Does high-speed rail affect the agglomeration of banks in China?

Yan Dong [a], Jun Huang [b], Ji Wu [c,*]

[a] *Research Institute of Economics and Management, Southwestern University of Finance and Economics, Chengdu, China*
[b] *Research Institute of Economics and Management, Southwestern University of Finance and Economics, Chengdu, China*
[c] *Research Institute of Economics and Management, Collaborative Innovation Center of Financial Security, Southwestern University of Finance and Economics, Chengdu, China*

ARTICLE INFO

ABSTRACT

This paper assesses the distributional impacts of high-speed rail (HSR) transportation on the agglomeration of banks in China. Using the data of county-level cities in China from 2000 to 2017, we find that the proliferation of HSR yields a positive impact on our indicators of banking agglomeration in large counties, while its impact in small counties is negative. Our results are consistent with a promotive effect of HSR lines to induce banking resources to be increasingly clustered toward large counties in China.

1. Introduction

The rapid proliferation of high-speed rail (HSR), typically characterized as rail transportation that operates at speeds exceeding 200 km per hour (Givoni, 2006), has been observed in many nations in recent years.[1] According to data from *International Union of Railways*, HSR is currently operational in 44 countries worldwide, with a total distance of over 52,000 km as of 2019. Additionally, approximately 23,000 km of HSR lines are currently under construction or planned. The policy objective driving the expansion of HSR has evolved from solely enhancing commuter transportation capacity to utilizing HSR as a tool for promoting economic integration and advancing overall economic and social well-being (Chen and Vickerman, 2019). Thanks to significantly enhanced market accessibility, the introduction of HSR has reshaped various economic aspects, including productivity (Wetwitoo and Kato, 2017), regional development (Amos et al., 2010; Ahlfeldt and Feddersen, 2018), employment concentration (Kim, 2000), labor's decisions regarding their work locations and mobility (Guirao et al., 2017; Willigers and van Wee, 2011), and even households' real estate demand (Chen and Haynes, 2015).

Since the debut of its first HSR line in 2008, China has rapidly surpassed numerous predecessors by constructing the most extensive HSR network globally. As of 2019, its total operational distance has exceeded 35,000 km, which accounts for almost two-thirds of the combined HSR distance worldwide. Currently, over 5000 km of HSR are under construction in China, accounting for approximately 44% of the total new HSR lines being built globally. In regards to the population residing along the routes of the HSR network, China has covered 95% of cities with a populace exceeding one million. Although already massive, China aims to further develop its HSR network by doubling its size to 70,000 km and linking all cities with a population of over 500,000 by 2035. In the online appendix (see

---

* Corresponding author at: Research Institute of Economics and Management, Southwestern University of Finance and Economics, Chengdu, Sichuan 611130, China.
  E-mail addresses: ydong@swufe.edu.cn (Y. Dong), 117020204012@smail.swufe.edu.cn (J. Huang), wuji@swufe.edu.cn (J. Wu).

[1] Despite the inauguration of the first modern high-speed railway, the *Shinkansen*, in Japan in 1964, HSR lines only became increasingly prevalent across countries, including both developed and developing ones, after the early 2000s.

Table A1), we present the total length of China' HSR network (in kilometers) for each year between 2007 and 2019, as well as a comparison to the global total length of HSR.

The advent of HSR in China gives rise to numerous inquiries regarding its economic and financial ramifications. In light of the accelerated flows of commuters and goods between cities and regions, are economic resources reallocated with the operation of HSR lines? And, does HSR have any distributional impacts on the agglomeration of economic resources in cities along these lines? These questions are only underexplored in existing literature. In this article, we focus on the spatially disparate impacts of HSR on the agglomeration of banks in small and large county-level cities on the lines, as banks hold a dominant position in the financing of Chinese enterprises and households.

From a theoretical standpoint, the results of HSR are ambiguous as the presence of HSR may yield an uncertain effect on the agglomeration of banks. On the one hand, HSR may foster significantly higher economic growth in small cities along its routes, thereby promoting increased demand for financial services. The increased demand for financial intermediation may lead to a greater number of banks establishing operations in small cities, as opposed to large cities where banks may have already opened offices. This could result in a "pull" effect, leading to a greater agglomeration of banks in small cities. On the other hand, however, HSR considerably reduces travel time, thereby reducing the mobility costs for economic agents to access financial institutions. This may prompt them to seek more favorable financial services from banks located in remote cities, rather than from those in their own city of residence. As banking institutions in large cities tend to possess superior human capital and offer more advanced financial services, the demand for financial services may be diverted from smaller cities. In other words, the decreased demand for financial services in small cities may lead to a "push" effect, causing banks to cluster toward larger cities. Hence, whether bank agglomeration would increase or decrease in small cities relies on the competing forces from the "pull" and "push" sides.

In this study, we use the data from county-level cities in China spanning from 2000 to 2017 to examine the impact of HSR on the agglomeration of banks in small and big cities. Our work conveys some key findings as follows: First, using the number of bank affiliates and that scaled by city population as the outcome variables, which proxy the agglomeration of banking resources in cities, we find a positive association between the operation of HSR lines and the agglomeration of bank affiliates, suggesting that HSR overall attracts banking resources to cities along the lines, in comparison to the periods before these cities were linked by the lines of HSR.

Second, large Chinese cities exhibit a greater cluster of banking affiliates, indicating a promotive force of HSR lines to agglomerate banking resources toward large cities. In contrast, after controlling for various potential determinants of banking agglomeration, the ceteris paribus impact of HSR in small cities is found to be negative. This finding suggests that HSR leads to a declined agglomeration of banks in small cities, in line with the hypothesis that the rapid transportation facilitated by HSR stimulates a stronger incentive of economic agents in smaller cities to pursue more advantageous financial services in distant, larger cities. This, in turn, prompts banks to expand their presence in these larger cities. Furthermore, the agglomeration of banking resources in favor of larger cities is evident not only across cities but also across broader regions.

Third, our conclusion is reinforced by the identification of distributional impacts of HSR operations on financial activities, including loans and deposits, in both small and large cities. This finding indicates that not only the accessibility but also the availability of financial resources is shifted more conspicuously toward large cities along the HSR lines, as opposed to smaller ones. As we use several real economic variables as outcome variables, including fixed asset investment and manufacturing and tertiary industry output, we also find consistent evidence that HSR has a more pronounced impact in large cities. This seemingly suggests notable differences in the real effects of HSR that may be attributed to unequal availability of bank financing.

To the best of our knowledge, our article is among the first ones to assess the impacts of HSR lines on the distribution of banking resources, particularly in China. This provides a supplementary contribution to the existing literature on the financial implications of HSR. Our research differs from many prior ones in various aspects. First, in contrast to previous studies that have examined the influence of HSR on the mobility of real economic activities, such as labor employment and investment, our research focuses on the impact of HSR on the allocation of banks. Our work bears important policy implications as the agglomeration of banks may have repercussions on the accumulation of real capital stock and consequently influence the long-term economic growth of the cities in question. Meanwhile, our investigation into the impact of HSR on the reallocation of financial resources enriches the existing literature on the redistributive effects of transportation development, as HSR has the potential to significantly improve market accessibility compared to conventional transportation infrastructures, such as highways and ordinary railways, and may probably have a more substantial impact.

Next, many previous studies have examined the varying effects of HSR by comparing outcome variables between core cities connected by HSR and peripheral cities without HSR. In contrast, we select in our sample only those cities that are linked by HSR lines, while excluding those cities that are not HSR-connected. Furthermore, we exclusively choose cities at the county level for our sample, excluding those with a higher administrative hierarchy (e.g., provincial capitals and prefecture-level cities) which typically possess greater economic might. One advantage of our practice is that, as these HSR-linked counties in our sample may be implicitly perceived as comparable in terms of economic importance when the HSR lines are designed and under operation, our finding on the distributional agglomeration of banks in large and small counties is less likely driven by the natural shift of financial resources from the economically peripheral cities to central cities.

Finally, we use a unique dataset that was manually collected for this study, distinguishing it from previous datasets in several ways. First, our data focus on the county-level cities in China, which are the lowest tier in the Chinese administrative hierarchy.[2] Our county-

---

[2] Chinese administrative hierarchy can be succinctly delineated as follows: nation – provinces (including autonomous regions, municipalities, and special administrative regions) – cities at the prefecture level – cities at the county level.

Table 1
Definition of variables and the main descriptive statistics.

Variables	Definition	Sources	N	Mean	Std. dev	Max	Min
Bank agglomeration							
BankNumber	The number of bank affiliates (in hundred) in sampled county-level cities	China Banking Regulatory Commission (CBRC)	5989	0.6986	0.6268	3.7300	0.0300
BankPopu	The number of bank affiliates scaled by the population in sampled county-level cities.		5989	0.1653	0.1326	1.1030	0.0074
High-speed rail							
HSR	The binary variable that is set equal to 1 since the following year when one (or more) HSR station is put into operation in the county, 0 otherwise.	The Ministry of Railways of China	5989	0.2371	0.4253	1	0
Control variables							
TownNumber	The number of administrative townships in each sampled county. The natural logarithm is applied to one plus the above indicator.	Chinese County Statistical Yearbooks, County Statistical Bulletins, City Statistical Yearbooks and Chinese Provincial Statistical Yearbooks	5989	2.3985	0.8445	4.3567	0
GeoSize	The geographic size of sampled counties (in square kilometers). The natural logarithm is applied to this variable.		5989	7.0105	1.0742	11.3120	2.8904
Population	The number of residents in sampled counties (in ten thousand). The natural logarithm is applied to this variable.		5989	4.0454	0.5626	6.1199	1.9459
GDP	The amount of gross domestic product in sampled counties (in hundred million). The natural logarithm is applied to this variable.		5989	4.7718	1.1393	9.2106	1.3218
FirmNumber	The number of large industrial enterprises in sampled counties. The natural logarithm is applied to this variable.		5989	4.6282	1.2369	8.1146	0
Payroll	The averaged payroll for employees in sampled counties (in ten thousand).		5989	2.9495	1.9352	11.8335	0.2594
PubFinance	The government fiscal revenue as a share of GDP in sampled counties.		5989	0.0538	0.0310	0.2062	0.0129
Education	The number of students in primary and middle schools in sampled counties (in ten thousand). The natural logarithm is applied to this variable.		5989	11.1612	0.6452	12.5665	9.3515

This table presents the definition of our main variables, the sources of data, and the main descriptive statistics of our major variables, including the mean, standard deviation, minimum, and maximum value.

level data allow us to examine the effects of HSR on economic activities in China from a micro-perspective, in contrast to some studies that rely on the data from higher levels of administrative hierarchy.[3] Second, we manually obtain the number of banks within the boundary of each county, which is used to construct our indicators of bank agglomeration, from the China Banking Regulatory Commission (CBRC). The county-level cities that are located along the HSR lines are identified by searching for information on operational HSR lines and their intermediate stops from the Ministry of Railways in China and its official ticketing website, www. 12306.com. Third, we match these county-level data with the information on some other potential determinants of banking agglomeration, which are manually extracted from various sources, including the Chinese County Statistical Yearbooks, County Statistical Bulletins, City Statistical Yearbooks, and Provincial Statistical Yearbooks.

