Effect of intercity electric multiple unit projects on regional economic development: evidence from a natural experiment in Zhongshan city, China

Heyuan Huang [a], Weikun Zhang [a], Meng Chen [a], Yiming He [b,*]

[a] College of Economics & Management, South China Agricultural University, China
[b] National School of Agricultural Institution & Development, South China Agricultural University, China

ARTICLE INFO

ABSTRACT

This paper analyzes the impact of the opening of an Intercity EUM (Electric Multiple Units) line, Guangzhou–Zhuhai Intercity Mass Rapid Transit, on GDP and population. By constructing the difference-in-difference (DID) model for whole 24 counties in Zhongshan of China, we found that with the opining of Intercity EMU, GDP growth and population inflow of the counties with Intercity EMU stations are significantly higher than those of without stations. Based on the Sharp Regression Discontinuity (SRD) Approach, we found that with the opining of Intercity EMU, the ratio of Guangzhou's GDP to that of other regions has obviously expanded, thus the opening of Intercity EMU has obviously widened the disparity of regional economic scale.

1. Introduction

The Chinese mainland launched its first high-speed rail line in December 26, 2009. By the end of 2016, the operating mileage of Chinese mainland high-speed railway reached 23,600 km, accounting for 65% of the world. Guangdong Province, as the most populous province in mainland China, has the highest economic total, and its high-speed railway construction is in the forefront of the country. As of the end of 2016, Guangdong Province high Speed railway density has reached 0.83 km of each hundred square kilometers.

With the popularization of high-speed railway and the improvement of related construction technology, the construction of EMU is also speeding up. There are three EMU lines has been built or completed in GuangDong Province: the Guangzhou-Shenzhen city rail transit, the Guangzhou-Zhuhai city rail transit and the Guangzhou-Foshan-Zhaoqing city rail transit. In addition, there are 2 more lines, the Guangzhou–Qingyuan city rail transit and the Guangzhou-Dongguan-Shenzhen city rail transit, will be opened in 2019.

According to the "Implementation Plan for Structural Reform of Infrastructure Provinces in Guangdong Province" issued on June 28, 2017, Guangdong Province will invest about 18.95 billion dollars from 2017 to 2020 to finish 22 EMU projects of the Pearl River Delta, to start the eastern Guangdong inter-city railway planning and construction. Basically formed EMU city rail transit network that centered in Guangzhou and its the operation mileage of intercity railway about 650 kilometers.

The importance of EMU lines in Guangdong is increasing. Because of its relatively low cost and convenient use, it can promote the flow of people along the route and economic exchanges, which may have a significant impact on regional economic development.

In terms of literature, Bester et al. (1996), Christiaans (2002), Huck et al. (2002), Drezner et al. (2002) show that the location game possesses an infinity of mixed strategy Nash equilibrium. Cournot game in a two-stage non-cooperative with location choice is derived by Dorta--González et al. (2004). Meagher and Zauner (2005), Benassi and Chirco (2008), Bárcena-Ruiz and Casado-Izaga (2008) investigate Hotelling's duopoly game of location. Pires (2009), Barr and Tassier (2010), Carbó Valverde and Fortes Escalona (2010), García et al. (2011), Loertscher and Muehlheusser (2011), Bárcena-Ruiz and Casado-Izaga (2012), and Sun (2012) explore the location game models.

The Guangzhou-Zhuhai inter city rail transit, starts from the Guangzhou, through the Shunde and Zhongshan, arrived in Zhuhai and Jiangmen. It was opened in January 3, 2011, and after 18 mouthes, the number of passengers trafficed by this EMU lines has reached 27 million. greatly promoted the Interregional exchanges of these cities. It is the second EMU line in Guangdong Province, and the first new construction of the opening of the train line since 2005 (Guang-Shen intercity motor vehicle line is based on the old line from the transformation). It is very important to investigate the actual effect of the operation of Guangzhou-Zhuhai intercity rail transit in the past five years, and will help us to

---

understand the influence of other EMU lines in Guangdong and even the whole country.

2. Theory

With the improvement of transport infrastructure, economic output and population migration along the way will be impacted. There are n motor vehicle routes in each area along the way, and the migration of population is more frequent. If these are included in the model, it will cause the nonlinear polynomial problem similar to the traveler (TSP). Whether the population will migrate to a regional development is generally determined by the willingness to pay in a given region. Assuming that the individual is limited in rationality, the individual will not migrate to the settlement of the region if it does not have a enough marginal willingness to pay for the individual of the act. To simplify the model, consider the existence of two regions, the development of these two areas of homogeneity, the two regions (A and B) each have a large leading manufacturers, and the production of products with substitutes, the behavior of individuals in the two regions move the midpoint. For this behavior individual, the choice of area A or area B, there is a marginal willingness to pay $\vartheta(\vartheta \in [0,1])$, its density function satisfies $\int_0^1 \rho(\vartheta) = 1$, and actually get marginal utilityv_i. This marginal utility is given by the product of region A and B, which satisfies the binomial distribution$(0,1)$. The utility function of the individual in the period i is

$$U_i = \alpha Q_i - \gamma Q_i |\vartheta - v_i| - 1/2 {}^* Q_i^2 - P_i Q_i - t |d_{Ai} - d_{Bi}| \tag{1}$$

Among them, the parameter $\alpha > 0$, which is the market size. P_i and $Q_i$represent the price and output of the product respectively. $\gamma > 0$ represent the marginal negative utility factor. $d_A$and d_B represent the individual vehicle traffic time costs for the individual area A and area B, respectively. t indicate the difference between these two time costs compared to the ratio of utility bundles.

