Spatial Disparity on the Impact of Real Estate Investment on Regional Economy in China: MGWR Approach

Colombo Economic Journal (CEJ)
Volume 2 Issue 2, December 2024: PP 87-106
ISSN 2950-7480 (Print)
ISSN 2961-5437 (Online)

Published by Department of Economics, University of Colombo, Sri Lanka
Website: https://arts.cmb.ac.lk/econ/colombo-economic-journal-cej/

Hua Zheng
Graduate School of International Development, Nagoya University, Nagoya, Japan

Correspondence: zhenghua97@yahoo.com

Received: 01 September 2024, **Revised**: 25 October, 2024, **Accepted**: 18 November 2024.

Abstract

This paper explores, for the first time, the regional disparity in the relationship between real estate investment and regional economic performance using quantitative and visual approaches. By introducing the MGWR (Multi-Scaled Geographically Weighted Regression) model, we conduct a regression analysis on cross-sectional urban data for 2020. The findings reveal that the greatest contribution to regional economies is made by economically developed regions along the eastern coast, as well as some inland cities such as Baotou, Hohhot, and Taiyuan, followed by the eastern and southeastern coastal areas and the central-western regions. In contrast, real estate investment in the western region contributes the least to regional economic development.

***Keywords**: Real Estate Investment, China, Regional Disparity, MGWR, Spatial Econometrics*
JEL Classification: C23, R10, R31

Introduction

Since market reforms of the real estate sector in 1998 China's real estate industry has developed rapidly at an average annual growth rate of 21.3%. Over the past two decades, it has become a crucial driver of China's economic growth, closely intertwined with the broader economy. The sector's highest growth rate was 27% in 2003, surpassing the 25% growth rate of fixed asset investment in the same year. As a major driver of China's economy, the real estate sector has spurred the development of various industries, including construction and finance, thereby boosting overall economic output. However, this real estate boom has also drawn a substantial amount of social wealth into the sector, which has potentially stifled the growth of other industries (Glaeser *et al.*, 2017). However, in recent years, China has faced a slowdown in economic growth and an imbalance in its industrial structure, with resource misallocation being a primary issue.

The real estate sector, due to its substantial investment requirements and long economic cycle, has emerged as a significant source of this resource misallocation, raising questions about its overall contribution to economic development. Additionally, there is a pronounced spatial imbalance in both economic development and real estate sector growth across regions in China.

Spatially, although the disparities in real estate investment among the eastern, central, and western regions of China have been narrowing in recent years, a clear pattern remains: the East is the most developed, the Central region is less developed, and the West remains underdeveloped. In addition to these inter-regional differences, intra-regional disparities are also significant. Real estate investment tends to be concentrated in capital cities and major urban centers such as Beijing, Shanghai, Guangzhou, and Shenzhen. Since 2011, China's 35 large and medium-sized cities, out of more than 300 cities nationwide, have consistently accounted for more than 50% of total real estate investment.

Similarly, the geographical distribution of China's regional economies clearly reflect the distribution of real estate investment, with the eastern coast leading, the central region being weaker, and the western region significantly lagging. While the relationship between real estate investment and regional economic development has been widely discussed in academic circles, most studies have focused on the temporal linkages and changes over time. However, studies exploring the relationship from a spatial perspective remain scarce.

To address this gap, this paper focuses on the spatial pattern of the impact of real estate investment on regional economies, providing a detailed analysis of regional differences across space. This study also innovatively considers real estate investment as a stock indicator, thereby addressing the limitations of most existing studies that rely on flow indicators. Additionally, this paper uses city-level data rather than the more commonly used provincial-level data, offering a more precise representation of spatial differences. This paper is structured into six sections: an introduction, a literature review, a theoretical framework, a methodology and data explanation, a discussion, and finally, the conclusions.

Literatures Review

Given the strong intersectoral linkages of the real estate sector with other industries and its significant contribution to economic growth, the real estate sector can be considered a key driver of economic development (Hirschman & Sirkin, 1958; Polenske & Sivitanides, 1990). Real estate investment can influence economic development through its impact on employment, savings, investments in other sectors, and labor productivity (Harris & Arku, 2006).

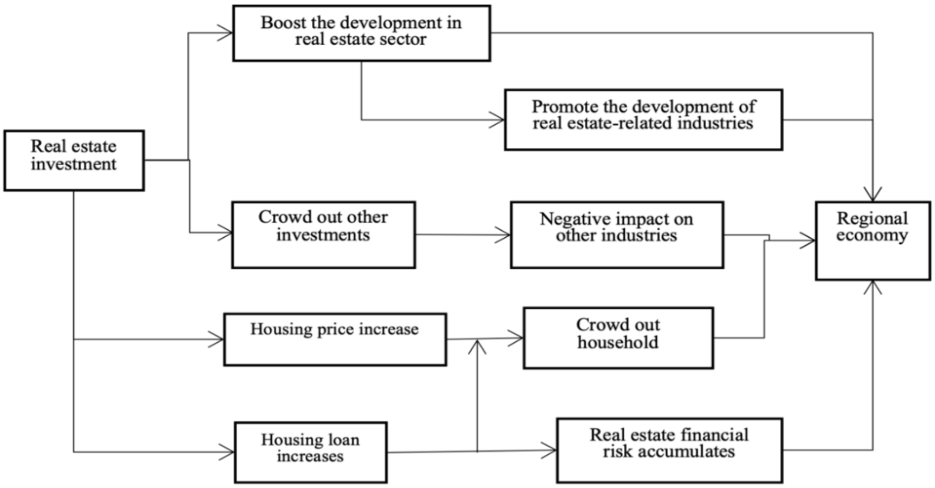
Since marketization of real estate investment in China, the sector has developed rapidly, attracting considerable attention from researchers. Liow and Yang (2005) foundthat both in the short and long run, real estate investment is a major driving force for economic growth. Li-xi (2014) empirically demonstrated that growth in real estate investment can stimulate development in the finance and insurance sectors, construction, non-metallic mineral products, hotels, restaurants, and other business activities. Song and Liu (2007) analyzed the role of the real estate sector in the Chinese economy from 1997 to 2002, finding that its contribution to the economy, as well as its push and pull effects (backward and forward linkages), were stronger in 2002 than in 1997. Furthermore, Li et al. (2008) indicated that the pull effect of real estate investment is stronger than its push effect. On the other hand, Ding and Lichtenberg (2010) emphasized that rapid growth in real estate poses a threat to China's balanced economic development. Liu and Chen (2011) found that while the short-term effects of real estate investment on economic growth are positive, they turn negative in the long term, suggesting that real estate investment may crowd out other investments and consumption, ultimately harming the national economy. Concerns about the impact of real estate investment on China's economy are growing and remain a controversial topic.