Our main findings in this study, albeit solely grounded on Chinese sample, are believed to be a useful reference for many other emerging economies. In recent years, there has been a growing trend of HSR operation in numerous emerging economies.[4] In emerging economies with only an underdeveloped network of traditional transportation, the operation of HSR could have a greater marginal impact on the allocative efficiency of resources. This, in turn, is likely to promote economic prosperity in the HSR-connected cities, resulting in increased bank agglomeration. Meanwhile, due to the constraints of transportation in many emerging economies, in-habitants residing in peripheral areas have traditionally had to rely on local banks, which may possess substandard services and human resources, thereby falling short of meeting the heightened demand for superior financial services among residents. Just as it has been observed in China, HSR may have divergent effects on cities with differing levels of financial service quality. The HSR lines that link the

---

[3] For example, Cao et al. (2013) use data from 49 major cities in China to investigate the accessibility impact of HSR.
[4] HSR has been operating in Turkey, Uzbekistan, and Morocco, connecting their major cities such as Ankara-Istanbul, Tashkent-Bukhara, and Casablanca-Tangier. HSR lines are presently under construction in Thailand, Indonesia, and Saudi Arabia, or have been included in some countries' plan, e.g., India and Malaysia.

major cities with outlying regions may create a significant impetus for inhabitants of the latter to seek out more advantageous financial services in other cities. This may result in a positive (negative) impact on the bank agglomeration in cities with superior (inferior) financial services as well.

The remainder of this manuscript is delineated in the subsequent sections. Section 2 presents our hypotheses in this study, based on the extant related literature. Section 3 introduces our data and main variables, followed by their descriptive statistics. Section 4 presents our econometric model and methodology. We report our primary empirical results in Section 5 and perform some additional analyses in Section 6. Section 7 concludes.

2. Hypothesis development

We develop our hypotheses based on the extant literature about HSR's economic impact. Since its inception, there has been a growing body of literature on the economic implications of HSR. Numerous studies have examined whether HSR contributes to greater economic integration, given its potential to significantly enhance intra-regional mobility and promote the development of a larger functional area. For instance, Blum et al. (1997) discuss how HSR may facilitate regional integration by enabling a greater degree of market integration in the near term, incentivizing the reallocation of households and enterprises in the medium term, and shifting production and trade in the long term. Cheng et al. (2015) evaluate the impact of HSR on enhancing economic integration by promoting regional competitiveness and economic cohesion in China and the European Union, respectively.[5]

The increased economic integration resulting from the introduction of HSR can be attributed to the improved market accessibility that encourages production, investment, and migration among the cities and areas connected by HSR. Kim (2000) examines the effects of HSR on population density and employment prospects in the HSR-connected capital region in Korea. Wetwitoo and Kato (2017) provide supportive evidence that the prefectures connected to HSR in Japan exhibit higher levels of productivity compared to their counterparts without HSR. Carbo et al. (2019) document some evidence of a positive impact on the number of firms located in provinces along the Madrid-Barcelona HSR line. Ahlfeldt and Feddersen (2018) conduct a study on the output impact of the Cologne-Frankfurt HSR line in Germany. They compared the actual output of counties with intermediate HSR stops to their potential output in the absence of HSR and found that the actual output in these counties notably exceeded their counterfactual level.

Although numerous previous studies have demonstrated the positive effects of HSR on the real economy in the regions it serves, there is a dearth of research examining its impact on financial aspects.[6] As banks may need to broaden their reach to meet the increased demand for financial services resulting from the introduction of HSR, we propose our first hypothesis that the agglomeration of banks overall increases in HSR-linked cities.

Hypothesis 1. *The operation of HSR promotes the agglomeration of banks in the linked cities as a whole.*

Our research is also indebted to previous studies that have examined the redistributive effects of conventional transportation infrastructure, particularly the inter-region development gap, as economic activities can be more readily relocated from peripheral cities to core cities.[7] Nevertheless, there are surprisingly only a few works investigating whether HSR may could lead to a significant reallocation of economic activities and resources. Based on their counterfactual simulation experiment, Sasaki et al. (1997) argue that the expansion of the Shinkansen network does not necessarily lead to the regional diversion of economic activities and population. Garmendia et al. (2008), using the survey approach, find that HSR can assist remote cities to "acquire territorial roles of greater importance" by attracting both intra- and inter-provincial migration and familial investment. Li et al. (2016) suggest that HSR results in an increase in investment "influx" in the cities connected by HSR lines, while the cities without HSR access experience a decline in investment.[8] The reallocation of labor and capital is likely to cause discrepant results in terms of production output. Ke et al. (2017) find that HSR lines promote greater local GDP growth rates. However, this effect is more pronounced in cities that are highly industrialized, more advanced in the tertiary industry, or possess better infrastructure. The inconclusive outcomes of HSR in different cities or economic activities shed only a vague light on how financial resources, specifically banking resources, may be distributed in different cities.

We propose our second hypothesis that HSR may have diverse effects on the agglomeration of banks in large and small cities. Since the operation of HSR allows residents in small cities to access financial services in remote large cities more easily, it may incentivize banks to cluster more conspicuously in large cities. In other words, HSR could yield a positive impact on the presence of banks in large

---

[5] Vickerman (2015) is skeptical about the efficacy of HSR in economic integration, as he finds only limited evidence to suggest that the operation of HSR has a significant impact on the economic development of the intermediate areas between HSR-connected metropolises.

[6] An exception is the study conducted by Hernández and Jiménez (2014), who find that both fiscal revenue and budget gap improve with the operation of HSR, using the municipalities' data from Spain.

[7] For instance, Boarnet (1998) suggests that public infrastructure investment in one location may draw production from other locations. Baum-Snow (2007) finds supportive evidence that the expansion of the interstate highway system in the U.S. has a significant impact on the migration of the population toward suburban areas. In a related research, Baum-Snow et al. (2017) also find a "diffusion effect" that led to the dispersion of the population from Chinese central cities to the surrounding areas due to the expansion of the urban railway and highway network. Facilitated transportation is likely to promote labor and capital toward major cities, but it can impede the economic development of marginal cities, which experiences an outflow of resources. Faber (2014) investigates the influence of China's National Trunk Highway System and finds an adverse effect on the GDP growth rate of the linked peripheral counties. Qin (2017) finds that the enhancement of railway infrastructure leads to a decrease in production and per capita GDP in the counties that are impacted by the upgrade.

[8] Li et al. (2016) also find that consumption in HSR-linked cities "outflows" into cities without HSR.

cities but a negative impact in small cities.

Hypothesis 2. *The operation of HSR lines has a positive effect on the agglomeration of banks in large cities but a negative effect in small cities.*

3. Data and variables

We have selected the sample period as 2000–2017 for our study. Our data are collected from 498 county-level cities in China. Cities with higher administrative hierarchies, such as provincial capitals, autonomous regions, municipalities, and prefecture-level cities, are excluded since we aim to assess the effects of HSR from a relatively micro-perspective. In addition to counties, we also incorporate municipal districts (*shi xia qu*) in our sample, which were traditionally the satellite counties of major cities prior to the administrative reform of "transforming counties to districts", if there are HSR stations within their boundaries.

All of our sample counties were connected to operational HSR lines by 2017. Limiting our sample to the HSR-connected counties helps alleviate some concerns of endogeneity. As HSR lines are commonly posited linking counties with greater economic importance, our sample rules out the economically marginal counties. Therefore, the bank agglomeration amid the operation of HSR, as we reveal in this study, is less likely attributable to the inherent tendency of banking resources to shift from less economically prosperous areas to more developed ones, a phenomenon that could occur even in the absence of HSR.

We define HSR in this paper in a broad sense. Specifically, HSR is characterized as a rail system designed to operate at speeds of 200 km per hour or higher, intended for the transportation of either commuters or freight. In some inter-municipal district zones, although the actual running speed of HSR trains could be lower than 200 km/h, we still define these lines as HSR if only their designed velocity is at least 200 km/h.[9] A county is defined as being linked by HSR if there is at least one station located on HSR lines within its borders. We use a binary variable, denoted as *HSR*, to capture the period since the connection of counties to the Chinese HSR network. As there may be a lagged impact of HSR on the distribution of economic activities and resources, we let *HSR* equal to 1 in the following years when a county establishes a station on HSR lines and otherwise 0. The counties located on HSR lines are identified by searching for information on operating HSR lines and their intermediate stops from the Ministry of Railways in China and its official ticketing website, www. 12306.com.

We employ two variables to measure the degree of bank agglomeration. The first one, denoted as *BankNumber*, represents the number of bank affiliates (in hundreds), including subsidiaries, branches, and sub-branches within the administrative jurisdiction of each county.[10] Nevertheless, since counties with a larger population may have a higher number of bank affiliates, we introduce a second variable, *BankPopu*, which represents the number of bank affiliates scaled by the county's population size, to measure the per capita agglomeration of banks in each county. A larger number of affiliates of banks is perceived as indicating a denser cluster of banks, and therefore, greater accessibility to banking resources. We obtain the data on the number of bank affiliates from the China Banking Regulatory Commission (CBRC) by manually identifying whether the registered address of a bank affiliate is located within the geographic limits of a county.[11]

We control for a range of county-specific characteristics as possible factors influencing bank agglomeration. Firstly, we control for the influence of county size, as larger counties may have a greater capacity to attract bank affiliates to operate within their jurisdiction. We use three variables as proxies for the size of counties, specifically the number of townships administered by counties (*TownNumber*), the geographic area of counties (measured in square kilometers) (*GeoSize*), and the population size (in ten thousand) in counties (*Population*). If banks were to increase their presence with the size of the host counties, these variables would be positively associated with the number of bank affiliates. However, the expenses associated with expanding their physical presence may hinder banks from establishing affiliates that can effectively serve the counties they are targeting. Alternatively, they may opt to increase their workforce in each affiliate to cover a wider business area.[12] Additionally, the size of counties may be a less relevant factor in comparison to the economic might of counties when banks are deciding to establish their affiliates.