If the behavior of the individuals to the area A and area B motor vehicle traffic time costs are different, with the passage of time, the improvement of the transport infrastructure, $|d_{Ai} - d_{Bi}|$converges to the range of $(0,1]$; If the behavior individual to region A obtains the actual marginal utility of 0 (meaning that the product of area A is undifferentiated or very small in contrast to the location of the behavioral individual), then the actual marginal utility of area B is 1 (meaning that the product of area B is different from the location of the behavioral individual). The utility function of the behavior individual to region A and region B can be obtained from Eq. (1):

$$U_{Ai} = \alpha Q_{Ai} - \gamma Q_{Ai} \vartheta - 1/2 {}^* Q_{Ai}^2 - P_{Ai} Q_{Ai} - t Q_{Ai} \tag{2}$$

$$U_{Bi} = \alpha Q_{Bi} - \gamma Q_{Bi}(1 - \vartheta) - 1/2 {}^* Q_{Bi}^2 - P_{Bi} Q_{Bi} - t Q_{Bi} \tag{3}$$

In the model, the price of the product of the area A and B is given, which is observed by the behavior individual. And the yield of the products of the regions A and B is not given, according to the behavior of individual unit effect of the firm's observation, if the individual behavior preference region A is strictly stronger than the region B, then satisfying:

$$\alpha - \gamma \vartheta - P_{Ai} - t |d_{Ai} - d_{Bi}| \geq \alpha - \gamma(1 - \vartheta) - P_{Bi} - t |d_{Ai} - d_{Bi}| \tag{4}$$

$$\vartheta \leq \frac{P_{Bi} - P_{Ai}}{2\gamma} + \frac{1}{2} \tag{5}$$

Similarly, if the behavior of individual preference area B is stronger than area A, then there is:

$$\vartheta \geq \frac{P_{Bi} - P_{Ai}}{2\gamma} + \frac{1}{2} \tag{6}$$

Regions A and B are determined by the unit utility function of the individual behavior, and the revenue function of region A and region B is:

$$\pi_A = P_A \int_0^{\frac{1}{2} + \frac{P_B - P_A}{2\gamma}} (\alpha - \gamma \vartheta - P_A - t) \rho(\vartheta) d\vartheta$$

$$= P_A(\alpha - P_A - t)\left(\frac{1}{2} + \frac{P_B - P_A}{2\gamma}\right) - \frac{\gamma P_A}{2}\left(\frac{1}{2} + \frac{P_B - P_A}{2\gamma}\right)^2 \tag{7}$$

$$\pi_B = P_B \int_{\frac{1}{2} + \frac{P_B - P_A}{2\gamma}}^1 (\alpha - \gamma \vartheta - P_B - t) \rho(\vartheta) d\vartheta$$

$$= P_B(\alpha - \gamma - P_B - t)\left(\frac{1}{2} + \frac{P_B - P_A}{2\gamma}\right) + \frac{\gamma P_B}{2}\left[1 - \left(\frac{1}{2} + \frac{P_B - P_A}{2\gamma}\right)^2\right] \tag{8}$$

With the passage of time and frequent regional transactions, $P_A = P_B = P^*$ and$Q_A = Q_B = Q^*$. Take the two order derivative of P_A of formula (7). Derivative of P_A first order and then $\alpha$and Derivative of P_A first order and then t, get

$$\frac{\partial^2 \pi_A}{\partial(P_A)^2}\bigg|_{P_A = P_B = P^*} = -\frac{2\alpha - 3P^* - 2t + \gamma}{2\gamma} < 0 \tag{9}$$

$$\frac{\partial^2 \pi_A}{\partial P_A \partial \alpha} = \frac{\gamma - P^*}{2\gamma} \tag{10}$$

$$\frac{\partial^2 \pi_A}{\partial P_A \partial t} = -\frac{\gamma - P^*}{2\gamma} \tag{11}$$

if $\gamma > P^*$, it means that $v_B - v_A > P^*/\gamma$, indicating that the economies of the two economies are heterogeneous. By formula (9) we can know that $\pi_A$is strict concave, there are $\partial P^*/\partial \alpha = -(\partial^2 \pi_A/\partial P_A \partial \alpha)/(\partial^2 \pi_A/\partial(P_A)^2) > 0$　$\partial P^*/\partial t < 0$`　$\partial \pi_A/\partial \alpha > 0$ and $\partial \pi_A/\partial t < 0$; if $\gamma < P^*$, it means that $v_B - v_A < P^*/\gamma$, indicating that the economic situation of two regional economies is homogeneous, getting that $\partial \pi_A/\partial \alpha < 0$and $\partial \pi_A/\partial t > 0$.

Proposition 1. If the two regional economies are heterogeneous and the traffic time cost is shortened, it is advantageous to optimize the regional resource allocation, enlarge the development scale of the economy and benefit the economic growth. If the two regional economies are homogeneous and the cost of transit time is shortened, the competition between economies is aggravated, and if the appropriate market size is maintained, resources are not skewed, which will be beneficial to economic growth.

