Dialect Diversity, Factor Agglomeration and City Size: Empirical Test Based on Satellite Night-Time Light Data

Ding Congming (丁从明)*[1] Huang Xueyang (黄雪洋) and Zhou Min (周敏)
The School of Public Policy & Administration, Chongqing University, Chongqing, China

Abstract: *Since the reform and opening up program was initiated in 1978, China's urbanization has made rapid progress, but urban development remains unbalanced and insufficient. From the perspective of social and cultural diversity, this paper explores the impact of dialect diversity on city size. Dialect diversity impedes the expansion of cities by causing a trust segmentation and impeding cross-regional factor flow and the factor agglomeration effect. Based on the regional dialect diversity indicator and the NPP-VIIRS city night-time light index of 2016, this paper offers an empirical study of the impact of dialect diversity on city size. Econometric results indicate that dialect diversity has a significantly negative impact on city size. On average, an increase of each dialect sub-category leads to a decrease in city size estimated by the night-time light index by 4.55%. Robustness test and causality identification reveal that the estimated results of this paper have a robust causal relationship. Further empirical research indicates that dialect diversity affects the expansion of city size by inhibiting the flow and agglomeration of labor, capital and technology factors. Our research suggests that the development of diverse and inclusive modern cities needs to balance the costs and benefits of cultural diversity and uniformity, break through cultural barriers, remove cultural prejudices, raise the level of social trust, and give play to the complementary effect of cultural diversity.*

Keywords: *dialect diversity, city size, night-time light index*
JEL Classification Code: Z13, O18, R58
DOI: 10.19602/j.chinaeconomist.2021.07.09

1. Introduction

Since the reform and opening up program was launched in 1978, China's urbanization rate had steadily increased from 17.9% in 1978 to 60.6% in 2019. Despite the sharp rise in the urbanization rate, the quality of China's urbanization has yet to improve. Amid the transition of China's economy from rapid growth to high-quality development, the spatial layout of cities "dominated by city clusters with coordinated development between large, medium-sized and small cities and small towns" needs to take shape. Unravelling the pattern of the spatial flow and agglomeration of economic factors is the key to optimizing the layout of spatial economy and giving play to factor agglomeration and economies of scale. These questions require academics to further explore the deep-seated determinants of the size and layout of cities.

* CONTACT: Ding Congming, email: hayekting@126.com.
Acknowledgement: Chongqing University Innovation Capacity Improvement Program "Cross-Jurisdictional Coordinated Development and Governance of Contiguous Poor Regions" (Grant No.2019CDSKXYGG0043); Chongqing University Fundamental Research Funds Program for the Central Universities "Regional Differences of Social Trust Modes and the Economic Effects" (Grant No. 2018CDJSK01XK06).

Cities expand under the interactive effect of economic, social and cultural factors. Such expansion is usually manifested in the growth of urban population, economy and land use given the quantitative and spatial attributes of the size of cities. Numerous studies have been carried out to identify the determinants of the size of cities from such perspectives as economic and political factors, natural conditions and transportation (Chen *et al.*, 2007; Duan, 2013; Sun *et al.*, 2019) with fruitful results. These factors have indeed played an important role in the formation and development of cities, but very few studies have been carried out to investigate the cultural diversity and size of cities from a cultural perspective. For China as a large country with great cultural diversity, language is an ideal variable for measuring cultural diversity as it is a key manifestation of culture. Despite a uniform written language, China has a variety of dialects across regions. Aside from the cost of communication, dialects have also resulted in trust segmentation among those speaking different dialects, thus affecting labor flow, technology diffusion and the formation of an integrated market (Liu *et al.*, 2015; Lin and Zhao, 2017; Ding *et al.*, 2018). Without a doubt, exploring the effects of dialect diversity on size of Chinese cities is of great significance to unravelling the underlying mechanism of China's urbanization and the economic and spatial layout of cities.