We next control for some variables that proxy the level of economic development in various counties. These variables include the amount of gross domestic product (in hundred million RMB) (*GDP*), the number of large industrial enterprises (*FirmNumber*),[13] the average payroll of employees (measured in ten thousand RMB) (*Payroll*), and the government's fiscal revenue as a share of GDP (*PubFinance*) in sample counties. These variables may have varying impacts on the agglomeration of banks. To be specific, higher output and a greater cluster of industrial enterprises may be indicative of higher demand for financial services, thereby leading to a greater presence of banks. A higher payroll in a city, on one hand, may attract a greater influx of human capital and encourage financial institutions to open additional offices. On the other hand, a higher payroll can also lead to increased labor costs for banks, which may have a negative impact on the number of affiliates that banks tend to establish (Shao et al., 2017). The ratio of fiscal revenue to local GDP may reflect the quality of infrastructure in our sample counties, given that infrastructure development is primarily funded by the fiscal revenue of local governments in China. Infrastructure can serve as a catalyst for the dissemination of resources, thereby incentivizing banks to establish operations in cities with better infrastructure quality.

---

[9] In our subsequent robustness checks, we restrict HSR to those with a designed velocity of 250 km/h and an actual operating speed of 200 km/h, specifically for commuter transportation.

[10] If a bank has its headquarters in one of the counties included in our sample, it is also considered as one of the "affiliates" of that bank.

[11] We have excluded policy banks and the Postal Savings Bank from our sample due to their non-profit nature.

[12] Unfortunately, information regarding the quantity of personnel in bank affiliates is unavailable.

[13] "Large industrial enterprises" refer to manufacturing firms that generate an annual operating revenue exceeding 20 million RMB in China.

Finally, we control for the number of students (in units of ten thousand) in primary and middle schools within our sample counties (*Education*) as a demographic variable that could potentially impact the allocation decisions made by banks. A county with a greater number of school-aged children is likely to possess a higher caliber of human capital, which may entice banks to open offices by capitalizing on the local human capital. Moreover, an increased enrollment of children in schools may indicate a relatively younger demographic that may demand more financial services compared to the elderly population. We have matched our data at the county level with the information of the above-introduced control variables, which were manually collected from various sources, including the Chinese County Statistical Yearbooks, County Statistical Bulletins, City Statistical Yearbooks, and Provincial Statistical Yearbooks. We have taken the natural logarithm of certain selective variables to mitigate the influence of their high values on the count of bank affiliates.[14] We provide a summary of the definition of our main variables and their data source in Table 1.

Table 1 also presents the major descriptive statistics of our variables, after we have ruled out the observations of our control variables that lie below the 1st percentile and above the 99th percentile in distribution. The agglomeration of banks, whether measured by the pure number of banks or the number of banks adjusted for county population, exhibits notable variation across counties. The mean of *BankNumber* (*BankPopu*) is 0.6986 (0.1653), and the standard deviation is 0.6268 (0.1326). The relatively large standard deviation, in comparison to the mean, suggests a broad range of our observations, as supported by the min-max range of [0.0300, 3.7300] ([0.0074, 1.1030]). In 2017, Yuwangtai, Henan had the lowest number of bank affiliates with only 5, whereas Yubei, Chongqing had the highest number of banks with 863 operating in the same year. In terms of the density of bank affiliates per capita, Tiexi county in Liaoning province had the highest agglomeration of banks in 2017, with 237 banks serving a population of 0.3 million. Meanwhile, the county with the lowest bank agglomeration is Yongqiao, located in the province of Anhui, with only 15 banks catering to a population of 1.9 million.

The mean value of *HSR*, the binary variable that indicates whether a county is connected to the network of HSR in China, is 0.2371. This suggests that in our sampled counties from 2000 to 2017, the county-year observations of HSR connection make up around 24% of all observations. We also note that counties exhibit significant variations in various aspects, such as their size, economic development, and size of school-going children. For instance, the mean value for the measure of the geographic size of counties is 7.0105, with a range from the minimum value of 2.8904 to the maximum value of 11.3120. Additionally, the highest value of the GDP measure in our sample is 9.2106, which is almost seven times greater than its lowest observation. As the variation in bank agglomeration may be driven by differences in county-specific characteristics, incorporating these variables into the analysis could alleviate the potential issues of omitted variables.

We also present the pairwise correlation of our main variables in the online appendix (refer to Table A2). The correlation between *HSR* and both *BankNumber* and *BankPopu* is positively and statistically significant, seemingly implying that bank agglomeration tends to increase with the expansion of HSR. Our indicators of bank agglomeration exhibit a positive correlation with population size, GDP, the number of large industrial enterprises, the mean level of employee payroll, the proportion of fiscal revenue to GDP, and the level of human capital. Conversely, they are negatively correlated with the number of townships and the geographic expanse of counties. HSR is significantly correlated with all the county feature variables. This indicates that the exclusion of these variables could potentially result in an overestimation or underestimation of the impact of HSR on the clustering of banks. The correlations between HSR and county characteristics are overall moderate, suggesting no serious multicollinearity problems when incorporating these variables jointly in our estimations.

4. Model

Our assessment of the impact of HSR on the agglomeration of banks unfolds in two steps. First, we investigate whether counties' connection to HSR lines promotes an overall increase in the number of bank affiliates, compared to the years preceding their HSR connection (Eq. 1). Next, we aim to detect if HSR yields distributive effects on big and small counties. In particular, having controlled for various other variables, we examine whether HSR may induce a positive impact on bank agglomeration in large counties, whereas a negative effect may occur in small counties (Eq. 2).

We use a difference-in-differences specification as below[15]:

$$Y_{it} = c + \beta \cdot HSR_{it} + \lambda \cdot X_{it} + County_i + Year_t + \varepsilon_{it} \tag{1}$$

$$Y_{it} = c + \beta \cdot HSR_{it} + \delta \cdot HSR_{it} * Big_i + \lambda \cdot X_{it} + County_i + Year_t + \varepsilon_{it} \tag{2}$$

where *i* and *t* refer to the county and year, respectively. *Y* denotes our measure of bank agglomeration, i.e., the number of banks in a county and the number of banks scaled by the size of population in the county. The variable *HSR* is binary and takes a value of 1 since year *t* + 1 if one (or more) HSR line passes through county *i* and at least one HSR station within the county's boundary becomes operational in year *t*. Otherwise, *HSR* takes a value of 0. *X* represents a set of county-specific time-varying feature variables, including the county's scale, level of economic development, and human capital, as introduced in the previous section. *County* represents the county-specific time-invariant factors that may affect the agglomeration of banks within that county. *Year* denotes the year-specific fixed effects, which may capture the nationwide shocks and the long-term trends that shape the clustering of banks over a period

---

[14] Certain variables are selected for smoothing through the use of natural logarithms, as they may capture the impact of county size.
[15] Our specification is similar to some earlier practices, such as Beck et al. (2010).

of time. c is the constant and ε is the idiosyncratic error term. In Eq. (1), a positive estimate of β is interpreted as an overall promotive effect of HSR that leads to a higher level of bank agglomeration in Chinese counties, as compared to the years prior to the operation of HSR in these counties.

In Eq. (2), we add an interactive term between *HSR* and the binary variable, *Big*, which differentiates larger counties from their smaller counterparts. We let *Big* be equal to 1 if the population of the county in 2017 exceeds the median value of the population distribution of the sampled counties in the same year.[16] In our robustness checks, we also experiment by defining a county as "big" based on its economic strength. That is, we let *Big* equal 1 if the county's GDP per capita in 2017 exceeds the median value of GDP per capita among the sample counties in the same year. As counties can be classified into large and small ones, the estimated coefficient β indicates the ceteris paribus effect of HSR on small counties, while the coefficient δ on the interactive term *HSR*Big* indicates how the effect of HSR on large counties differs from that on small counties. We anticipate that the coefficient δ may be positive, as the promotive impact of HSR in encouraging banks to establish affiliates may be more conspicuous in large counties. However, the coefficient β might be negative if HSR facilitates improved access to banking services for residents of small counties from larger, more distant cities. In other words, if the "push" effect, which diminishes the demand for financial services in small counties, outweighs the "pull" effect that entices banks to extend their operations into small counties, then it can be inferred that HSR has a detrimental impact on the bank agglomeration in small counties.

We employ the fixed-effects estimator in our examinations, which allows for correlations between the time-invariant county-specific variable and the other regressors. The Hausman test also suggests that the fixed-effects estimator is preferable to the random-effects estimator. We use robust standard errors that account for heteroskedasticity and within-panel serial correlation, and cluster standard errors at the county level during our estimations.

5. Results

5.1. The impact of HSR lines on the agglomeration of banks

We present the results of our estimations for Eqs. (1) and (2) in Table 2. In columns (1)–(3), our dependent variable is *BankNumber*, i.e. the number of bank affiliates in various counties. We use *BankPopu* as the dependent variable in columns (4)–(6), which represents the quantity of bank affiliates scaled by the population size of counties. In columns (1) and (4), we first include *HSR*, the binary indicator of whether a county is connected to the Chinese HSR network, as the regressor, along with the fixed effects of year and county. In columns (2) and (5), we also incorporate the covariates that measure the diversity in scale, economic might, and human capital of the counties. In columns (3) and (6), we include the interactive term *HSR*Big* to detect the distributional effects of HSR operation on the level of bank agglomeration in large and small counties.