In fact, the regional development is generally unbalanced because of the different resources endowments and the inclination of government policies in China, that is$\gamma > P^*$. Considering the behavior of the individual to the same traffic time as the region A and region B, the utility function of behavior individual to region A and region B can be obtained by formula (1):

$$\widehat{U}_{Ai} = \alpha Q_{Ai} - \gamma Q_{Ai} \vartheta - 1/2 {}^* Q_{Ai}^2 - P_{Ai} Q_{Ai} \tag{12}$$

$$\widehat{U}_{Bi} = \alpha Q_{Bi} - \gamma Q_{Bi}(1 - \vartheta) - 1/2 {}^* Q_{Bi}^2 - P_{Bi} Q_{Bi} \tag{13}$$

The process of deriving the model is exactly the same as described above, and is not repeated here. Because of the different regional development, general $\gamma > \widehat{P}$, can get the revenue function of the regional A manufacturers $\widehat{\pi}_{Ai}$about market scale α's first order condition is greater than $0(\partial \widehat{\pi}_{Ai}/\partial \alpha > 0)$, the region A product price about market scaleα's first order condition is also greater than $0(\partial \widehat{P}_{Ai}/\partial \alpha > 0)$, So we can get the following proposition:

Proposition 2. The opening of the train line, greatly shorten the transportation time, converge to the transportation time cost is consistent, the price of the product rises with the market scale enlargement, the local economy gross also will rise with the increase.

No matter the time cost of transportation is very different or the same,

the balanced output of the region and the profit of the manufacturers are different. Combining formula (2), formula (3), formula (12) and formula (13), a comparison of the two is obtained:

$$Q_{Ai} = Q_{Bi} = Q^* = \frac{1}{2}(\alpha - P^* - t) - \frac{\gamma}{8} \leq \widehat{Q}_{Ai} = \widehat{Q}_{Bi} = \widehat{Q}^* = \frac{1}{2}\left(\alpha - \widehat{P}^*\right) - \frac{\gamma}{8}$$

(14)

$$\pi^* = \widehat{\pi}^* - \frac{t}{2}P^* < \widehat{\pi}^*$$

(15)

Proposition 3. The equilibrium yield of the same area of transportation time cost and the equilibrium income of the manufacturer are more than that of the different transportation time cost.

Proposition 4. At the beginning of the opening of the train line, the regional economy and market scale will be enlarged, and with the regional division of labor and the optimization of the train line, the transportation time cost is shortened, the regional transaction is more and more frequent, and the regional product output and GDP will rise.

3. Methodology

3.1. Econometric methodology

This section of the analysis using the DID(Differences in Differences) model. As the influence of intercity train line is mainly through the docking site, and this effect decreases with increasing distance. Because of its high construction cost, the general planning and construction of the Intercity railway line is coordinated by the national or provincial government, and its route and site setting must take into account economic benefits, regional development balance, line distance, terrain conditions, construction difficulty and other factors. Therefore, whether to set up a station in a certain area is not entirely determined by the level of economic development, and it has certain randomness from the angle of economic development. Therefore, it is possible to view the construction of Guangzhou-Zhuhai intercity rail and its site as "quasi-experiment" of an external impact on economic growth along the route.

The DID model groups samples according to whether they are impacted, and introduces grouping and staging of two dummy variables. By analyzing the grouped dummy variables, the staged dummy variables and the intersection of two dummy variables in the regression model, the actual impact magnitude of the impact on the impacted sample can be obtained after the second order difference, which can solve the problem of internal and external validity during the implementation of "quasi-experiment".

Therefore, this article sets up the area with the stations in the experimental group, in the control group. Study on the impact of setting up a site on the regional GDP and migration trend of the Guangzhou-Zhuhai intercity rail. The overall structure of the DID model is as follows:

$$Y_{ij} = \gamma + \beta_1 G_i + \beta_2 T_j + \beta_3 G_i T_j + \sum_{k=1}^{n} \alpha_k x_k + \varepsilon_{ij}$$

(16)

Since the construction of intercity train line needs to consider economic benefits, this grouping is not completely random from the perspective of economic development and needs to be adjusted by adding other covariates. In addition, the introduction of covariates can reduce the variance of disturbance items and make the estimation more accurate. x_k in the above equation is the covariate. When the sample is in the experimental group, $i = 1$; when the sample is in the tracking period, $j = 1$. G_i is a group dummy variable, and T_j is a staged dummy variable, and the specific values are as follows:

$$G_i = \begin{cases} 1 & if \quad i = 1 \\ 0 & else \end{cases}$$

(17)

$$T_j = \begin{cases} 1 & if \, j = 1 \\ 0 & else \end{cases}$$

(18)

According to the analysis, by a first order difference performed for each group, the variation of the dependent variables in the two period for each group can be obtained, which can eliminate the impact caused by the change of the internal factors such as industrial structure change in the period of two. Then, to eliminate the influence of individual factors among groups by a second order difference, and the actual impact of impact on the impacted samples is obtained. Make $D = (Y_{11} - Y_{10}) - (Y_{01} - Y_{00})$, obviously, there is:

$$D = (Y_{11} - Y_{10}) - (Y_{01} - Y_{00}) = \beta_2 + \beta_3 + \varepsilon_{11} - \varepsilon_{10} - \beta_2 + \varepsilon_{01} - \varepsilon_{00}$$
$$= \beta_3 + \Delta\varepsilon$$

(19)

Among them, $\Delta\varepsilon$ is the result of second order difference of random perturbation term, the expectation is 0. In this way, the coefficient β_3 of the intersection of two dummy variables can be used to evaluate the significance of the policy effect.