China's regional economic development is highly uneven (Chen et al., 2011), and the impact of real estate on regional economies is also likely to vary across different regions. However, there has been limited research addressing this question. Zhang et al. (2012) identified a threshold effect and different patterns of Granger causality among the eastern, central, and western regions of China. Ren et al. (2014) confirmed that the impact of the real estate sector on GDP and employment varies spatially by studying 30 sample cities using input-output tables.

Previous literature has two main shortcomings. First, it often overlooks the spatial disparity of the impact of real estate investment on regional economies. This is especially relevant in China, where economic disparities are vast, and economic structures vary across regions, meaning that the efficiency of real estate investment's impact on regional economies also differs spatially. This paper addresses this gap by applying the MGWR model to capture the impact on each region individually. Second, much of the previous research relies on provincial-level data, which is not ideal for revealing regional differences in a country as large and diverse as China. To address this, city-level data will be used in this analysis to better capture internal differences. As a result, this study aims to provide more effective guidance for regional development. Additionally, this paper incorporates geographical influences into the analysis to comprehensively examine the spatial differences in the impact of real estate investment on regional economies.

Theoretical Frameworks

According to the Cobb-Douglas production function, capital stock plays a critical role in the economic development of a country or region. Among various types of investments, real estate investment is particularly influential in driving economic growth. This process can be illustrated as shown in Figure below. On the one hand, real estate investment can stimulate the growth of the real estate industry and its related sectors, thereby accelerating economic development. On the other hand, if a significant portion of social wealth is concentrated in the real estate sector, investments in other sectors may be constrained, ultimately hindering their development—especially in manufacturing. Additionally, when housing prices become excessively high, residents are forced to allocate a larger share of their income and savings toward housing. In China's case, both developers and homebuyers often rely heavily on bank loans. When these loans grow too large, financial stability may be threatened.



Source: Created by Author

Figure 1: The Influencing Framework of Real Estate Investment on the Regional Economy

Methodology and Data Selection

To better understand the spatial variations in the relationship between real estate investment and regional economies, this analysis employs spatial econometrics, specifically the MGWR model. Unlike traditional statistical methods, spatial econometrics incorporates spatial weights, enabling the analysis of geographical relationships among regions.

Geographically Weighted Regression (GWR)

Before delving into the MGWR model, it is helpful to first introduce the GWR model. The GWR model can be seen as a localized version of the OLS (Ordinary Least Squares) model, but with a key difference in sampling methodology. While OLS calculates relationships by including the full sample, GWR focuses only on samples within a specific geographical distance or unit referred to as the "bandwidth." GWR assigns weights to regions within this bandwidth and computes results specific to each region.

In terms of coefficients, OLS provides an average effect across all samples. In contrast, GWR, by assigning the highest weight to a specific region and reducing the weight for neighboring regions within the bandwidth, yields results specific to the region with the highest weight. The fundamental concept behind GWR is that relationships between two factors in space may be relatively constant and similar within a certain distance or unit but may differ beyond that range. The former is spatial homogeneity, and the latter is spatial heterogeneity (Gu et al., 2021). The GWR model allows for the embedding of spatial geographical information into the estimation of coefficients, providing estimates for each sample point by applying information from relevant sub-samples rather than relying on the entire sample (Fotheringham, Charlton & Brunsdon, 1998).

The GWR model can be expressed as follows:

$$y_i = \sum_{j=0}^{m} \beta_j(u_i, v_i)x_{ij} + \varepsilon_i \quad (1)$$

Where y_i represents the explained variable, (u_i, v_i) is the location information, u_i is the easting, and v_i is the northing. m is the total amount of X variables; $\beta_j(u_i, v_i)$ is the j^{th} coefficient for city i. X_{ij} is the observation value of the variable j at the sample point i; ε_i is the error term.

Multiscale Geographically Weighted Regression (MGWR)

MGWR is an extension of GWR that relaxes the assumption that all spatially varying processes operate at the same spatial scale. Instead, MGWR allows the bandwidth for each independent factor to be flexible (Fotheringham et al , 2017). The core idea is that different explanatory variables exert varying degrees of spatial influence on the dependent variable. In practice, this means that different sample sizes are assigned to different variables, depending on the spatial influence of those variables. The formula is as follows:

$$y_i = \sum_{j=0}^{m} \beta_{bwj}(u_i, v_i)x_{ij} + \varepsilon_i \quad (2)$$

Compared to equation (1), the difference of equation (2) is the bwj in β_{bwj} indicates the bandwidth used for calibration of the jth conditional relationship, different variable j has different bandwidth bwj. Sample size for different variable is different. (u_i, v_i) is the location information, j is jth variable; i is the city number i, m is the number of variables. ε_i is the error term.

Data Selection

This paper aims to identify the spatial differences in the impact of real estate investment on regional economies. The dependent variable in the analysis is GDP per capita (lngdppc), while the primary explanatory variable is real estate investment per capita (lnreinv). Several control variables are also included: R&D funds per capita (lnsci), urbanization rate (U), the ratio of bank loans to GDP (loan_gdp), the share of high school students in the total population (hs), the number of people working in the tertiary sector per 10,000 population (lnterp), and the ratio of secondary industry to GDP (sec_gdp).

The lnsci variable represents R&D funds, which can be seen as a proxy for technological progress, and its impact on regional economies is expected to be positive. Urbanization reflects better resource agglomeration, which can boost regional economies and is also expected to have a positive impact. A higher ratio of bank loans to GDP suggests increased investment, while lnterp and sec_gdp indicate an upgrading of the industrial structure, both of which are anticipated to positively influence regional economic performance. Table 1 provides a summary of the variables used in the analysis.