We find that, firstly, *HSR* enters our regressions positively and significantly, at least at a 5% significance level, when the interaction term *HSR*Big* is not included. This suggests, on average, a promotive effect of the linkage to HSR lines on the agglomeration of banks in Chinese counties. As we note the number of bank affiliates in the hundreds, the estimated coefficient on *HSR* in column (2), which is 0.0678, suggests that, the average number of banks in a typical county tends to increase by 6.78 due to its connection with HSR lines. With a mean value of 69.86 for *BankNumber*, our estimate suggests that, after controlling for other factors, the presence of HSR lines is associated with a 9.7% increase in bank agglomeration in Chinese counties. Using *BankPopu* as the outcome variable also produce similar findings. The estimate of *HSR* in column (5), which is 0.0152, is interpreted as a rise of 1.52 percentage points in the ratio of the number of banks to the population of the county. This increase is equivalent to nearly a 9.2% rise from the average level of bank numbers adjusted for county residents, which is 0.1653. These results are consistent with *Hypothesis 1*.

Secondly, after incorporating the interactive term *HSR*Big* in our regressions, we find that the estimate for *HSR*Big* is positive and statistically significant, while the estimate for *HSR* turns out to be significantly negative. As the binary variable *Big* distinguishes our sampled counties into small and large ones, based on the median value in the distribution of county residents in 2017, the positive estimate on *HSR*Big* suggests that HSR facilitates the agglomeration of banks in larger counties. However, the negative estimate on the stand-alone term *HSR* suggests a contrary impact of HSR on the quantity of banks in smaller counties. That is, ceteris paribus, the level of bank agglomeration tends to decrease, rather than increase, in response to the operation of HSR. Quantitatively, when using *BankNumber* as the dependent variable, connection to HSR lines leads to an average increase of 12.5 banks in large counties,[17] while in small counties, the number of banks decreases by nearly 14.8. Our finding on the discriminatory impacts of HSR on bank agglomeration in large and small counties is consistent with *Hypothesis 2*.

Our finding is consistent with the "push" hypothesis, which proposes that HSR can reduce the cost for residents in small counties to access more favorable financial services from remote large counties, resulting in lower demand for local financial services. Although HSR may have a positive impact on productivity and economic growth in small counties, as some prior works have documented, and may attract bank affiliates to provide financial intermediation, this effect appears to be outweighed by the drain of local financial

---

[16] We conducted an experiment to classify counties as either large or small based on the median value of the sampled counties' population or GDP per capita in 2006, which was the year when HSR was introduced in China. We found that our results were qualitatively consistent. The results are available upon request.

[17] Use the estimates in column (3), Table 2, as an example. The effect of HSR on the quantity of banks in large counties is represented by $\beta + \delta = -0.1479 + 0.2730 = 0.1251$. This suggests that, on average, the number of banks in large counties tends to increase by 12.51 following their integration into the HSR network. The result for $\beta + \delta$ is statistically significant at the 1% level of significance. We report the *F*-statistic in Table 2.

Table 2

The impact of HSR on the agglomeration of banks.

Dependent variable						
	BankNumber			BankPopu		
	(1)	(2)	(3)	(4)	(5)	(6)
HSR	0.0627***	0.0678***	−0.1479***	0.0115**	0.0152***	−0.0252***
	(0.0223)	(0.0240)	(0.0291)	(0.0047)	(0.0051)	(0.0070)
HSR*Big			0.2730***			0.0511***
			(0.0335)			(0.0079)
TownNumber		−0.1180***	−0.1194***		−0.0244***	−0.0246***
		(0.0421)	(0.0389)		(0.0089)	(0.0083)
GeoSize		−0.0364	−0.0329		−0.0025	−0.0018
		(0.0580)	(0.0547)		(0.0122)	(0.0116)
Population		−0.1085	−0.1343		−0.0848***	−0.0897***
		(0.0792)	(0.0815)		(0.0218)	(0.0215)
GDP		−0.0509	−0.0339		−0.0072	−0.0040
		(0.0319)	(0.0291)		(0.0068)	(0.0063)
FirmNumber		0.0286	0.0259		0.0041	0.0036
		(0.0215)	(0.0209)		(0.0047)	(0.0046)
Payroll		0.0902***	0.0850***		0.0200***	0.0190***
		(0.0189)	(0.0185)		(0.0041)	(0.0040)
PubFinance		0.4733	0.6690*		0.1552*	0.1918**
		(0.3972)	(0.3799)		(0.0883)	(0.0859)
Education		0.3200***	0.3073***		0.0747***	0.0723***
		(0.0686)	(0.0664)		(0.0153)	(0.0148)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
N	5989	5989	5989	5989	5989	5989
Adj. R^2	0.4931	0.5487	0.5667	0.5227	0.5753	0.5878
F-statistic			21.18***			20.41***

This table reports the results of our baseline regressions. Our dependent variable is *BankNumber*, which measures the number of banks in county-level cities, and *BankPopu*, which measures the number of banks scaled by the population in counties. *HSR* is a binary variable that is equal to 1 since the next year when one (or more) HSR station is opened, and 0 otherwise. *Big* is a binary variable that is equal to 1(0) if a county is defined as a big (small) county according to the median value in the distribution of Chinese county-level cities' population in 2017. Among the other control variables, *TownNumber* measures the number of administrative townships in each sampled county. *GeoSize* represents the geographic size of sampled counties. *Population* is the number of residents in sampled counties. *GDP* measures the quantity of gross domestic product in sampled counties. *FirmNumber* gauges the number of large industrial enterprises in each sampled county. *Payroll* represents the average payroll for employees in sampled counties. *PubFinance* measures the government fiscal revenue as a share of GDP in sampled counties. *Education* captures the number of students in primary and middle schools in sampled counties. We also control the year-specific fixed effects (*Year FE*) and the county-specific fixed effects (*County FE*). Robust errors in parentheses. F-statistic reports the significance of the sum of the coefficient of *HSR* and that of *HSR*Big*. ***, ** and * denote the statistical significance on 1%, 5% and 10%, respectively.

demand from small counties caused by HSR. As a result, the increased demand for financial services in large counties creates a "pull" force that stimulates a greater clustering of banks in these areas.[18]

We checked the dynamics of the relationship between HSR connections and bank agglomeration to validate our DID framework since it takes time for banks to allocate their affiliates after the connection of counties with HSR lines. We revise our specification by incorporating a series of dummy variables to delineate the year-by-year effect of HSR:

$$Y_{it} = c + \sum_{\tau=1}^{5} \delta_{-\tau} \cdot HSR_{i,t-\tau} + \sum_{\tau=1}^{5} \delta_{+\tau} \cdot HSR_{i,t+\tau} + \lambda \cdot X_{it} + County_i + Year_t + \varepsilon_{it} \qquad (3)$$

where the dummy variables for HSR connection, the "HSRs," are equal to 0, except that $HSR_{t-\tau}$ equals 1 for counties in the τth year prior to their HSR connection, and $HSR_{t+\tau}$ equals 1 for counties in the τth year following the connection. We employ a 10-year window, spanning from 5 years before the HSR connection until 5 (or more) years after the connection. By excluding the current year of connection, our estimates on *HSRs* can be viewed as the dynamic impact of HSR on bank agglomeration in comparison to the year of connection. Our results are presented in Table 3. Still using *BankNumber* and *BankPopu* as the dependent variables, we find that the estimated coefficients for HSR_{t-5}, HSR_{t-4} … HSR_{t-1} are all statistically insignificant. This suggests that there is no apparent trend of bank

---

[18] We also find that the agglomeration of banks is affected by various other factors. The coefficient on the number of townships is significantly negative in our regressions, probably reflecting that the high brick-and-mortar expenses to open new affiliates may discourage banks from expanding their operations in counties with a higher number of townships. Alternatively, banks may distribute their operations across a broader region among their existing affiliates, rather than creating new physical locations. The mean employee salary, fiscal income as a percentage of GDP, and the ratio of school-aged children exhibit a positive association with bank agglomeration. This seemingly implies that banks are more likely to cluster in counties with higher household income, better infrastructure, and greater availability of human capital.

agglomeration in counties prior to their connection with HSR lines. In contrast, the estimates for HSR_{t+1}, HSR_{t+2} ... HSR_{t+5} are all positive and exhibit statistical significance in our tests. This indicates banks' tendency to adjust their presence in response to the operation of HSR in connected counties. Furthermore, the size of our estimates for HSR_{t+1}, HSR_{t+2}, ... HSR_{t+5} seemingly increases over time, suggesting a rising degree of bank clustering in counties linked to the HSR network. This trend may be attributable to banks having more time to reorganize the distribution of their branches.[19]

We also conducted a placebo test by assuming that the counties were connected to HSR lines a decade prior to the actual year of connection. For instance, if HSR started to operate in a county in 2013, we falsify this occurring in 2004. By restricting our dataset to 1999–2006, we re-estimate Eq. (1) and report the results in Table 4. Our earlier findings disappear in this placebo test as the estimated coefficients on *HSR* turn out to be negative and lack statistical significance. These findings suggest that there is no apparent trend in shaping the agglomeration of bank affiliates in Chinese counties before HSR lines go through.

5.1. The impact of HSR lines on financial intermediation and real economic activities

Next, having observed the competing effects of HSR on the cluster of banks in large and small counties, we assess whether the financial intermediation activities of banks, such as deposit collection and loan provision, would also be affected by the operation of HSR. One possibility is that, despite adjustments in terms of the number of affiliates in counties, banks may not alter the amount of financial resources available to their clientele, leading to a relatively modest impact on real economic activities. To examine this conjecture, we obtain the quantity of aggregate bank deposits and loans in each county and regress them on the same covariates as previously mentioned. Our results are reported in Table 5.

In Panel A of Table 5, we solely include *HSR* while excluding *HSR*Big* in our estimations. We find that *HSR* enters our regressions positively and significantly, suggesting an overall surge in banks' deposits and loans subsequent to the connection of counties to HSR lines. We further use three variables related to real economic activities, specifically the amount of fixed asset investment, the output of manufacturing industries, and the output of service industries as the dependent variables. We find that the estimated coefficients of *HSR* are also positive and statistically significant. These findings imply an increase in investment and production financed by banks' intermediation following the HSR connections in the counties.[20]

In Panel B of Table 5, we introduce the interactive term *HSR*Big* into our regressions. The resulting estimates allow us to investigate whether banking activities and real economic outcomes differ between large and small counties following their integration into the HSR network. We find consistent results that the coefficients for the stand-alone *HSR* are significantly negative, while those for the interaction term *HSR*Big* are significantly positive. We interpret our results as follows: in conjunction with the reallocation of their affiliate presence, banks also adjust the quantity of their financial services in favor of large counties. Deposits and loans of banks tend to rise in larger counties, leading to an increase in fixed assets and output as a result. However, the retreat of bank affiliates from smaller counties causes a reduction in the amount of bank funding available, leading to a contractionary impact on the real economy of these counties.[21] Our findings emphasize the need for caution regarding the potentially counterproductive effects of HSR connections, which may impact not only the banking market structure but also the real economy in smaller counties.