The construction of the intercity rail line improves the traffic conditions along the line. On the one hand, it can reduce the cost of interregional division of labor, so that the hold market is enlarged and can provide more jobs for residents along the route. On the other hand, more convenient traffic makes life easier for residents along the routes and eventually improve the region's appeal to residents. These changes will ultimately affect the economic and social development. In order to investigate the direct influence of the construction of Guang-Zhu City rail on the areas along the line, this section will analyze the GDP and the population transfer number respectively.

3.2. Data description

In view of the small scope of the research object, only four municipal and one district administrative units are involved. If the DID model is analyzed directly using municipal data, the sample size is too small to affect the measurement results, so the data of the municipal administrative units is not used in this section.

In addition, the current statistical yearbook in mainland China is generally published only to the relevant data of the next level of administrative units in the statistical units, so the public statistics are generally up to the county (distinct) level. As the only pilot of Guangdong province governing county, Shunde district has a considerable degree of financial autonomy. However, due to its establishment still belongs to Foshan, the specific data below the district level is not disclosed in the yearbook. Zhuhai and Jiangmen as Municipal city, its county (district) level data can be found in the yearbook, but because Zhuhai only three districts, and Jiangmen district only three districts and four counties, its sample size is still less, will affect the accuracy of the model. In addition, Guangzhou as the capital of Guangdong province and South China's transportation hub center, from Guang-Zhu City rail opened till now, there are other intercity rail lines, such as the Guang-Fo subway, which have different impacts on different districts. If you use the data below the Guangzhou city level, it may cause some interference. Therefore, the relevant data of the above three cities and one distinct are not used.

Zhongshan city is one of the only two cities in mainland China to implement the municipal-town two-level administrative system (another is Dongguan city). Unlike other districts that implement the municipal-county-town three-level administrative system, the data of 24 township-level administrative units of Zhongshan city all have detailed records in the Statistical Yearbook. On the one hand, the size of the sample is relatively sufficient; On the other hand, the regional division of the township administrative units is finer, which makes the relevant data of the district administrative units more detailed and deeper, and can reflect the facts more objectively. In addition, the Guangzhou-Zhuhai intercity rail transit runs through the whole territory of Zhongshan
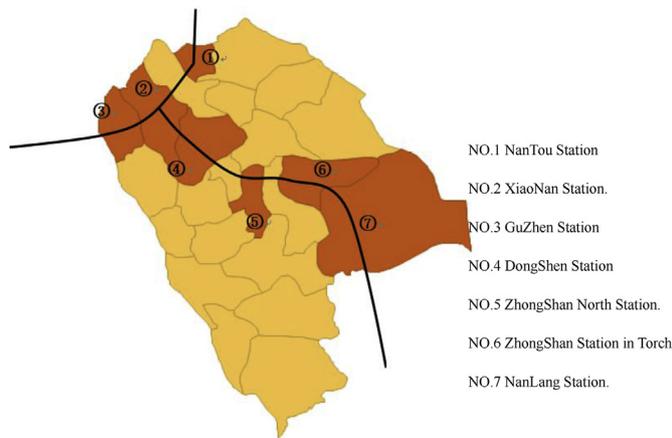
Fig. 1. Map of Intercity EMU project in Zhongshan.

NO.1 NanTou Station
NO.2 XiaoNan Station.
NO.3 GuZhen Station
NO.4 DongShen Station
NO.5 ZhongShan North Station.
NO.6 ZhongShan Station in Torch
NO.7 NanLang Station.

from north to south and has 6 stations in Zhongshan. Its impact on Zhongshan is relatively more significant. Finally, from 2010 to 2015, Zhongshan's GDP has been fairly stable at the second and third place in the four cities and one district in this area which makes Zhongshan city more representativeness.

Therefore, in this section, the data of DID model metering analysis for economic performance selects the relevant data of all 24 town-level administrative units in Zhongshan city. Among them, 6 towns (districts) with Stations of Guangzhou-Zhuhai intercity rail transit will be used as the experimental group and the other 18 towns (districts) will be used as the control group. Since the Guangzhou-Zhuhai intercity rail transit was opened on January 3, 2011, it will be established as a base period in 2010. In addition, in order to lengthen the experimental span as far as possible, this paper selects the last year that the Statistical Yearbook has published data, that is, 2015 as the experimental period. All the data in this section derive from the 2011 and 2016 annual Zhongshan Statistical Yearbook. Details of relevant sites are provided in Fig. 1 and Table 1.

4. Results

4.1. DID experiment

GDP is one of the most important indicators to measure economic development. In view of the fact that GDP is also affected by the size of the population and the size of the investment, this section adds covariance to these two aspects.

In view of the population size, on the one hand, because Guangdong Province is the largest imported labor force in mainland China, permanent population of non-native population accounts for a large proportion. Zhongshan city is located in the Pearl River Delta region, from 2010 to 2015, its registered population in the permanent population accounted for less than 50%. On the other hand, due to the impact of the current household registration system, the household registration population is different from the immigrant population in terms of real estate investment and education, and there are many differences in their labor quality. In view of this, the household population scale, denoted by re, and the resident population, denoted by in, are introduced to measure the size of the population.