Table 1: Variable Definition and Notation

Variable	Notation
lngrppc	Natural logarithm of GDP per capita
lnreinv	Natural logarithm of real estate investment per capita
lnsci	Natural logarithm of R&D fund per capita
U	Urbanization
loan_gdp	The share of loans outstanding to GDP
hs	The share of high school student to total population
sec_gdp	The share of secondary sector to GDP
lnterp	Natural logarithm of the employees in tertiary sector per 10000 people

Source: *China City Statistics Yearbook*
Note: Created by Author

Modeling

Firstly, we build OLS model. Based on the Cobb-Douglas production function, we assume that, as below:

$$Y = AK^{\alpha}L^{1-\alpha} \tag{3}$$

Where Y is the total production, A is including all technological development; K is social fixed investment; L is the labor force, α, $1-\alpha$ are the contributions of capital and labor force to total production respectively. Then, we divide both sides of the equation by L, and then take the natural logarithm:

$$\ln y = \ln A + \alpha \ln k \tag{4}$$

Where y is the GDP per capita, k is the capital stock per capita. Next, we add some control variables and get OLS model as below:

$$lngrppc_i = \beta_0 + \beta_1 lnreinv_i + \beta_2 lnsci_i + \beta_3 \ln U_{i+} \beta_4 lnloan_{gdp_i} + \beta_5 \ln hs_i + \beta_6 \sec\_gdp_i + \beta_7 \ln Interp_i + \varepsilon_i \tag{5}$$

Where *i* is the sample city *i*. ε_i is the error term.
Based on the OLS model, GWR model is built as below:

$$\ln grppc_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)\ln reinv_i + \beta_2(u_i, v_i)\ln sci_i + \beta_3(u_i, v_i)\ln U_i + \beta_4(u_i, v_i)\ln loan_{gdp_i} + \beta_5(u_i, v_i)\ln hs_i + \beta_6(u_i, v_i)\sec \quad gdp_i + \beta_7(u_i, v_i)\ln lnterp_i + \varepsilon_i \tag{6}$$

Where (u_i, v_i) is the location information of city *i*.

Based on the GWR model, MGWR model is conducted as below:

$$\begin{aligned} \ln grppc_i = {} & \beta_{bw0}(u_i, v_i) + \beta_{bw1}(u_i, v_i) \ln reinv_i + \beta_{bw2}(u_i, v_i) \ln sci_i \\ & + \beta_{bw3}(u_i, v_i) \ln U_i \\ & + \beta_{bw4}(u_i, v_i) \ln loan_{gdp_i} \\ & + \beta_{bw5}(u_i, v_i) \ln hs_i + \beta_6(u_i, v_i) \sec_{gdp\quad i} + \beta_{bw7}(u_i, v_i) \ln lnterp_i \\ & + \varepsilon_i \end{aligned} \tag{7}$$

β_{bwj} is the coefficient under j^{th} bandwidth (number of sample data). (u_i, v_i) is same as above in GWR model.

Data Source and Indicator Selection

(1) Explained variable: lngrppc; we take the nominal GRP (Gross Regional Production) from *China City Statistical Yearbook* in 2020. Sample size is 283.

(2) Main explaining variable: lnreinv; It is the real estate investment per capita, which in the Cobb-Douglas function is the stock indicator. Since the data in China city Yearbook is the investment flow, we need to change it into stock. And so far, the commonly used method of measuring capital stock is the perpetual inventory method pioneered by Goldsmith in 1951(Goldsmith, 1951). The function is as below:

$$K_t = K_{t-1}(1-\sigma) + I_t \tag{8}$$

Where K_t is the capita stock in year t, σ is the depreciation rate of capita stock, I_t is the investment in year t. To be consistent with the intrinsic meaning of the perpetual inventory method, and under the assumption that the relative efficiency of capital goods decreases in a geometric manner, an approach consistent with that of Zhang (2003) was adopted for the calculation of the depreciation rate of the real estate investment stock, and the balance depreciation method representing decreasing geometric efficiency was chosen.

$$d_\tau = (1-\sigma)^\tau, \tau = 0,1, \ldots \tag{9}$$

Where d_τ is the relative efficiency of capital goods, in other words, is the marginal production efficiency of old capital goods relative to new capital goods. σ is same as above, the depreciation rate of capital stock, τ is period. In the study by Huang Yongfeng et al, the relative efficiency of capital goods, $d\tau$, was replaced by the statutory residual value rate in China, with a value of 3-5%, and the middle value of 4% was used in equation above (Huang, Ren and Liu, 2002). Since the construction period of real estate investment is generally about 3 years, and the sales period and renovation period are 1-5 years, considering the impact of the later use and refurbishment of the housing stock, a depreciation period of 10 years is chosen to calculate the depreciation rate of real estate investment, and the result of σ is 27.5% according to the geometrically decreasing balance depreciation method listed in equation above. For the initial investment stock in 2000, we take the average geometric growth rate from 2000 to 2020. According to the methodology be given by Hall and Jones to compute the capital stock for each country in 1960. We use the following equation to compute the initial investment stock in 2000.

$$K = \frac{I}{\theta + \sigma} \tag{10}$$

Where θ is the average geometric growth rate from 2000 to 2020, σ is same as above, the depreciation rate of capital goods.

This paper uses panel data from 2000 to 2020 to computing the capital stock of real estate investment. Considering the differences in inflation over time and space, we have used the inflation in 2000 as the basis and inflation in Beijing as the benchmark for deflating the price data for 283 cities in 21 years.

(3) Control variables: All data for control variables in Table 1 are taking from *China City Statistical Yearbook* in 2020.