5.2. Endogeneity

Our previous estimation and analysis rest on the presumption that the HSR connection of the counties is not influenced by the level of bank agglomeration. However, it could be contended that HSR lines may first link the counties with higher bank agglomeration, thus causing a potential issue of reverse causality that could generate biased estimates. We present a diagram in Fig. 1 that illustrates the

---

[19] We conducted an experiment by including the interactive terms of $HSR_{t-\tau}$ and $HSR_{t+\tau}$ with *Big* (i.e., $HSR_{t-\tau}$*Big* and $HSR_{t+\tau}$*Big*) into the specification of Eq. (3). We find that the estimated coefficients for $HSR_{t+\tau}$*Big* remain significantly positive, and those for $HSR_{t+\tau}$ are still significantly negative. We view these results as consistent with our baseline finding that HSR lines are likely to result in increased (decreased) levels of bank agglomeration in large (small) counties. However, it is noteworthy that we also observe positive estimates for the coefficients on $HSR_{t-\tau}$*Big* and negative estimates for $HSR_{t-\tau}$. These results seemingly imply a rise (decline) in bank agglomeration in large (small) counties prior to their linkage to the HSR network. We conjecture two potential explanations for these results: Firstly, the trajectory of HSR routes is usually publicly disclosed during the planning phase, which allows banks to adjust their allocations years in advance of the completion of these routes. Secondly, the estimated coefficients on HSR_{t-5}*Big*, HSR_{t-4}*Big* ... HSR_{t-1}*Big*, HSR_{t+1}*Big*... HSR_{t+4}*Big*, and HSR_{t+5}*Big* exhibit an upward trend, likely because the linkage to HSR lines imposes an additional "push" impact on small counties. This effect drives residents of small counties to seek financial services in larger cities, leading to a more pronounced agglomeration of banks in large counties.

[20] Quantitatively, our estimates on *HSR* implies that the volume of bank deposits, bank loans, fixed asset investment, manufacturing industry output and tertiary industry output increase by 26.25 million RMB, 61.54 million RMB, 23.97 million RMB, 108.82 million RMB, and 23.59 million RMB, respectively, after a country has been connected to the HSR network.

[21] In big counties, bank deposits, bank loans, firms' fixed asset investment, the output of manufacturing industries, and the output of service industries increase by 49.01 million RMB, 102.40 million RMB, 52.54 million RMB, 169.99 million RMB, and 43.39 million RMB, respectively. However, in small counties, these factors decrease by 68.83 million RMB, 112.89 million RMB, 84.37 million RMB, 121.18 million RMB, and 50.26 million RMB, respectively.

Table 3
Dynamics of HSR connection and bank agglomeration.

Variables	BankNumber	BankPopu
	(1)	(2)
HSR_{t-5}	−0.0041	−0.0007
	(0.0101)	(0.0023)
HSR_{t-4}	−0.0152	−0.0032
	(0.0111)	(0.0025)
HSR_{t-3}	−0.0093	−0.0020
	(0.0121)	(0.0027)
HSR_{t-2}	−0.0133	−0.0034
	(0.0120)	(0.0027)
HSR_{t-1}	0.0083	0.0016
	(0.0128)	(0.0028)
HSR_{t+1}	0.0577***	0.0124***
	(0.0220)	(0.0047)
HSR_{t+2}	0.0938***	0.0217***
	(0.0294)	(0.0063)
HSR_{t+3}	0.1310***	0.0285***
	(0.0393)	(0.0083)
HSR_{t+4}	0.1237**	0.0268***
	(0.0479)	(0.0101)
HSR_{t+5}	0.1958***	0.0414***
	(0.0652)	(0.0136)
Control variables	Yes	Yes
Year FE	Yes	Yes
County FE	Yes	Yes
N	5989	5989
Adj. R^2	0.5530	0.5792

This table reports the results of HSR's time dynamic effects. $HSR_{t-\tau}$ represents the τth year before a county is connected to HSR lines, and $HSR_{t+\tau}$ is the τth year after the county is connected. We include the same county-specific characteristics in our estimation. We also control the year-specific fixed effects (*Year FE*) and the county-specific fixed effects (*County FE*). Robust errors in parentheses. ***, ** and * denote the statistical significance on 1%, 5% and 10%, respectively.

Table 4
A placebo test.

Dependent variable		
	BankNumber	BankPopu
	(1)	(2)
HSR	−0.0199	−0.0044
	(0.0121)	(0.0027)
Control variables	Yes	Yes
Year FE	Yes	Yes
County FE	Yes	Yes
N	2293	2293
Adj_R2	0.3126	0.3178

This table reports the results of HSR on the agglomeration of banks by assuming that the HSR lines were under operation 10 years earlier and using the data during 1999–2006. We include the same county-specific characteristics in our estimation. We also control the year-specific fixed effects (*Year FE*) and the county-specific fixed effects (*County FE*). Robust errors in parentheses. ***, ** and * denote the statistical significance on 1%, 5% and 10%, respectively.

year in which counties were connected to HSR, as well as the average number of banks, and the average ratio of bank number to county population in these counties prior to their HSR connection.[22] Our analysis does not reveal any discernible pattern between the timing of counties' HSR connection and the level of bank agglomeration. We also conducted a reverse causality test by regressing the binary variable HSR_{it} on our indicator of bank agglomeration (i.e., *BankNumber* or *BankPopu*) in year *t-1*. In other words, we assess whether the level of bank agglomeration in a given county may have a role in attracting HSR routes to traverse through that county. We have

---

[22] To provide greater clarity, we first regress our indicators of bank agglomeration on year dummies to remove any temporal trends. Subsequently, we illustrate the residuals of our regression in relation to the years of counties' HSR connections. Our approach is consistent with some previous studies (e.g. Beck et al., 2010).

Table 5
The impact of HSR on financial intermediation and economic activities.

Dependent variable					
	Bank deposits	Bank loans	Fixed asset investment	Manufacturing output	Service output
	(1)	(2)	(3)	(4)	(5)
Panel A					
HSR	0.2625***	0.6154***	0.2397***	1.0882***	0.2359***
	(0.0859)	(0.1916)	(0.0813)	(0.2954)	(0.0573)
Control variables	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
N	5138	5225	5970	5956	5957
Adj. R^2	0.6011	0.3309	0.5414	0.4105	0.5864
Panel B					
HSR	−0.6883***	−1.1289***	−0.8437***	−1.2118***	−0.5026***
	(0.1039)	(0.2227)	(0.1205)	(0.3281)	(0.0765)
HSR*Big	1.1784***	2.1529***	1.3691***	2.9117***	0.9365***
	(0.1125)	(0.3354)	(0.1449)	(0.3456)	(0.0913)
Control variables	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
N	5138	5225	5970	5956	5957
Adj. R^2	0.6292	0.3520	0.5620	0.4291	0.6172

This table reports the impact of HSR on various financial intermediation and economic activities. The dependent variables are bank deposits, bank loans, the amount of fixed assets investment, the output of manufacturing industries, and the output of service industries. In Panel A, we only include the variable *HSR* and the other control variables in our estimations. In Panel B, we add the interactive term *HSR*Big*. *HSR* is a binary variable that is equal to 1 since the next year when one (or more) HSR station is opened, and 0 otherwise. *Big* is a binary variable that is equal to 1(0) if a county is defined as a big (small) county. We include the same county-specific characteristics in our estimation. We also control the year-specific fixed effects (*Year FE*) and the county-specific fixed effects (*County FE*). Robust errors in parentheses. ***, ** and * denote the statistical significance on 1%, 5% and 10%, respectively.

found no plausible evidence to suggest that bank agglomeration has any significant impact on the likelihood of a county being connected to the HSR network. The coefficients for *BankNumber* and *BankPopu*, as well as their interactive terms with *Big*, are statistically insignificant in all regressions (see Table A3 in the online appendix), seemingly implying that the level of bank agglomeration in a county is less likely to have a significant impact on its likelihood of being selected for the HSR network.

Additionally, we employ the IV-2SLS approach to mitigate potential endogeneity concerns. We employ the count of HSR stations in the adjacent counties of the focal county as the instrumental variable. We first regress *HSR* and *HSR*Big* on our instrumental variable and other covariates in our first-stage estimation, and then replace *HSR* and *HSR*Big* with their predicted values in the second-stage regression. Our results are presented in Table 6, where Panel A reports the second-stage results. Consistent with our baseline finding, we still find that HSR may have diverse effects on counties of different sizes. When we include only *HSR* in covariates, the estimated coefficient is significantly positive, indicating an overall promotive effect of HSR lines on the agglomeration of banks after the counties are connected by HSR. Upon adding the interactive term *HSR*Big* to the regressions, we find that the coefficients on *HSR*Big* remain positive and highly significant. This suggests that the agglomeration of banks tends to increase in larger counties. In contrast, the estimated coefficient for the stand-alone term of *HSR* is significantly negative, consistent with our benchmark result that the presence of HSR has a ceteris paribus negative impact on banking agglomeration in smaller counties.

In the first-stage regression (Panel B, Table 6), the estimates on our instrumental variable are positive and statistically significant, indicating that our instrumental variable is significantly associated with the presumed endogenous variables.[23] We present the *F*-statistics of the Cragg-Donald Wald test, as recommended by Stock and Yogo (2005), to assess the weakness of instrumental variables. We find that the *F*-statistics in our tests are considerably greater than the critical values (at the 5% significance level) suggested by Stock and Yogo (2005). This leads us to reject the hypothesis that our instruments are weak. We also check the potential endogeneity of counties' HSR-connection status by performing the Durbin-Wu-Hausman test. We find that the outcomes are statistically significant, thereby rejecting the hypothesis that the HSR-connection status of counties may be exogenous. However, it should be noted that the results of our Durbin-Wu-Hausman test are only marginally significant in all of the tests. Hence, we view them as only limited evidence for the endogeneity of counties' HSR-connection status.