For the scale of investment, investment in fixed assets is the most important aspect. first of all, the total investment in fixed assets, denoted by fai, is introduced. In addition, Zhongshan is located in the Pearl River Delta, which is the earliest opening area in mainland China, the impact of foreign investment is also an important factor in this region. The total amount of foreign capital actually utilized, denoted by fiu, is introduced here as another measure of investment size. To simplify the model, we use a linear production function, according to Eq. (1), getting:

$$GDP_{ij} = \gamma + \beta_1 G_i + \beta_2 T_j + \beta_3 G_i T_j + \alpha_1 in_{ij} + \alpha_2 re_{ij} + \alpha_3 fai_{ij} + \alpha_4 fiu_{ij} + \varepsilon_{ij}$$
(20)

Measurement results are shown in Table 2.

Among them, the $\_I\_diff\_1$ is the β_3 of formula 1 and formula 2, and the coefficients of the orthogonal item of the group dummy variable and the staged dummy variable. From the measurement results, the orthogonal item is significantly positive at the level of 5%, P value is 0.047, indicating that the setting of the intercity EMU station could significantly increase the GDP of the site-setting area. In addition, the four covariance

Table 1
Related data of Experimental Group.

Variable	Sample	Average	SD	Minimum	Maximum	Skewness	Kurtosis	Remake
gdp	12	1660031	1112388	671642	4433056	1.387104	4.149126	GDP
first	12	18167.17	22706.71	1128	64162	0.9507608	2.295248	Output of the first industry
second	12	1025929	949514.6	272101	3525964	1.713809	5.028032	Output of the second industry
third	12	615935.3	381490.9	200496	1322560	0.6830774	2.166397	Output of the third industry
inhabitant	12	187458	78931.8	108041	324590	0.5903652	1.98156	Number of resident population
registered	12	95454.83	56421.85	40305	171552	0.5445705	1.507428	Registered population
Fai	12	509226.8	382877.5	114406	1528849	1.691639	5.285004	Fixed-asset investment
Fiu	12	5115.417	7398.166	465	21303	1.67418	4.01109	Foreign capital investment
miip	12	971.25	1679.016	56	5971	2.458449	7.85212	Number of people moving from other parts of the province
miop	12	576.8333	538.8567	134	1615	0.8381531	2.073153	Number of people moving from other provinces
moip	12	398.75	846.0068	28	3051	2.883977	9.596691	Number of people moving to other parts of the province
moop	12	307.25	297.5079	65	955	1.164902	3.042542	Number of people moving to other provinces

Table 2
GDP in the area along the line affected by EMU.

gdp	Coef.	Std. Err.	t	P > t	[95% Conf.Interval]	
_It_1	226532.8	84115.23	2.69	0.010	56529.56	396536
_Ig_1	-823.3254	116540.2	-0.01	0.994	-236359.8	234713.2
_I_diff_1	302011.5	147460.3	2.05	0.047	3983.123	600039.9
in	4.336874	0.9144192	4.74	0.000	2.488764	6.184984
re	0.8077585	1.494011	0.54	0.592	-2.211751	3.827268
fai	1.12199	0.2329263	4.82	0.000	0.6512282	1.592752
fiu	44.28472	13.42499	3.3	0.002	17.15181	71.41764
_cons	-291385.9	80897.07	-3.6	0.001	-454885	-127886.8

Table 3
Population migration along the line is affected by EUM.

Variables	sn	sw	z
t	-197.7	-14.94	-212.6
	(268.4)	(60.89)	(281.0)
g	-212.6	-43.63	-256.3
	(361.3)	(81.96)	(378.2)
_diff	1,136**	-108.5	1,028*
	(489.7)	(111.1)	(512.7)
first	-0.00395	-0.000505	-0.00446
	(0.00607)	(0.00138)	(0.00636)
second	-0.000355	0.000284***	-7.14e-05
	(0.000229)	(5.19e-05)	(0.000240)
third	0.00109***	0.000279***	0.00137***
	(0.000379)	(8.59e-05)	(0.000396)
Constant	82.21	-78.97	3.241
	(263.5)	(59.79)	(275.9)
Observations	48	48	48
R-squared	0.352	0.602	0.414

Note:Standard errors in parentheses ***p < 0.01, **p < 0.05, *p < 0.1.

variables, in addition to the total registered population, were significantly positive in the 1% level, in line with expectations. However, the effect of the total registered population on GDP is not significant, with P value of only 0.592. It is clear that the total registered population has little impact on the overall output of Zhongshan city, a region that has a large influx of immigrants and is not fully reflected in the household register.

The construction of intercity rail lines improves the traffic situation along the area, and the more convenient traffic will increase the attractiveness of the area to the residents, thus affecting the trend of population change to a certain extent. And this change will eventually play a role in the economic and social development. This section will make the measurement analysis of the influence of Guangzhou-Zhuhai inter city rail transit on the household register migration of the site settings.

In view of the impact of economic development, the migration of household registration is joined by output of three major industrial, fixed assets investment and actual utilization of foreign capital as a covariant variables. In addition, social welfare security is an important factor affecting the migration intention, but because the samples used in this section are from Zhongshan, there is no difference in the social welfare policy and the concrete implementation between towns from experimental groups or control groups. Therefore, the covariance of social welfare guarantee is not specially introduced here.