(4) To address missing value, we used the miss Forest package in R studio for imputation. (Amount of missing data/Amount of total data=104/4811)

Table 2: Descriptive Statistics for MGWR Analysis

Variable	Obs	Mean	Median	Std.Dev	Min	Max
lngrppc	283	7.73	7.65	0.65	6.30	10.13
lnreinv	283	6.55	6.36	0.05	4.23	9.24
lnsci	283	1.53	1.43	1.21	-0.97	6.25
U	283	0.37	0.31	0.24	0.05	1.00
loan_gdp	283	1.05	0.88	0.58	0.33	3.71
hs	283	0.02	0.01	0.02	0.00	0.13
sec_gdp	283	0.45	0.46	0.09	0.15	0.71
lnterp	283	6.32	6.18	0.52	5.40	8.49

Source: *China City Statistics Yearbook,* National Bureau of Statistics
Note: Computed by Author using Stata

Table 2 above presents the descriptive statistics for all the variables used in the analysis. Urbanization is measured by the ratio of the urban population (those with non-agricultural residency) to the total population, with a maximum value of 1. For instance, in cities like Beijing and Shanghai, which are nearly fully urbanized, there is essentially no rural population within the city limits. The average year-end loan balance to GDP ratio exceeds 1, partly due to accommodative monetary policies. The ratio of secondary industry output to GDP is used as an indicator of a city's degree of industrialization. The table shows an average value of 0.45, which is close to 0.5, reflecting the success of industrialization reforms implemented in China following the reform and opening-up period. The number of people employed in the tertiary sector serves as a proxy for the level of structural upgrading and technological progress.

Result and Discussion

Regression Results

Before proceeding to the MGWR model, we conducted a Variance Inflation Factor (VIF) test for each variable to avoid multicollinearity. If the VIF result is less than 10 and the 1/VIF value is greater than 0.1, it can be considered that there is no multicollinearity among the variables. As shown in Table 3, the VIF for all variables is below 4.0, confirming that multicollinearity is not an issue, allowing us to proceed with the MGWR analysis.

Table 3: VIF Result for MGWR Analysis

Variable	VIF	1/VIF
lnreinv	3.51	0.28
lnsci	2.53	0.36
U	2.01	0.50
loan_gdp	2.45	0.41
hs	2.75	0.36
sec_gdp	1.44	0.69
lnterp	3.52	0.28

Source: *China City Statistics Yearbook,* National Bureau of Statistics
Note: Computed by Author using Stata

To facilitate better comparison and understanding of the MGWR model, we conducted analyses using both the OLS (Ordinary Least Squares) model and the MGWR model. The results for both models were generated using the MGWR 2.2.1 software application. Table 4 and Table 5 below present the results of the OLS and MGWR models, respectively. The MGWR model, with its higher adjusted R^2, lower residual sum of squares, and lower AIC (Akaike Information Criterion), can be considered a superior model for interpreting the data.

Table 4: Results of OLS Model

lngrppc	Coef.	St.Err	t-value	p-value	Sig.
Intercept	-0.000	0.024	-0.000	1.000	***
lnreinv	0.388	0.045	8.638	0.000	***
lnsci	0.180	0.038	4.700	0.000	***
U	0.085	0.034	2.512	0.012	**
loan_gdp	-0.240	0.038	-6.378	0.000	***
hs	0.089	0.040	2.229	0.026	**
sec_gdp	0.175	0.029	6.086	0.000	***
lnterp	0.406	0.045	9.014	0.000	***
R-squared	0.842		Residual sum of squares		44.80
Adjust R-squared	0.838		Number of obs.		283.00
AIC	297.501		AICc		300.16

Source: *China City Statistics Yearbook,* National Bureau of Statistics
Note: Computed by Author using Stata; *** p<0.01, ** p<0.05, * p<0.1. Coef. refers to the influence coefficient, St.Err is the standard error, Sig. is the significance.

Table 5: Summary of Results for MGWR Model

lngrppc	Coef.	SD	min	median	max
Intercept	-0.079	0.173	-0.561	-0.061	0.245
lnreinv	0.356	0.064	0.248	0.362	0.479
lnsci	0.221	0.106	0.006	0.252	0.427
U	0.059	0.081	-0.127	0.067	0.221
loan_gdp	-0.251	0.016	-0.269	-0.256	-0.204
hs	0.123	0.001	0.120	0.123	0.125
sec_gdp	0.256	0.014	0.229	0.255	0.282
lnterp	0.370	0.155	0.063	0.331	0.800
R-squared		0.938	Residual sum of squares		17.525
Adjust R-squared		0.923	Number of obs.		283
AIC		129.769	AICc		159.098

Source: *China City Statistics Yearbook,* National Bureau of Statistics

Note: Computed by Author using MGWR 2.2.1; Coef. refers to the influence coefficients of all observations, SD is the standard deviation.

It is important to note that while the OLS model is a global regression with only one set of coefficients for the entire dataset, the MGWR model is a local regression, producing a set of coefficients for each city resulting in 283 coefficients in this analysis. The coefficients presented in

Table 5 are the mean values derived from these 283 local coefficients. When comparing the coefficients between the OLS and MGWR models, the signs of each coefficient are consistent, and the coefficients for each variable are similar. However, it is noteworthy that the maximum and minimum values of the coefficient for urbanization (U) have opposite sign, indicating that urbanization progress has varying impacts on regional economies. A detailed plot of these local coefficients will be presented later. It can be tentatively observed that in both the OLS and MGWR models, real estate investment has a higher coefficient of influence on the regional economy than the other factors, which are 0.388 and 0.356(the coefficient of which in MGWR is averaged across the board), respectively. Initially, it is possible to see the relatively high returns of real estate investment to the regional economy.

Bandwidth Analysis

In addition to summarizing the results, Table 6 lists the bandwidths from the MGWR model, which are crucial for understanding the spatial variation in the impact of real estate investment on regional economies. Bandwidths can be expressed in two forms: distance or unit points. Using distance bandwidths can result in too few sample points in areas with low urban density, which could affect the accuracy of the results—particularly in China, where geographic distribution varies significantly.

As mentioned earlier, bandwidth represents the sample range within which the MGWR model applies its local regression, meaning that within this bandwidth, the similarity of the impact from the independent variable (X) to the dependent variable (Y) is high. Generally, a larger bandwidth indicates greater spatial similarity, while a smaller bandwidth suggests greater spatial variability. This can be reflected in the coefficients by the fact that variables with wider bandwidths have smaller differences in impact coefficients, while on the contrary, variables with smaller bandwidths have larger spatial differences in impact coefficients. This difference is due to the changing sample when calculating the effect of variable X on variable Y in different regions (here the sample is a re-sampling of the full sample, which is equal to the bandwidth, which is always smaller than the total sample size).

