result at 30 km radius found in Kyushu region is also supported by the suggestion from Bernard et al. *(15)*. Therefore, based on the result from this study, it can be concluded that there is a significant effect from the proximity to HSR stations to the local economy, but the proximity range could vary upon the characteristics of such local economy.

Analysis under DID or PSM framework requires the clear distinction between the treatment group and the control group. In our case, we distinguish the treatment group from the control group by the proximity of new HSR stations under the implicit assumption that the local economy in one group is different from that in the other group. However, it should be noted that our analysis results could not confirm whether the introduction of HSR stations leads to regional economic development or not. For instance, if the effective range of HSR service really differs across locations, then further study with deeper consideration of the accessibility, geographical condition, and spatial interaction between economies should be further elaborated. Additionally, the discussion regarding the causal relationship of indirect effects, such as the agglomeration effect from HSR *(13)*, would be also valuable for understanding the effect of HSR. Furthermore, it is also interesting to discuss the effect of HSR through HSR performance such as HSR usage *(16)* or service frequency *(23)* for a deeper understanding of the mechanism between HSR service and economic development.

REFERENCES

1. Yin, M., L. Bertolini, and J. Duan. The effects of the high-speed railway on urban development: International experience and potential implications for China. *Progress in Planning*, Vol. 98, No. 2, 2015, pp. 1–52.

2. Albalate, D., and G. Bel. High-Speed Rail: Lessons for Policy Makers from Experiences Abroad. *Public Administration Review*, Vol. 72, No. 3, 2012, pp. 336–349.

3. De Rus, G. *Economic Analysis of High Speed Rail in Europe*. Fundación BBVA, Bilbao, 2012.

4. Cao, J., X. Cathy, Y. Wang, and Q. Li. Accessibility impacts of China's high-speed rail network. *Journal of Transport Geography*, Vol. 28, 2013, pp. 12–21.

5. Shen, Y., J. De Abreu, and L. M. Martínez. HSR Station Location Choice and Its Local Land Use Impacts on Small Cities: A Case Study of Aveiro, Portugal. *Procedia - Social and Behavioral Sciences*, Vol. 111, 2014, pp. 470–479.

6. Chen, C. L., and P. Hall. The impacts of high-speed trains on British economic geography: a study of the UK's InterCity 125/225 and its effects. *Journal of Transport Geography*, Vol. 19, No. 4, 2011, pp. 689–704.

7. Chen, C., and P. Hall. The wider spatial-economic impacts of high-speed trains: a comparative case study of Manchester and Lille sub-regions. *Journal of Transport Geography*, Vol. 24, 2012, pp. 89–110.

8. Han, J., Y. Hayashi, P. Jia, and Q. Yuan. Economic Effect of High-Speed Rail: Empirical Analysis of Shinkansen's Impact on Industrial Location. *Journal of Transportation Engineering*, Vol. 138, No. 12, 2012, pp. 1551–1557.

9. Garmendia, M., J. M. Ureña, and J. M. Coronado. Long-distance trips in a sparsely populated region: The impact of high-speed infrastructures. *Journal of Transport Geography*, Vol. 19, No. 4, 2011, pp. 537–551.

10. Kim, H. W., D. H. Lee, and H. S. Park. The impact of Gyeongbu High Speed Rail construction on regional economic growth. *KSCE Journal of Civil Engineering*, Vol. 17, No. 6, 2013, pp. 1206–1212.

11. Bouf, D., and C. Desmaris. *Spatial equity and high speed trains : the example of France*. Publication halshs-01137902v2. 2015.

12. Chen, G., and J. D. A. e Silva. Estimating the provincial economic impacts of high-speed rail in Spain: An application of structural equation modeling. *Procedia - Social and Behavioral Sciences*, Vol. 111, 2014, pp. 157–165.

13. Graham, D. J., and P. C. Melo. Assessment of Wider Economic Impacts of High-Speed Rail for Great Britain. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2261, 2012, pp. 15–24.

14. Sasaki, K., T. Ohashi, and A. Ando. High-speed rail transit impact on regional systems: does the Shinkansen contribute to dispersion? *The Annals of Regional Science*, Vol. 31, 1997, pp. 77–98.

15. Bernard, A. B., A. Moxnes, and Y. U. Saito. Production Networks, Geography and Firm Performance. *CEP Discussion Paper*, No. 1435, 2016.

16. Wetwitoo, J., and H. Kato. High-speed rail and regional economic productivity through agglomeration and network externality: A case study of inter-regional transportation in Japan. *Case Studies on Transport Policy*, Vol. 5, No. 4, 2017, pp. 549–559.

17. Pelechrinis, K., C. Zacharias, M. Kokkodis, and T. Lappas. Economic impact and policy implications from urban shared transportation: The case of Pittsburgh's shared bike system. *PLoS ONE*, Vol. 12, No. 8, 2017, pp. 1–21.

18. Kim, J. Y., and J. H. Han. Straw effects of new highway construction on local population and employment growth. *Habitat International*, Vol. 53, 2016, pp. 123–132.

19. Yoshino, N., and U. Abidhadjaev. Impact of Infrastructure Investment on Tax: Estimating Spillover Effects of the Kyushu High-Speed Rail Line in Japan on Regional Tax Revenue. *ADBI Working Paper 574*, 2016.

20. Li, Z., and H. Xu. *High-Speed Railroad and Economic Geography: Evidence From Japan*. Publication 485. Manila, 2016.

21. Hernández, A., and J. L. Jiménez. Does high-speed rail generate spillovers on local budgets? *Transport Policy*, Vol. 35, 2014, pp. 211–219.

22. Yoshino, N., and U. Abidhadjaev. An Impact Evaluation of Investment in Infrastructure: The Case of the Railway Connection in Uzbekistan. *ADBI Working Paper 548*, 2015.

23. Jia, S., C. Zhou, and C. Qin. No difference in effect of high-speed rail on regional economic growth based on match effect perspective? *Transportation Research Part A: Policy and Practice*, Vol. 106, 2017, pp. 144–157.

24. Funderburg, R. G., H. Nixon, M. G. Boarnet, and G. Ferguson. New highways and land use change: Results from a quasi-experimental research design. *Transportation Research Part A*, Vol. 44, 2010, pp. 76–98.

25. Li, H., D. J. Graham, and A. Majumdar. The impacts of speed cameras on road accidents: An application of propensity score matching methods. *Accident Analysis and Prevention*, Vol. 60, 2013, pp. 148–157.

26. Cao, X. J., and J. Schoner. The influence of light rail transit on transit use : An exploration of station area residents along the Hiawatha line in Minneapolis. *Transportation Research Part A*, Vol. 59, 2014, pp. 134–143.

27. Komikado, H., and H. Kato. Can high speed rail foster regional innovation? Empirical Analysis in Japan. *Presented at Transportation Research Board 2018 Annual Meeting, Washington, D.C. January, 2018*.

28. Bertrand, M., E. Duflo, and S. Mullainathan. How Much Should We Trust Differences-in-Differences Estimates? *The Quarterly Journal of Economics*, Vol. 119, No. 1, 2004, pp. 249–275.

29. Fukao, K. Service Sector Productivity in Japan: The key to future economic growth. *RIETI Policy Discussion Paper*, 2010.