We have also employed the IV-2SLS methodology to estimate the impact of HSR lines on financial intermediation and real economic activities. As previously stated, we use the quantity of total bank deposits and loans in each county as indicators of financial

---

[23] When we only include *HSR* in our specification, the dependent variable in our first-stage regression is correspondingly *HSR*. However, it is important to note that when we include both *HSR* and *HSR*Big* in our specification, we run *separate* regressions for these two variables on our IV in the first-stage. This allows us to obtain two estimated coefficients, which are presented in Panel B of Table 6.

Panel A
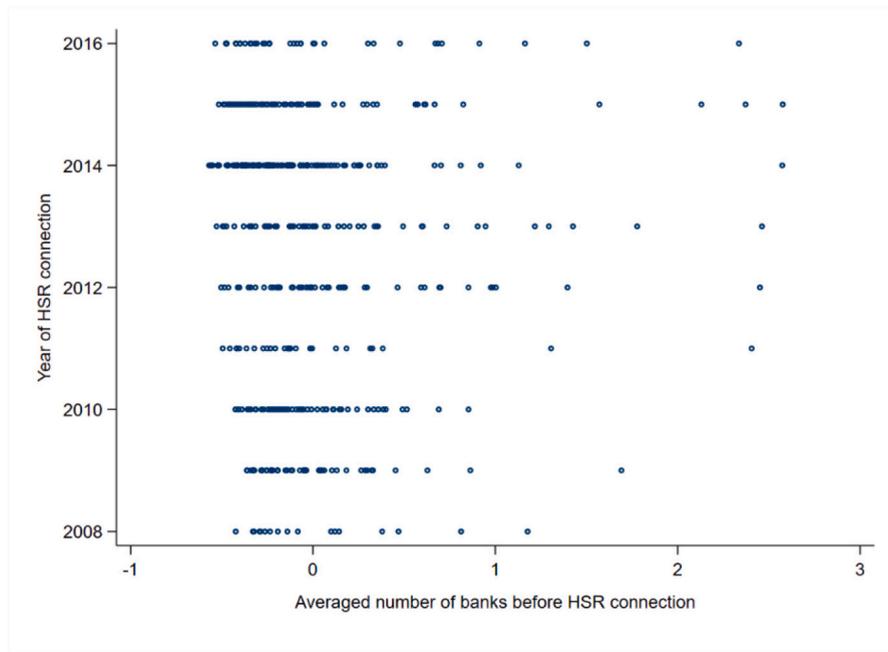
Panel B
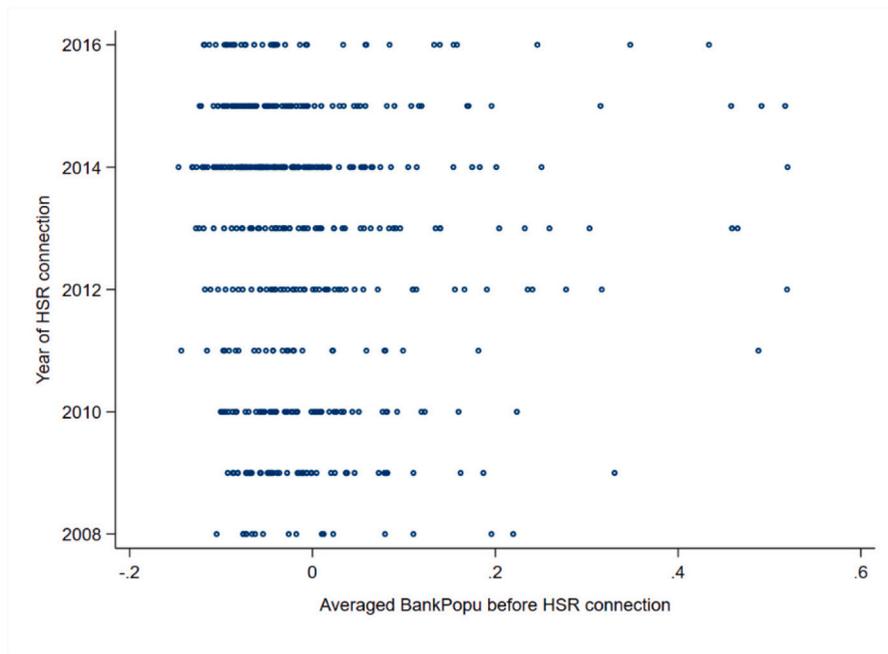
Fig. 1. The time of HSR connection and the agglomeration of banks.

In Panel A, the figure shows a scatter plot of the average number of banks in counties before their connection with HSR lines and the year of their connection. In Panel B, the figure shows a scatter plot of the average ratio of the number of banks to county population before the counties' connection with HSR lines and the year of their connection. The *t*-statistics for the association in the figures of Panel A and B are −0.64 and − 0.44, respectively.

Table 6

The impact of HSR on the agglomeration of banks: IV-2SLS estimation.

	BankNumber		BankPopu	
	(1)	(2)	(3)	(4)
Panel A:Second stage				
HSR	0.1295***	−0.1176***	0.0279***	−0.0187**
	(0.0378)	(0.0368)	(0.0082)	(0.0087)
HSR*Big		0.2975***		0.0561***
		(0.0334)		(0.0078)
Control variables	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
N	5980	5980	5980	5980
Adj. R^2	0.5096	0.5296	0.5387	0.5526
Panel B:First Stage				
The coefficient of IV on HSR	0.2885***	0.2840***	0.2885***	0.2840***
	(0.0131)	(0.0229)	(0.0131)	(0.0229)
The coefficient of IV on HSR*Big HSR*Big		0.4428***		0.4428***
		(0.0136)		(0.0136)
Cragg-Donald	4160.473	2048.531	4160.473	2048.531
Durbin-Wu-Hausman	17.07	17.41	15.33	15.92

This table reports our result when using the IV-2SLS estimator. The dependent variable is *BankNumber* in column (1)–(2) and *BankPopu* in column (3)–(4), respectively. We use the number of HSR stations in the neighboring counties of focal counties as the instrumental variable. In Panel A, we report the second-stage results, and in Panel B the results of the first stage. *HSR* is a binary variable that is equal to 1 since the next year when one (or more) HSR station is opened, and 0 otherwise. *Big* is a binary variable that is equal to 1(0) if a county is defined as a big (small) county. We include the same county-specific characteristics in our estimation. We also control the year-specific fixed effects (*Year FE*) and the county-specific fixed effects (*County FE*). *Cragg-Donald* is the *F* statistics of the Cragg-Donald Wald test as suggested by Stock and Yogo (2005), and *Durbin-Wu-Hausman* reports the statistic of the Durbin-Wu-Hausman test. Robust errors in parentheses. ***, ** and * denote the statistical significance on 1%, 5% and 10%, respectively.

intermediation. We employ the amount of fixed asset investment, the output of the manufacturing sector, and the output of the service sector to gauge the activities of real economy. The results are reported in Table 7. In line with our previous results, we find that the integration of a county into the HSR network overall creates a positive impact on both financial intermediation and real economy. When including only *HSR* in our covariates, we find its estimated coefficient to be positive and statistically significant in all regressions. However, the differential impacts of HSR lines on counties of varying sizes become apparent when we introduce the interactive term *HSR*Big* into our estimations. We find again that the coefficient on *HSR*Big* is significantly positive. In contrast, the estimate for *HSR* is consistently negative and significant across all regressions. We interpret our findings as compelling evidence of the divergent effects of HSR lines on both large and small counties. The quantity of banks' deposits and loans tends to augment in large counties, while it declines in their smaller counterparts. As banks play a dominant role in funding the investment and production of enterprises, the latter two also exhibit a contrasting trend in large and small counties: Investment and output in large counties increase with the operation of HSR, while they decline in small counties. We perceive our above-mentioned results as favorable evidence that aligns with our baseline finding, once the endogeneity concerns have been ameliorated.

5.3. Robustness checks

We conduct a series of robustness checks for our baseline result in this subsection. First, we expand our sample by including some counties that remain unaffected by HSR lines, specifically those counties that are never connected to the HSR network in China. However, given the substantial variation across counties in terms of economic and demographic characteristics, it is likely that these factors influence the level of bank agglomeration and introduce potential biases in our estimation. To alleviate this concern, we exclusively incorporate "comparable" counties into our sample by using the propensity score matching technique to match each county linked to the HSR network with a counterpart that is not. To be precise, we perform a logistic regression analysis using the (one-year lagged) observations of three variables, namely counties' GDP, population, and fiscal revenue, to assess the odds of a county being linked to the HSR network. After applying the "one-to-one" matching ratio, we then employ the caliper matching algorithm, which sets the maximum permissible distance between propensity scores at 0.01, to identify a "comparable" county that is disconnected to the HSR for each of the HSR-connected counties. A total of 443 counties along the HSR lines have been successfully matched with an equal number of counties that are not connected to the HSR network.[24]

After including the counties that are connected to the HSR and those that are not in our sample, we have re-estimated our model

---

[24] It is noteworthy that the difficulty of matching increases with the number of county characteristics under consideration. The greater the number of variables we choose, the fewer counties that are disconnected from HSR we can recognize as "comparable" to the counties connected to HSR. Due to this trade-off, we have opted to select the aforementioned three county characteristics (i.e., counties' GDP, population, and fiscal revenue) to ensure a sufficient number of counties to be matched.

Table 7

The impact of HSR on financial intermediation and real economic activities: IV-2SLS estimation.