The largest bandwidth, 282, is assigned to the share of high school students in the total population. This is nearly equivalent to a global bandwidth, given the sample size of 283, meaning that the impact of high school education on economic development does not vary significantly across different regions. This indicates that education's contribution to economic development is consistently strong across China. Similarly, the share of loan balances of financial institutions to GDP has a bandwidth of 250, indicating minimal spatial heterogeneity, which suggests that the importance of loans to economic development is relatively uniform across regions.

Table 6: Bandwidths for Each Variable in MGWR Model

Variable	Bandwidth	ENP_j	Adj t-val(95%)	DoD_j
Intercept	43	14.882	2.958	0.522
lnreinv	117	4.165	2.528	0.747
lnsci	51	10.995	2.860	0.575
U	76	8.005	2.755	0.632
loan_gdp	250	1.955	2.245	0.881
hs	282	1.222	2.054	0.964
sec_gdp	259	2.070	2.267	0.871
lnterp	47	12.655	2.906	0.550

Source: *China City Statistics Yearbook*, National Bureau of Statistics

Note: Computed by Author using MGWR 2.2.1; ENP_j is the effective number of parameters; Adj t-val (95%) denotes the adjusted t value adjusted t values; DoD_j is the degree of dependency.

The bandwidth for R&D funding is 51, which is relatively small, indicating significant spatial heterogeneity. In China, more than half of the country's R&D expenditure comes from large and medium-sized enterprises. These companies will only invest in R&D if they expect a substantial return on investment. However, the outcomes of R&D investment vary from place to place due to differences in know-how, the quality of human capital, and other factors, which ultimately contribute to varying degrees of economic development (Boeing, Eberle & Howell, 2022). The bandwidth for urbanization and employees in the tertiary sector per 10,000 people is also relatively small, at 76 out of

283 cities. The common understanding is that higher urbanization leads to better regional economic outcomes. Since 2014, the Chinese government decided to cancel the division between non-agricultural and agricultural registered permanent residence. To promote this change, once people move from agricultural to non-agricultural residency for reasons such as study or employment, they cannot move back. The urbanization calculation here is based on the ratio of the population with non-agricultural residency to the total population. Consequently, urbanization may not uniformly contribute to economic development.

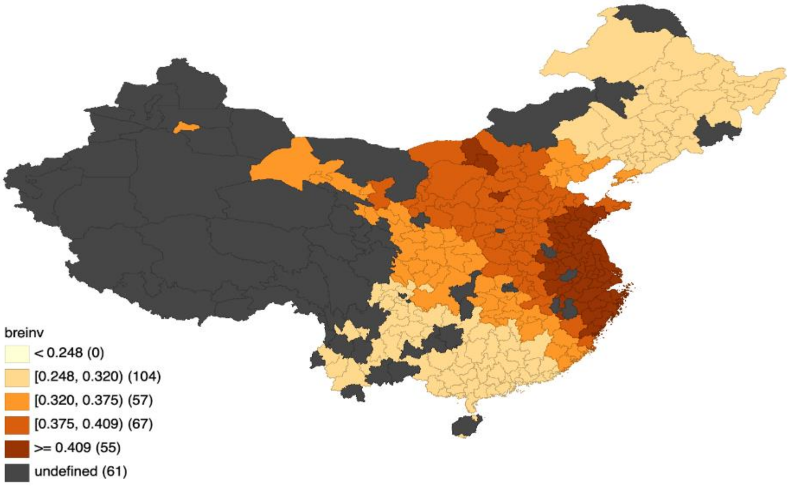
In terms of employees in the tertiary sector per 10,000 people, the small bandwidth suggests significant spatial variability. This variation is likely due to the differing nature of tertiary industries across regions. For example, in the more developed eastern areas, the financial and information technology sectors, which require less human capital but offer high returns, dominate. In contrast, in the southwest, tourism a labor-intensive industry with relatively lower returns is more prominent. The ratio of the secondary sector to GDP, with a bandwidth of 259, implies less spatial variation. As the backbone of the real economy, the contribution of the secondary sector to economic development is more consistent across regions, provided there are no significant differences in social productivity and that marginal returns do not diminish significantly. The Chinese government's commitment to achieving carbon neutrality by 2060 has led to higher efficiency demands on enterprises in resource use. Recent advancements in information technology have also facilitated closer technological exchanges between enterprises, resulting in minimal differences in the efficiency of secondary industry enterprises across cities.

The bandwidth for real estate investment is 117, roughly equivalent to one-third of the sample size of 283, indicating that the contribution of real estate investment to regional economies is less consistent across space. As mentioned earlier, the impact of real estate investment on economic development is multidimensional and involves multiple sectors, making its channels of influence more complex. Variations in human resources, financial development, and policy factors across regions contribute to significant differences in how real estate investment affects regional economies. The bandwidth of the intercept is the smallest at 43, suggesting that there are inherently large differences in the levels of regional economies between cities, beyond the effects of the listed control variables.

One of the key advantages of the MGWR model is its ability to precisely estimate the impact coefficients of explanatory variables on dependent variables in each city, which plays a critical role in this study. This capability allows for a detailed understanding of the spatial distribution of impacts. These impact coefficients are plotted in Figure 1.

Spatial Distribution of Real Estate Investment Impact

Figure 1 illustrates the spatial distribution of the impact of real estate investment on regional economies. The colors in the graph represent magnitudes, with darker shades indicating larger magnitudes. A clear trend emerges from the figure.



Source: *China City Statistics Yearbook*, National Bureau of Statistics

Note: breinv refers to the beta coefficient of real estate investment to the regional economy; The number in parentheses indicates the number of cities with impact coefficients in that range.