In terms of dependent variables, in order to more objectively reflect the change in the trend of population migration, we separately measured the number migration in the province of Zhongshan in this year, denoted by **sn**, the number migrations outside the province, denoted by **sw**, and the total migration population, denoted by **w**. **sn** is the difference between the number of **miip,** immigrating from the province, and the number of **miop**, emigrating to the province in this year. **sw** is the difference between the number of **moip,** migrating from outside the province, and the **moop,** emigrating to the outside province, in this year. **z** is the sum of **sn** and **sw**.

$$sn_{ij} = \gamma + \beta_1 G_i + \beta_2 T_j + \beta_3 G_i T_j + \alpha_1 first_{ij} + \alpha_2 second_{ij} + \alpha_3 third_{ij} + \varepsilon_{ij}$$
(21)

$$sw_{ij} = \gamma + \beta_1 G_i + \beta_2 T_j + \beta_3 G_i T_j + \alpha_1 first_{ij} + \alpha_2 second_{ij} + \alpha_3 third_{ij} + \varepsilon_{ij}$$
(22)

$$z_{ij} = \gamma + \beta_1 G_i + \beta_2 T_j + \beta_3 G_i T_j + \alpha_1 first_{ij} + \alpha_2 second_{ij} + \alpha_3 third_{ij} + \varepsilon_{ij}$$
(23)

Measurement results are shown in Table 3.

The number migration in the province of Zhongshan in this year is significantly positive at 5% level. That is, the sites of EMU line setting significant increases in the attractiveness of residents from other parts of the same province of that region. The intersection of the total number of

migrants was significantly positive at 10% level, at 1028, that is, the opening of Guangzhou-Zhuhai intercity rail transit allowed the site to set an additional annual increase of about 1028 citizens. On the contrary, the intersection of the number people Immigrate to Zhongshan outside the province is not significant, that is, the influence of Guangzhou-Zhuhai intercity rail transit on the emigration from other provinces to Zhongshan is not significant. In addition, the effect of primary output value on three items is not significant, and the second industry output value only has a significant positive effect on the number of Zhongshan migration in the outside of the province, while the tertiary industry was significant at the level of 1% to the three items.

4.2. SRD experiment

This section uses the SRD(Sharp Regression Discontinuity) model for analysis. The idea of regression discontinuity design is that there is a continuous variable X_i, it determines the processing variable D_i by a breakpoint, and this continuous variable X_i is called the assignment variable. The random distribution of samples in the small neighborhood $[\delta - \mu, \delta + \mu]$ of X_i is regarded as "quasi experiment". By estimating LATE(Local Average Treatment Effect), it is possible to reduce whether the dependent varible has a breakpoint in $X_i = \delta$. Among them,

$$LATE = \lim_{X_i \to \delta+} E(Y_i) - \lim_{X_i \to \delta-} E(Y_i)$$
(24)

In this paper, the time variable is used as X_i. Since its opening at the first quarter of 2011, the impact of the Guangzhou-Zhuhai intercity rail transit line is starting. So, the breakpoint at the individual treatment concept from 0 to 1 is consistent with the assumption of breakpoint regression.

As a local random experiment, the introduction of covariables does not affect consistency of the breakpoint regression estimator. However, introducting covariables that can explain the dependent variable, can effectively would reduce the disturbance variance and improve the accuracy of estimation effectively. Because of the differences of industrial structure, the seasonal shocks are not the same to diffenent cities. In order to reduce the related fluctuations, this section introduces the first production ratio, denoted by **f**, the second production ratio, denoted by **s**, to measure the differences in industrial structure. The ratio of total import and export volume to GDP, denoted by **zp**, was introduced as a covariate to measure the extent of foreign trade dependence. The regression equation is as follows:

$$Y_i = \alpha + \beta X_i + \sigma D_i + \gamma D_i X_i + f_i + s_i + zp_i + \varepsilon_i$$
(25)

Among them,

$$D_i = \begin{cases} 1 & if\ X_i > 2011 \\ 0 & else \end{cases}$$
(26)

The original hypothesis of the test is:

$$H_0 \equiv \lim_{X_i \to \delta+} Y_i - \lim_{X_i \to \delta-} Y_i = 0$$
(27)

Because regression discontinuity design is a local random test, only the observed values near the breakpoint, in principle, should be used. Excessive observation periods may affect the progress of the test. In order to compress the range of observations as much as possible under the premise of guaranteeing the number of samples, this section uses quarterly statistics. In view of the lack of Shunde district eight quarters of data in 2009 and 2010, that is, the lack of all the data before the breakpoint, in order to reduce the error, this section does not use the relevant data of Shunde district. The data used in this section are all from Guangdong Provincial Statistics Network. The data of the four quarters in 2009 are calculated from the actual data and growth rate in 2010. The general characteristics of the data are given in Table 1.