	Bank deposits		Bank loans		Fixed asset investment		Manufacturing output		Service output	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A:Second stage										
HSR	0.4645***	−0.5870***	0.8748**	−1.0617***	0.4958***	−0.7415***	2.6522***	−0.0200	0.3494***	−0.4671***
	(0.1589)	(0.1392)	(0.3626)	(0.3876)	(0.1796)	(0.1819)	(0.5624)	(0.4887)	(0.1286)	(0.1080)
HSR*Big		1.2787***		2.3332***		1.4879***		3.2154***		0.9829***
		(0.1224)		(0.3452)		(0.1543)		(0.4275)		(0.0965)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	5130	5130	5213	5213	5962	5962	5945	5945	5949	5949
Adj. R^2	0.5681	0.5985	0.2770	0.2998	0.5020	0.5248	0.3505	0.3722	0.5514	0.5852
Panel B:First Stage										
The coefficient of IV on HSR	0.2949***	0.2836***	0.2959***	0.2854***	0.2885***	0.2830***	0.2886***	0.2840***	0.2898***	0.2850***
	(0.0146)	(0.0244)	(0.0143)	(0.0244)	(0.0131)	(0.0229)	(0.0131)	(0.0229)	(0.0131)	(0.0229)
The coefficient of IV on HSR*Big		0.4410***		0.4400***		0.4426***		0.4427***		0.4427***
		(0.0145)		(0.0142)		(0.0136)		(0.0136)		(0.0136)
Cragg-Donald	4075.414	2017.088	4243.913	2096.426	4141.163	2037.448	4138.159	2036.412	4230.473	2078.785
Durbin-Wu-Hausman	17.45	24.59	4.09	6.46	13.21	13.99	75.57	74.89	9.28	7.31

This table reports the impact of HSR on various financial intermediation and economic activities. The dependent variables are bank deposits, bank loans, the amount of fixed assets investment, the output of manufacturing industries, and the output of service industries. In Panel A, we report the second-stage results, and in Panel B the results of the first stage. *HSR* is a binary variable that is equal to 1 since the next year when one (or more) HSR station is opened, and 0 otherwise. *Big* is a binary variable that is equal to 1(0) if a county is defined as a big (small) county. We include the same county-specific characteristics in our estimation. We also control the year-specific fixed effects (*Year FE*) and the county-specific fixed effects (*County FE*). *Cragg-Donald* is the *F* statistics of the Cragg-Donald Wald test as suggested by Stock and Yogo (2005), and *Durbin-Wu-Hausman* reports the statistic of the Durbin-Wu-Hausman test. Robust errors in parentheses. ***, ** and * denote the statistical significance on 1%, 5% and 10%, respectively.

Table 8

The impact of HSR on bank agglomeration: Adding the sample of comparable disconnected counties.

Dependent variable	BankNumber		BankPopu	
	(1)	(2)	(3)	(4)
HSR	0.0401*	−0.0674**	0.0088*	−0.0165**
	(0.0240)	(0.0296)	(0.0052)	(0.0082)
HSR*Big		0.1348***		0.0326***
		(0.0432)		(0.0106)
Control variables	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
N	9900	9900	9900	9900
Adj. R^2	0.5312	0.5450	0.5556	0.5635

This table reports our results on the impact of HSR on bank agglomeration when we add non-HSR-connected counties into our sample. Our dependent variable is *BankNumber*, which measures the number of banks in county-level cities, and *BankPopu*, which measures the number of banks scaled by the population in counties. *HSR* is a binary variable that is equal to 1 since the next year when a county is connected with the HSR lines, and 0 otherwise. *Big* is a binary variable that is equal to 1(0) if a county is defined as a big (small) county. We include the same county-specific characteristics in our estimation. We also control the year-specific fixed effects (*Year FE*) and the county-specific fixed effects (*County FE*). Robust errors in parentheses. F-statistic reports the significance of the sum of the coefficient of *HSR* and that of *HSR*Big*. ***, ** and * denote the statistical significance on 1%, 5% and 10%, respectively.

and presented the outcomes in Table 8. Our results remain consistent as we still find that, by using either the number of banks or the number of banks scaled by county population as the dependent variable, the estimated coefficient for *HSR* is significantly positive when we exclude the interactive term *HSR*Big* in our regressions. However, when adding this interactive term into our estimation, the coefficient on *HSR* turns out to be significantly negative, while that on *HSR*Big* is highly positive. Our above practice confirms our baseline finding that, although having an overall positive impact on the agglomeration of banks in Chinese counties, the operation of HSR promotes a more pronounced clustering of banks in larger counties but leads to a decrease of bank agglomeration in smaller counties.[25]

Second, we propose a narrower definition of HSR lines, specifically those with a designed velocity exceeding 250 km/h, an actual operating speed of over 200 km/h, and intended solely for commuter transportation. In other words, we exclude the HSR lines that have a lower designed speed, those that operate de facto at a speed below 200 km/h (especially the intra-city lines), and those intended solely for freight transportation. We re-estimate Eqs. (1) and (2) and report our results in the online appendix (refer to Table A4). We find that our results are consistent with those before, indicating that the implementation of HSR systems results in varying levels of bank agglomeration in both large and small counties.

Third, our prior practice categorizes counties as either large or small based on their population size. However, it is plausible to argue that banks may prioritize allocating their affiliates and resources to economically developed counties, rather than those with high population density. To check if our earlier results overstate the motivation of banks to distribute their reach in conjunction with the proliferation of HSR lines, we alternatively define "large" ("small") counties as those with a GDP per capita above (below) the median level of distribution in 2017. We report our results in the online appendix (refer to Table A5). We still find consistent results as before. The interactive term *HSR*Big* enters our regressions positively and significantly, suggesting again that the agglomeration of banks augments in counties with higher GDP per capita after their integration into the Chinese HSR network. The estimate on the stand-alone term *HSR* is negative and statistically significant, indicating again a counterproductive impact of HSR on the number of banks in counties with lower GDP per head. Our results confirm the competing effects of HSR on promoting banks to allocate their affiliates and resources among different counties. Notably, the coefficients for the stand-alone *HSR* are notably less negative than their baseline equivalents. This finding probably implies a weaker impact of HSR in inducing a withdrawal of banks from less prosperous counties. This is likely due to the fact that banks must sustain their presence to ensure the availability of fundamental financial services, even in underdeveloped counties that are economically underdeveloped.

6. Some extended analyses

In this section, we extend our analysis to several dimensions. Firstly, we investigate whether our findings may differ if we take into account counties of a "medium" size. We use the 25th and 75th percentiles of the population distribution of all counties to classify our sampled counties into three categories: large, medium, and small. We next re-estimate Eq. (2) by using different combinations of sampled counties. For the sake of conciseness, the findings are presented in the in the online appendix (refer to Table A6).

We still use *BankNumber* and *BankPopu* as the dependent variables, respectively, in columns (1) through (4) and columns (5) to (8). In columns (1) and (5), we use the same sample as before, which include the big, medium-sized, and small counties. We limit our sample to exclusively big and medium-sized counties in columns (2) and (6), whereas we solely incorporate large and small counties in

---

[25] We highly appreciate the suggestion of a subject editor to use a matched sample to check our baseline findings.

columns (3) and (7). In the above practice, we let the binary variable *Big* be equal to 1 for counties with a population greater than the 75th percentile in the distribution of counties' population in 2017, and 0 for all other counties. Finally, we only use the medium-sized and small counties in columns (4) and (8), and the binary variable *Big* correspondingly is set to 1 to represent the medium-sized counties and 0 for small counties. We find that the estimates on *HSR*Big* are positive, while the estimates on *HSR* are consistently negative. These findings indicate once more that HSR stimulates a more conspicuous promotive impact on the agglomeration of banks in relatively larger counties, whereas it has the opposite effect in their smaller ones.[26]

Secondly, it could be argued that banks may adjust the allocation of their affiliates across counties within the same prefecture. That is to say, banks may expand their presence in larger counties while decreasing it in smaller ones, yet still maintain the same number of affiliates within the prefecture. We therefore resort to the data of prefecture-level cities to check if our findings would be altered. We calculate the number of banks per prefecture, and scale it by the population size of each prefecture. These values are then used as our dependent variables. Similar to our earlier practices, the dummy variable *Big* is used to denote the prefectures with a population more than the median value in the distribution of all prefectures' populations in 2017. Our results based on the prefecture-level cities are reported in the online appendix (see Table A7), in which we find the coefficient on *HSR*Big* is significantly positive, but that on *HSR* is significantly negative again. Our results suggest that the reallocation of bank affiliates and resources occurs not only across counties but also across prefectures. In other words, the operation of HSR may allow residents to access financial services outside of their prefecture, leading banks to redistribute their presence across prefectures.

We next test whether the reallocation of bank affiliates can be even region-wise. In other words, our focus is on identifying whether HSR would have varying effects on bank agglomeration across different blocs of provinces and municipalities. Following the prevailing approach in related literature, we categorize the counties in our sample into two groups: those situated in the eastern region of China and those located in the central and western regions.[27] Accordingly, we construct a binary variable *East*, which takes a value of 1 if a county is located in an eastern province or municipality of China, and 0 otherwise. To investigate whether the impacts of HSR on counties in the eastern/central and western regions of China differ, we interact *HSR* with *East* and use it to replace our original interaction term, *HSR*Big*. We then reconduct our estimations and report the results in the online appendix (refer to Table A8). When using *BankNumber* as the dependent variable, we first use only the sampled counties from the east of China and those from the middle and west, respectively, in columns (1) and (2). We next use all of our sampled counties and include the interactive term *HSR*East* into the regression in column (3), where the stand-alone term *HSR* captures the impact of HSR connections on the bank agglomeration in the counties of the middle and west of China, and *HSR*East* is interpreted as the impact in the counties in the east of China.

In column (1), we note a significantly positive estimate of *HSR*, indicating that HSR promotes a greater concentration of banks in a typical eastern county. In contrast, the estimate of *HSR* in column (2) is negative, when using only the counties from the central and western regions of China as the sample. However, this result is not statistically significant. This finding suggests that in regions of China outside of the east, the agglomeration of banks tends to decrease as counties become linked to HSR lines. In column (3), using the counties from both the eastern, middle, and western parts of China, we find that the estimates on *HSR* and *HSR*East* are significantly negative and positive, respectively, in our regression. This presents some suggestive evidence that banks may cluster more heavily in the eastern region of China while decreasing their presence in other regions. As we use *BankPopu* as the dependent variable in columns (4)–(6), we find only consistent results. Our results, although not conclusive, support the conjecture that banks may relocate their affiliates across regions, resulting in varying effects on the agglomeration of banks in counties situated in different regions of China. In terms of the agglomeration of banking resources, it appears that the eastern region of China is comparatively benefited than its central and western counterparts.