Figure 1: Coefficients of Real Estate Investment on Regional Economy

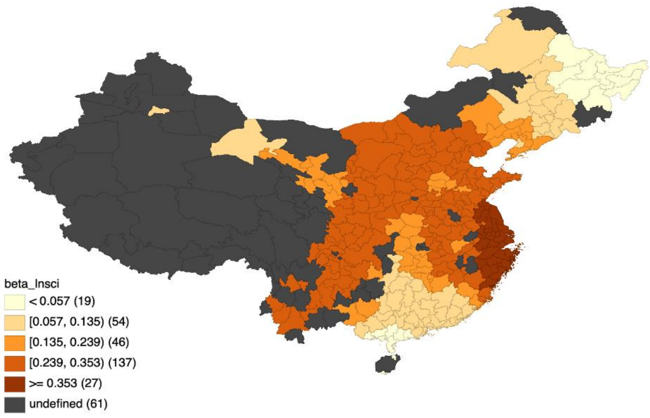
The most significant impacts of real estate investment on regional economies are concentrated along the eastern seaboard, including Shandong, Zhejiang, Shanghai, and Jiangsu, as well as in some inland cities, most notably Baotou (the capital city of Inner Mongolia), Hohhot (a city in Inner Mongolia), and Taiyuan (the capital city of Shanxi province). This is followed by the central region, including Henan, Shanxi, and Hubei, as well as areas in the northeast and southwest, where the coefficient of real estate investment on regional economies diminishes. Due to its status as a financial center, Shanghai attracts a variety of financial resources, covering the eastern coast of Jiangsu and Zhejiang. Additionally, the Yangtze River Delta industrial base, with its high industrial efficiency and financial resources, coupled with the lower level of economic development in surrounding provinces, draws talent to these three areas. This creates a robust demand for real estate in these regions, which is not as pronounced elsewhere.

The high returns on real estate investment in the East can be attributed to the East's superior economic performance and the large number of labor opportunities, combined with the commodity and investment nature of real estate. As a commodity, more real estate is built in the East, which makes more productive economic activities possible on the business side, through the clustering effect of industries, which makes economic returns better than elsewhere; and on the other hand, on the residential side, through the provision of a living environment, which makes it possible for residents to devote themselves to productive activities with peace of mind. All this contributes to the regional economy. As an investment, rising housing prices are also an important driver of GDP growth. The Eastern region has a greater urbanization drive than other regions, with more people moving from rural to urban areas, creating more demand for housing. More real estate investment, on the one hand, consumes this demand for housing and balances out the price increase due to higher demand, but on the other hand, it raises the average price of housing by increasing real estate investment per unit area. And the existence of speculative behavior is also driving up prices.

The northern cities of Baotou, Hohhot, and Taiyuan also show a high contribution of real estate investment to their economies. Despite their inland locations, these cities rank relatively low in overall economic performance within China, with Baotou at No. 92, Hohhot at No. 98, and Taiyuan at No. 50 out of 344 cities. In this context, the significant contribution of real estate investment to the regional economy highlights an issue: the unbalanced industrial structure of these areas. The high impact coefficient of real estate investment on the local economy suggests that a substantial portion of economic growth in these cities depends heavily on real estate development.

Conversely, the Southwest and Northeast regions exhibit the lowest contribution of real estate investment to local economies. The Southwest has historically lagged behind other regions in economic development due to its rugged terrain, which hinders development, and poor transportation infrastructure. As a result, the housing market in this region is more focused on meeting residential demand than on speculation and investment, leading to a less significant contribution from real estate investment. The Northeast, which was a hub of industry—particularly heavy industry in the 1970s and 1980s, has struggled since the depletion of its resources and the lack of new industries to replace the heavy industry. This has led to a significant brain drain and the emergence of "empty cities," exacerbating the region's economic challenges.

Figure 2 illustrates the spatial distribution of the impact of R&D funding on regional economies. The small bandwidth for R&D funding indicates substantial spatial heterogeneity. Nevertheless, a clear pattern emerges, with the largest contributions of R&D funding to the economy concentrated in Shanghai and its surrounding areas, followed by a transition to the central western and northern regions, and then a decline towards the south and northeast. The Northeast and southern regions show the lowest contribution from R&D funding.



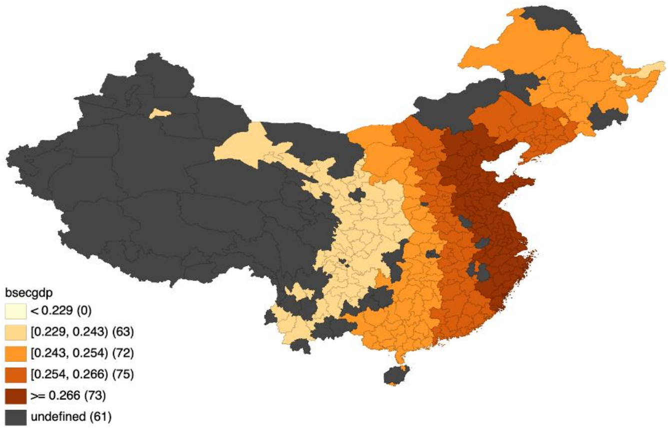
Source: Created by Author using Geoda

Note: beta_lnsci refers the beta coefficient of R&D fund to the regional economy; The number in parentheses indicates the number of cities with impact coefficients in that range.

Figure 2: Coefficients of R&D Fund on Regional Economy

In 2020, the total economic output of the eastern Yangtze River Delta region accounted for 24% of the national economy, underscoring its significant economic strength. The region is home to 20 Global 500 enterprises, which further highlights its economic clout. The concentration of enterprises in this area attracts substantial R&D funding, and combined with a continuous influx of talent, this results in a high return on R&D investments. The central region and some parts of the western regions also benefit from R&D funding, though to a slightly lesser extent. Since more than half of China's R&D funds come from private companies, these enterprises are typically more inclined to invest in R&D only when they anticipate a favorable return on investment, leading to relatively stable returns overall. In contrast, the Northeast's lack of significant returns on R&D investment tends to be closely related to the region's talent drain and the absence of new industry development. Illustrates the spatial distribution of the contribution of secondary industries to regional economies. Although the bandwidth is large and spatial differences are minimal, a clear trend of east-west variation is still apparent. The largest contributions of the secondary industry ratio to the economy are concentrated in the eastern coastal regions, stretching from north to south, including Beijing, Tianjin, Shandong, Jiangsu, Shanghai, Zhejiang, and Fujian. This encompasses the Beijing-Tianjin Industrial Base and the Yangtze River Delta Industrial Base. These are the industrial bases of the Pearl River Delta, which are transitioning towards the central region, and the industrial bases of Liaoning, Central, and South China in the northeast.