In the four cities and one distinct of the Guangzhou-Zhuhai inter city rail transit crossing, the Guangzhou is the most economically developed

Table 4

Change of GDP ratio of other cities and GuangZhou.

g2011	Coef.	Std. Err.	z	P > z	[95% Conf.Interval]	
lwald	-9.560952	6.935322	-1.38	0.168	-23.15393	4.032029
lwald50	-6.114962	4.848555	-1.26	0.207	-15.61796	3.388031
lwald200	-1.472858	4.576923	-0.32	0.748	-10.44346	7.497746

Table 5

Change of GDP ratio of other cities and GuangZhou with covariate variables.

g2011	Coef.	Std. Err.	z	P > z	[95% Conf.Interval]	
lwald	-10.09926	4.622810	-2.18	0.029	-19.15980	-1.038722
lwald50	-1.492242	3.281893	-0.45	0.649	-7.924534	4.940251
lwald200	-0.1336211	3.868907	-0.03	0.972	-7.716539	7.449297

Table 6

Continuity check results of covariate conditional density at breakpoint.

g2011	Coef.	Std. Err.	z	P > z	[95% Conf.Interval]	
f	-0.01115	0.022551	-0.49	0.621	-0.05535	0.033052
s	0.024052	0.018708	1.29	0.199	-0.01262	0.060720
zp	-0.08718	0.879419	-0.10	0.921	-1.81081	1.636451
lwald	-9.56095	6.935322	-1.38	0.168	-23.1539	4.032029
f50	-0.00673	0.020575	-0.33	0.744	-0.04705	0.033598
s50	-0.00724	0.015162	-0.48	0.633	-0.03696	0.022475
zp50	-0.08718	0.752950	-0.12	0.908	-1.56293	1.388577
lwald50	-6.11496	4.848555	-1.26	0.207	-15.6180	3.388031
f200	-0.00402	0.014295	-0.28	0.779	-0.03204	0.023998
s200	-0.00761	0.016148	-0.47	0.637	-0.03926	0.024035
zp200	-0.08140	0.519849	-0.16	0.876	-1.10029	0.937482
lwald200	-1.47286	4.576923	-0.32	0.748	-10.4435	7.497746

region. From 2010 to 2016, its GDP has remained at 1.63 to 1.84 times of the total GDP of other four regions. The economic development gap between the other regions is relatively small. From 2010 and 2016, GDP of the most economically developed region of these four regions is only 1.72 times as much as GDP of the least economically developed one. Therefore, in order to analyze the changes in the balance of regional development, Guangzhou is set as the central area here, and the other cities are regarded as peripheral areas. The ratio of GDP in other cities to Guangzhou's GDP is taken as the research object.

Due to the lack of quarterly data, and if the use of quarterly data to calculate the proportion of the first, second production and the ratio of total volume of imports and exports to GDP, it may also cause unnecessary fluctuations due to the impact of different industries during the off and peak season. Here we use the annual data to estimate the size of the three covariates.

Because there are some differences in the economic scale of the three cities, in order to observe the trend more intuitively, this section uses the ratio of standardized municipalities to Guangzhou's GDP as the analytic variable. The ratio of each period in the city is divided by the ratio of GDP

in the first quarter of 2011 to that of Guangzhou (i.e., the ratio of GDP of the breakpoint period) to 100, to observe the change of the ratio between the period before and after the breakpoint.

In view of the difference in the economic size of these four cities, in order to directly analyze the trend, in this section, the ratio of GDP in these cities to Guangzhou's GDP would be standardized for the first quarter of 2011, and the results would be used as interpreted variables.

This section uses the triangular kernel for SRD, empirical results without adding covariates show in Table 4, the results of adding covariates show in Table 5. Among them, lwald, lwald50, lwald200 respectively represent the optimal bandwidth, 0.5 times the optimal bandwidth and 2 times the optimal bandwidth.

According to Table 4 and Table 5, no metter the covariates are added or not, the coefficients are all negative under the three bandwidths. After adding the covariate, the variance decreases, among which, the P value of the group variable decreases from 0.168 to 0.029 under the optimal bandwidth, that is, rejects the null hypothesis H_0 at the level of 5%. No matter whether covariates are added or not, the coefficients of the grouping variables under the optimal bandwidth are basically the same, which are -10.09926 and -9.560952 respectively. That is, the impact of which makes the ratio of GDP of the other three cities and Guangzhou reduced to about 90% of the original ratio, the opening of the intercity rail lines obviously more favorable to the developed area of Guangzhou.

In order to prevent the effect of covariance on the estimation of discontinuity in the breakpoint period, this paper examines the continuity of the conditional density of the covariance in the breakpoint period. The results are shown in Tables 6, 7, and 8.

According to the result of Table 6, the P values of concomitant variables f, s and zp at bandwidths of 50%, 100% and 200% are all are greater than 0.199, which means all of the concomitant variables has no the break point in the first quarter of 2011. The break point of GDP ratio of other cities to GuangZhou with covariate variable is not result from the break point of the concomitant variables. The opining of Guangzhou–Zhuhai Intercity Mass Rapid significantly expanded the economic gap between the the area along the line.

Table 7

Related data of Control Group.