Finally, we assess the influence of HSR on the agglomeration of different types of banks. The incentive for banks to reallocate their affiliates may vary depending on their type for at least three reasons. First, the banks with monopolistic advantages probably have only a weak incentive to shift their branches to more competitive markets, while the banks with competitive disadvantages may be more likely to "follow their customers" by increasing their presence in larger counties when their clients seek financial services from remote large counties. Second, it may be more difficult for banks with greater social responsibilities, such as the state-owned banks, to relocate their affiliates even in face of declining demand, since they must ensure the provision of essential financial services in local areas. Third, banks may encounter varying degrees of budgetary limitations or implicit geographical constraints that impede their capacity to establish a greater number of branches.

We categorize the sampled banks into five groups: state-owned banks, joint-stock commercial banks, city commercial banks, rural commercial banks, and foreign banks.[28] Like our prior practice, the number of each type of banks and that scaled by the size of county residents are used as two proxies of agglomeration for each type of banks. As our interest lies in whether different types of bank may alter their presence differently across large and small counties, we only report the results based on the estimation of Eq. (2) in the

---

[26] By defining the size of counties alternatively using the 25th and 75th percentiles in the distribution of all counties' GDP per capita, we obtain results that are consistent with our earlier findings.

[27] Based on the classification by the National Development and Reform Commission in China, the provinces and municipalities that belong to the eastern region are: Beijing, Fujian, Guangdong, Guangxi, Hainan, Hebei, Jiangsu, Liaoning, Shandong, Shanghai, Tianjin, and Zhejiang. The provinces and municipalities located in the central and western regions of China comprise Anhui, Chongqing, Gansu, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangxi, Jilin, Qinghai, Shaanxi, Shanxi, Sichuan, Xinjiang, and Yunnan. During our sample period, three additional provinces, namely Inner Mongolia, Ningxia, and Xizang, were not integrated into China's HSR network.

[28] The category of rural commercial banks encompasses rural credit unions, rural cooperative banks, and village banks. Private banks are excluded as they rarely have affiliates in counties.

online appendix (refer to Table A9).

We first use *BankNumber* as the outcome variable. We observe that, statistically significant in most cases, the estimated coefficients on *HSR* are negative, while those on *HSR*Big* are positive for all categories of banks. These findings indicate a shared motivation for all types of banks to expand their affiliate presence in large counties, while comparatively decreasing it in small counties. However, there are notable variations in the magnitude of our estimates among various categories of banks. With the exception of foreign-owned banks, whose presence in both small and large counties is barely affected by HSR, state-owned banks exhibit the lowest coefficient on *HSR* (−0.0159) compared to other bank groups. This suggests that, on average, the number of affiliates of state-owned banks decreases by 1.59 in small counties following their connection to the Chinese HSR network. In comparison, the *HSR* coefficient for rural commercial banks is −0.0714, indicating that the quantity of rural commercial banks is likely to decrease by 7.14 in a typical small county. Either because of the market power of state-owned banks, which allows them to enjoy higher monopolistic returns in small counties, or their social responsibility to maintain basic financial services, they may have only a lower incentive to shrink their presence in small counties. In contrast, the competitive disadvantages associated with rural commercial banks may make it difficult for them to "lock in" their clients when the latter have lower accessibility costs to opt for remote banks. This, in turn, creates a stronger incentive for rural commercial banks to downsize their affiliates in small counties. Meanwhile, rural commercial banks are found to increase their presence in large counties more greatly in comparison to other types of banks, as their estimated coefficient on *HSR*Big* is 0.1206, which is interpreted as meaning that the number of rural commercial banks' affiliates in a large county, on average, increases by 4.92 since the county is connected by HSR lines. The coefficient of *HSR*Big* for city commercial banks is only 0.0268. This implies that the number of their affiliates in a large county would increase by only 0.89 after being connected to HSR lines. This is likely a result of their budgetary constraints or implicit geographical limitations that impede the expansion of their branches. As we use *BankPopu* as our dependent variable, our findings remain consistent.

7. Conclusion

Amid the proliferation of HSR in numerous countries, there have been some studies associated with the economic outcomes of HSR, while how HSR shapes the distribution of economic resources is still underexplored. This paper makes a supplementary contribution to the extant literature by investigating the effect of HSR on the agglomeration of bank affiliates. We use the data from county-level cities in China, which currently boasts the largest network of HSR lines in the world. Our research may shed some light on how economic prosperity in the county-level cities of China may be influenced by the operation of HSR, particularly through the availability of banking resources.

Our baseline results suggest an overall promotive impact of counties' connection with HSR lines on the agglomeration of banks. However, when we differentiate counties based on their size, we observe consistent evidence that HSR has contrasting effects in these two categories of counties. HSR has a positive impact on the agglomeration of banks in large counties by encouraging a greater number of bank affiliates to operate in these areas. However, ceteris paribus, HSR has a negative effect on the number of banks in small counties. Our findings are consistent with the hypothesis that the lower accessibility cost owing to HSR transport "pushes" banks to reduce their presence in small counties, as their residents can access more favorable financial services in remote cities more easily, but "pulls" banks to trace the shift of financial demand by increasing their cluster in large counties.

Furthermore, as we investigate the impact of HSR on the financial activities of banks and various real economic indicators, we also observe mixed results regarding the impact of HSR connections in both large and small counties. This implies that banks not only alter their presence with the operation of HSR, but also change the availability of banking resources, which may thereby reshape some real economic activities and the prosperity of counties. This finding highlights the need for caution regarding the counterproductive effects of HSR, in addition to its benefits related to accelerated transportation and lowered accessibility costs.

The statement by authors

This work has not been published previously and is not under consideration for publication elsewhere. And its publication is approved by all authors.

Data availability

Data will be made available on request.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ememar.2023.101047.

References

Ahlfeldt, G.M., Feddersen, A., 2018. From periphery to core: measuring agglomeration effects using high-speed rail. J. Econ. Geogr. 18, 355–390.

Amos, P., Bullock, D., Sondhi, J., 2010. High-Speed Rail: The Fast Track to Economic Development? World Bank Working Paper.

Baum-Snow, N., 2007. Did highways cause suburbanization? Q. J. Econ. 122, 775–805.

Baum-Snow, N., Brandt, L., Henderson, J.V., Turner, M.A., Zhang, Q., 2017. Roads, railroads, and decentralization of Chinese cities. Rev. Econ. Stat. 99, 435–448.

Beck, T., Levine, R., Levkov, A., 2010. Big bad banks? The winners and losers from bank deregulation in the United States. J. Financ. 65, 1637–1667.

Blum, U., Haynes, K.E., Karlsson, C., 1997. The regional and urban effects of high-speed trains. Ann. Reg. Sci. 31, 1–20.

Boarnet, M.G., 1998. Spillovers and the locational effects of public infrastructure. J. Reg. Sci. 38, 381–400.

Cao, J., Liu, X.C., Wang, Y., Li, Q., 2013. Accessibility impacts of China's high-speed rail network. J. Transp. Geogr. 28, 12–21.

Carbo, J.M., Graham, D.J., Anupriya Casas, D., Melo, P.C., 2019. Evaluating the causal economic impacts of transport investments: Evidence from the Madrid–Barcelona high speed rail corridor. J. Appl. Stat. 46, 1714–1723.

Chen, Z., Haynes, K.E., 2015. Impact of high speed rail on housing values: an observation from the Beijing-Shanghai line. J. Transp. Geogr. 43, 91–100.

Chen, C.-L., Vickerman, R., 2019. Quantifying the economic and social impacts of high-speed rail: some evidence from Europe and the People's Republic of China. In: ADBI Working Paper Series No. 962.

Cheng, Y., Loo, B.P.Y., Vickerman, R., 2015. High-speed rail networks, economic integration and regional specialisation in China and Europe. Travel Behav. Soc. 2, 1–14.

Faber, B., 2014. Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System. Rev. Econ. Stud. 81, 1046–1070.

Garmendia, M., de Ureña, J.M., Ribalaygua, C., Leal, J., Coronado, J.M., 2008. Urban residential development in isolated small cities that are partially integrated in metropolitan areas by high speed train. Eur. Urban Regional Stud. 15, 249–264.

Givoni, M., 2006. Development and impact of the modern high-speed train: a review. Transp. Rev. 26, 593–611.

Guirao, B., Lara-Galera, A., Luis Campa, J., 2017. High speed rail commuting impacts on labour migration: the case of the concentration of metropolis in the Madrid functional area. Land Use Policy 66, 131–140.

Hernández, A., Jiménez, J.L., 2014. Does high-speed rail generate spillovers on local budgets? Transp. Policy 35, 211–219.

Ke, X., Chen, H., Hong, Y., Hsiao, C., 2017. Do China's high-speed-rail projects promote local economy? - new evidence from a panel data approach. China Econ. Rev. 44, 203–226.

Kim, K.S., 2000. High-speed rail developments and spatial restructuring: a case study of the capital region in South Korea. Cities 17, 251–262.

Li, X., Huang, B., Li, R., Zhang, Y., 2016. Exploring the impact of high speed railways on the spatial redistribution of economic activities - Yangtze River Delta urban agglomeration as a case study. J. Transp. Geogr. 57, 194–206.

Qin, Y., 2017. 'No county left behind?' The distributional impact of high-speed rail upgrades in China. J. Econ. Geogr. 17, 489–520.

Sasaki, K., Ohashi, T., Ando, A., 1997. High-speed rail transit impact on regional systems: does the Shinkansen contribute to dispersion? Ann. Reg. Sci. 31, 77–98.

Shao, S., Tian, Z.H., Yang, L.L., 2017. High speed rail and urban service industry agglomeration: evidence from China's Yangtze River Delta region. J. Transp. Geogr. 64, 174–183.

Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Andrews DWK Identification and Inference for Econometric Models. Cambridge University Press, New York, pp. 80–108.

Vickerman, R., 2015. High-speed rail and regional development: the case of intermediate stations. J. Transp. Geogr. 42, 157–165.

Wetwitoo, J., Kato, H., 2017. High-speed rail and regional economic productivity through agglomeration and network externality: a case study of inter-regional transportation in Japan. Case Stud. Transport Policy 5, 549–559.

Willigers, J., van Wee, B., 2011. High-speed rail and office location choices. A stated choice experiment for the Netherlands. J. Transp. Geogr. 19, 745–754.