Essentially, the emergence of industrial bases has led to a higher concentration of enterprises, facilitating better upstream and downstream linkages, promoting cooperation, and stimulating competition. These factors collectively contribute to enhanced economic development. However, numerically, there is little variation in the spatial contribution of the secondary sector's GDP ratio to overall economic development.

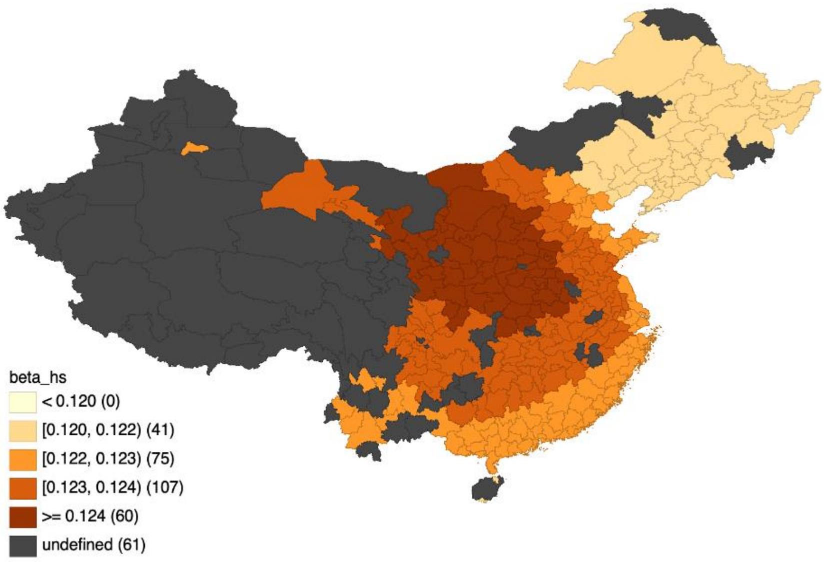


Source: Created by Author using Geoda

Note: bsecgdp is the coefficient of share of the secondary sector on the regional economy; The number in parentheses indicates the number of cities with impact coefficients in that range.

Figure 3: Coefficients of Share of Secondary Sector to GDP on Regional Economy

The contribution of the ratio of high school students to the total population to the regional economy is depicted in . Similar to the secondary sector, the contribution of high school students to economic development shows minimal spatial variation, although slight differences do exist. The highest impact coefficients are concentrated in the central part of the country, particularly in the central and western regions, with a gradual decrease towards the east and northeast. The regions where high school students contribute most significantly to the regional economy include Shaanxi, Gansu, Shanxi, and Inner Mongolia. Except for Shaanxi, the other three provinces are educationally underdeveloped. With the implementation of the Western Development Strategy, a large number of industries and companies have relocated to these regions, especially in Shaanxi province. This relocation has enabled these areas to better and more rapidly translate educational achievements into productivity. Although there are spatial differences in the impact, they are not pronounced, and the overall contribution of education to the economy is relatively consistent. This consistency further underscores the idea that the level of education has a broadly positive effect on China's economic development.



Source: Created by Author using Geoda

Note: beta_hs is the coefficient of share of high school student to total popu-lation on regional economy; The number in parentheses indicates the number of cities with impact coefficients in that range.

Figure 4: Coefficient of the Share of High School Student to Total Population on Regional Economy

Conclusions

The primary objective of this thesis is to explore the spatial disparity and distribution of the impact of real estate investment on regional economies in China. By using the MGWR model, this study estimates the coefficients of real estate investment's impact on the economies of 283 cities in China in 2016. The analysis reveals that there is inherent spatial heterogeneity in the economic impact of real estate, which is further influenced by variations in economic structure, labor quality, and economic efficiency across different cities. The results indicate a trend where the influence of real estate investment is strongest in the eastern regions and several inland cities, gradually weakening in the central, southwestern, and northeastern parts of China. The significant contribution of real estate investment to economic development in a few economically backward inland cities also highlights the structural imbalance in these areas' economies. To address this imbalance, timely economic restructuring and the vigorous development of the real economy should be key policy objectives for these regions.

Real estate investment emerges as the largest contributor to economic development among the factors studied, raising concerns about the potential risks of over-reliance on real estate for economic growth. The value of R&D investment shows spatial variability, with the majority of cities experiencing high returns. This suggests that sustained investment in R&D and innovation generally has a positive effect on economic development. The spatial differences in the contribution of the secondary sector to the economy are minimal, indicating the widespread importance of the real economy to China's overall economic growth. The higher contribution of the eastern regions, combined with the returns from R&D, suggests that the advantages of technological innovation are more pronounced in the East.

Acknowledgments: I sincerely thank CEJ for providing valuable feedback and constructive suggestions that greatly improved the quality of this research. Their insights were instrumental in refining my ideas and enhancing the overall clarity of the study.

References

Boeing, P., Eberle, J. and Howell, A. (2022) ‘The impact of China’s R&D subsidies on R&D investment, technological upgrading and economic growth’, *Technological Forecasting and Social Change,* 174, p. 121212. Available at: https://doi.org/10.1016/j.techfore.2021.121212.

Chen, J., Guo, F. and Wu, Y. (2011) ‘Chinese urbanization and urban housing growth since the mid-1990s’, *Journal of Housing and The Built Environment*, 26, pp. 219–232. Available at: https://doi.org/10.1007/s10901-011-9214-0.

Ding, C. and Lichtenberg, E. (2010) ‘Land and Urban Economic Growth in China*’, *Journal of Regional Science,* 51(2), pp. 299–317. Available at: https://doi.org /10.1111/j.1467-9787.2010.00686.x.

Fotheringham, A.S., Charlton, M.E. and Brunsdon, C. (1998) ‘Geographically Weighted Regression: A Natural Evolution of the Expansion Method for Spatial Data Analysis’, *Environment and Planning A: Economy and Space,* 30(11), pp. 1905–1927. Available at: https://doi.org/10.1068/a301905.