Variable	Sample	Average	SD	Minimum	Maximum	Skewness	Kurtosis	Remake
gdp	36	767024.4	399779.3	155803	1874955	0.6657001	3.102528	GDP
first	36	26056.92	17898.47	908	73236	0.6083737	2.617837	Output of the first industry
second	36	367938.6	172008.8	92691	769575	0.4512259	2.496653	Output of the second industry
third	36	373028.4	316906.5	53834	1683521	2.178778	9.203611	Output of the third industry
inhabitant	36	113411.8	49660.85	31410	223460	0.4505494	2.759182	Number of resident population
registered	36	54281.31	24997.33	7804	103608	-0.129985	2.086856	Registered population
Fai	36	302141	175353.5	57600	726140	1.089805	3.365686	Fixed-asset investment
Fiu	36	1590.75	1791.494	0	6502	1.143728	3.259765	Foreign capital investment
miip	36	278.2222	444.2123	29	2623	4.338217	22.99272	Number of people moving from other parts of the province
miop	36	223.9722	226.1468	0	872	1.54867	4.645412	Number of people moving from other provinces
moip	36	122.8889	221.0578	4	1292	4.415258	23.31192	Number of people moving to other parts of the province
moop	36	115.0278	92.60129	0	396	1.530005	4.807416	Number of people moving to other provinces

Table 8
Location of station.

Station Number in Fig. 1	Station Name	Location
1	NanTou Station	NanTou town
2	XiaoNan Station	XiaoNan town
3	GuZhen Station	GuZhen town
4	DongShen Station	DongShen town
5	ZhongShan North	ShiQi town
6	ZhongShan Station	Torch Development zone
7	NanLang Station	NanLang town

5. Conclusion

EMU lines will be a key point for the construction of infrastructure in China for in future a long time. On the one hand, the construction of EMU line needs a large amount of funds in the early stage, and it will involve a series of issues such as land expropriation and noise in the area along the line. On the other hand, it has a significant role in promoting the development of regional economic integration and the construction of urban agglomerations. Therefore, the research on the actual influence of the construction of Intercity train line is of great significance.

Based on the results of the DID model constructed in this paper, we can know that, in terms of longitudinal impact, the opening of intercity rail line is conducive to the migration of people from other areas to the area along the rail line, speeding up the population agglomeration along the area, and promoting the improvement of economic aggregates along the region, and increasing the absolute speed of economic growth along the region. However, through the horizontal comparison of the break-point regression, we find that the opening of the Intercity railway line will further expand the economic development gap between the first and the after areas along the region, making the relative speed of post-development economic growth slowed down, which may further aggravate the unbalanced development of the area along the line in turn.

Declarations

Author contribution statement

Heyuan Huang: Conceived and designed the experiments; Wrote the paper.

Weikun Zhang: Performed the experiments.

Meng Chen: Analyzed and interpreted the data.

Yiming He: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Bárcena-Ruiz, Juan Carlos, Casado-Izaga, F.Javier, 2008. ''Timing of endogenous bargaining over costs and firms' locations. J. Econ./Zeitschrift Fur Nationalokonomie 95 (2), 149–166.

Bárcena-Ruiz, Juan Carlos, Casado-Izaga, F.Javier, 2012. Location of public and private firms under endogenous timing of choices. J. Econ./Zeitschrift Fur Nationalokonomie 105 (2), 129–143.

Barr, Jason, Tassier, Troy, 2010. ''Endogenous neighborhood selection and the attainment of cooperation in a spatial prisoner's dilemma game. Comput. Econ. 35 (3), 211–234.

Benassi, Corrado, Chirco, Alessandra, 2008. An elasticity Approach to equilibrium and preference concentration in the hotelling game. J. Econ./Zeitschrift Fur Nationalokonomie 94 (2), 125–141.

Bester, Helmut, de Palma, André, Leininger, Wolfgang, Thomas, Jonathan, von Thadden, Ernst-Ludwig, 1996. A noncooperative analysis of Hotelling's location game. Games Econ. Behav. 12 (2), 165–186. Retrieved. http://wrap.warwick.ac.uk/19070/.

Carbó Valverde, S., Escalona, Miguel Ángel Fortes, 2010. Spatial cournot equilibrium: do branches matter? Ann. Reg. Sci. 44 (2), 377–407.

Christiaans, Thomas, 2002. Regional competition for the location of new facilities. Ann. Reg. Sci. 36 (4), 645–661.

Dorta-González, Pablo, Santos-Peñate, Dolores R., Suárez-Vega, Rafael, 2004. Cournot oligopolistic competition in spatially separated markets: the stackelberg equilibrium. Ann. Reg. Sci. 38 (3), 499–511.

Drezner, Tammy, Drezner, Zvi, Shiode, Shogo, 2002. A threshold-satisfying competitive location model. J. Reg. Sci. 42 (2), 287–299.

María Dolores, García, Pelegrín, Blas, Fernández, Pascual, 2011. Location strategy for a firm under competitive delivered prices. Ann. Reg. Sci. 47 (1), 1–23.

Huck, Steffen, Müller, Wieland, Vriend, Nicolaas J., 2002. ''the east end, the west end, and king's cross: on clustering in the four-player hotelling game. Econ. Inq. 40 (2), 231–240.

Loertscher, Simon, Muehlheusser, Gerd, 2011. Sequential location games. Rand J. Econ. 42 (4), 639–663.

Meagher, Kieron J., Zauner, Klaus G., 2005. Location-then-price competition with uncertain consumer tastes. Econ. Theor. 25 (4), 799–818.

Pires, Cesaltina Pacheco, 2009. Location choice under delivered pricing: a reinterpretation. Ann. Reg. Sci. 43 (1), 199–213.

Sun, Chia Hung, 2012. Sequential location in a discrete directional market with three or more players. Ann. Reg. Sci. 48 (1), 101–122.