Fotheringham, A.S., Yang, W. and Kang, W. (2017) ‘Multiscale Geographically Weighted Regression (MGWR)’, *Annals of the American Association of Geographers,* 107(6), pp. 1247–1265. Available at: https://doi.org/10.1080/ 24694452.2017.1352480.

Glaeser, E. *et al.* (2017) ‘A Real Estate Boom with Chinese Characteristics’, *Journal of Economic Perspectives,* 31(1), pp. 93–116. Available at: https://doi.org/10.1257/jep.31.1.93.

Goldsmith, R.W. (1951) ‘A Perpetual Inventory of National Wealth’, in *Studies in Income and Wealth, Volume 14*. NBER, pp. 5–73. Available at: https://www.nber.org/ books-and-chapters/studies-income-and-wealth-volume-14/perpetual-inventory-national-wealth (Accessed: 15 May 2022).

Gu, H. *et al.* (2021) 'Analyzing the distribution of researchers in China: An approach using multiscale geographically weighted regression', *Growth and Change*, 52(1), pp. 443–459. Available at: https://doi.org/10.1111/grow.12453.

Harris, R. and Arku, G. (2006) 'Housing and economic development: The evolution of an idea since 1945', *Habitat International*, 30(4), pp. 1007–1017. Available at: https://doi.org/10.1016/j.habitatint.2005.10.003.

Hirschman, A.O. and Sirkin, G. (1958) 'Investment Criteria and Capital Intensity Once Again', *The Quarterly Journal of Economics*, 72(3), pp. 469–471. Available at: https://doi.org/10.2307/1882237.

Huang, Y., Ren, R. and Liu, X. (2002) 'Capital stock estimates in Chinese manufacturing by perpetual inventory approach', *China Economic Quarterly*, 1(2), pp. 377–396.

Li, Y., Song, Y. and Liu, C. (2008) 'Exploring the Increasing Role of the Real Estate Sector in the Chinese Economy: 1997–2002', *Pacific Rim Property Research Journal*, 14(3), pp. 279–297. Available at: https://doi.org/10.1080/14445921.2008.11104259.

Liow, K.H. and Yang, H. (2005) 'Long-Term Co-Memories and Short-Run Adjustment: Securitized Real Estate and Stock Markets', *The Journal of Real Estate Finance and Economics*, 31(3), pp. 283–300. Available at: https://doi.org/10.1007/s11146-005-2790-6.

Liu, Z. and Chen, L. (2011) 'Research on Chinese Real Estate Development and the Future Trends', *Asian Social Science*, 7(9). Available at: https://doi.org/10.5539/ass.v7n9p207.

Li-xi, Z. (2014) 'Urbanization and Regional Differences of Investment in Real Estate Development: Evidences Based on the Dynamic Panel Data Model', *Contemporary Finance and Economics* [Preprint].

Polenske, K.R. and Sivitanides, P. (1990) 'Linkages in the construction sector', *The Annals of Regional Science*, 24(2), pp. 147–161. Available at: https://doi.org/10.1007/BF01579729.

Ren, H., Folmer, H. and Van der Vlist, A.J. (2014) 'What role does the real estate–construction sector play in China's regional economy?', *The Annals of Regional Science*, 52(3), pp. 839–857. Available at: https://doi.org/10.1007/s00168-014-0613-5.

Song, Y. and Liu, C. (2007) 'An Input–Output Approach for Measuring Real Estate Sector Linkages', *Journal of Property Research*, 24(1), pp. 71–91. Available at: https://doi.org/10.1080/09599910701297697.

Zhang, J., Wang, J. and Zhu, A. (2012) 'The relationship between real estate investment and economic growth in China: a threshold effect', *The Annals of Regional Science*, 48(1), pp. 123–134. Available at: https://doi.org/10.1007/s00168-010-0388-2.

Appendixes

Bandwidth selection for GWR model and MGWR model

Model calibration for GWR model can be conducted using weighted least squares, the estimator for the coefficients at location (u_i, v_i) is shown in below:

$$\hat{\beta}(u_i, v_i) = [X^T W(u_i, v_i)]^{-1} X^T W(u_{i,} v_i) y \qquad (a)$$

Where the X is the design matrix and W(u_i,v_i) is the n x n spatial weighting matrix for location (u_i,v_i) whose off-diagonal elements are zero and whose diagonal elements denote the geographical weighting of observed data for point *i*. W(u_i,v_i) is the same for each relationship due to the same bandwidth being used for all the relationships in the model. And then a goodness-of-fit measure such as AICc is calculated where AICc is defined by:

$$AICc = 2nln(\hat{\sigma}) + nln(2\pi) + n\frac{n + tr(S)}{n - 2 - tr(S)} \qquad (b)$$

Where $\hat{\sigma}$ is the estimated standard deviation of the error term and tr(S) is the trace of the hat matrix S. The optimal bandwidth is that which minimizes AICc.

There are multiple ways to conduct the spatial weight matrix, one of which used for GWR model is the Gauss function:

$$w_{ij} = \exp\left(-\left(\frac{d_{ij}}{\beta}\right)^2\right) \qquad (c)$$

Where β is referred to as the bandwidth, refers to the non-negative decay parameter as a function of the distance between the weights. The larger the bandwidth, the slower the weight decay, the smaller the bandwidth, the faster the weight shrink. When the bandwidth is 0, only the weights on the regression points are 1, and the weights of all other observation points are infinitely close to 0. On the contrary, when the bandwidth is infinite, the weights of all observation points are infinitely close to 1.

The bandwidth for GWR is relatively straightforward, because only a single bandwidth is required. But in the case of MGWR, it allows the conditional relationships between the response variable and the different predictor variables to vary at different spatial scales. The bandwidth selection for MGWR model is using back-fitting algorithms. Following the logic of GAM, $\beta_{bwj}x_j$ in MGWR is defined as the jth additive term $\mathcal{F}_j$, resulting in the GAM-style MGWR:

$$y = \sum_{j=0}^{m} \mathcal{F}_j + \varepsilon \qquad (d)$$

The basic idea of back-fitting is to calibrate each term in the model with a smoother assuming that all the other terms are known. In the case of the Gaussian MGWR model, the smoother is the GWR estimator defined in equation (a).