Academics have carried out an abundance of research on cultural and dialect diversity. Xu *et al.* (2015) found that dialect diversity impeded economic growth by putting up barriers to knowledge and technology diffusion. Lin and Zhao (2017) found that the differences of dialect prevented advanced technologies from diffusing to less developed regions. Ruan and Wang (2017) proved that regions with greater language differences also had more disparate market systems. Ding *et al.* (2018) found that dialect diversity would increase regional market segmentation and impede the formation of an integrated domestic market. Meanwhile, dialect diversity also has an adverse impact on the external openness of cities (Li *et al.*, 2017). Liu *et al.* (2015) demonstrated an inverted U-shaped pattern in the flow of labor force across regions with different dialects. From a more microscopic firm perspective, Dai *et al.* (2016) proved that dialect consistency between a company's board of directors and management helps reduce the cost of proxy. This paper is primarily concerned with how dialect diversity would affect the size of cities. Dialects would affect social trust (Huang and Liu, 2017) with an adverse impact on technology diffusion and labor migration, causing market segmentation and inhibiting the spatial agglomeration effect. Obviously, the above studies imply that dialect diversity may impede the expansion of cities.

2. Theoretical Hypotheses and Research Design

2.1 How Dialect Diversity Influences the Size of Cities

Language is a vehicle of information communication and exchange and a bridge of people-to-people interactions. Yet regional dialects will present barriers to language communication and increase the psychological distance between people. Dialects affect the level of social trust under the mechanism of group identification and screening with effects on the transmission and diffusion of factors and technologies such as human, financial and material factors, thus creating implicit barriers. Based on existing literature, this paper identifies the following influences of dialects and the theoretical channels through which dialects influence the size of cities:

2.1.1 Identification effect of dialects influences the level of public trust by creating trust barriers

Dialects are a signal and symbol of group identity. People tell each other's identity from their accent, which forms the basis for the level of trust between them. Each dialect is spoken in a specific region, and the regional heterogeneity of dialects is an important dimension for classifying ethnical groups and identities (K. Pendakur and R. Pendakur, 2002). In an environment of diverse dialects, those who speak the same dialect come from the same region with a similar cultural background, making it more likely for them to have similar ways of thinking and communicate in ways more acceptable to each other. Such similarities and mutual understanding tend to remove barriers to trust. Based on a study on Shanghai's labor market, Chen *et al.* (2014) found that the Shanghai dialect would influence an individual's income and access to job opportunities by enhancing intra-group identity rather than raising the cost of communication among strangers. In examining the impact of dialects on the entrepreneurship of migrant populations, Wei *et al.* (2016) found that the ability to speak local dialects had helped migrants enhance their social identity and reduce employment and entrepreneurial discrimination, thus spurring entrepreneurship.

2.1.2 Trust barriers of dialect diversity impede factor flow and agglomeration, thus affecting city expansion

City is the manifestation of the agglomeration effect. City development is a process of the regional agglomeration of human, financial and material factors (Chen and Huang, 2008). As can be learned from existing literature, dialects influence factor flow primarily via the following mechanisms: (1) Dialects impede the migration of labor factor. Li and Meng (2014) found that workers considered barriers to communication in standard mandarin and shared cultural background as key determinants in deciding whether or not to migrate to another region. Considering the complementary and recognition effects of dialects, Liu *et al.* (2015) found that with the shortening distance of dialects, the effect of dialects on labor migration turned from inhibitive to stimulative. (2) Dialects influence the flow of technology factors. Lin and Zhao (2017) found that the difference of dialects would impede technology diffusion by impeding institutional diffusion. Xu *et al.* also believed that dialect diversity would impede the flow and agglomeration of labor and technology factors and pose implicit barriers to the spatial agglomeration of urban factors, which is unfavorable to the expansion of cities. Based on these analyses, this paper puts forth the following hypothesis:

Research hypothesis: Dialects form trust segmentation among groups speaking different dialects, which impedes the sufficient flow of urban labor, capital and technology factors and thus the expansion of cities.

2.2 Explanations on Research Design

Based on the above research hypothesis, this paper employs the following linear model:

$$\ln psum_i=\alpha+\beta \mathrm{dia}_\mathrm{i}+\psi \mathrm{X}_\mathrm{i}+\varepsilon_i \quad (1)$$

Where, the explained variable $\mathrm{lnpsum_i}$ is the night-time light index and denotes the logarithm of the average luminance of lights in cities of city cluster i. It is calculated by dividing the night-time light luminous value of the built areas of all cities within a city cluster by the number of cities in the city cluster and then taking the natural logarithm of the result; $\mathrm{dia_i}$ is the key explanatory variable of this paper and denotes the number of dialects in city cluster i; $\mathrm{X_i}$ is other control variables, including economic, transportation and political factors such as market size, geographical distance of the city cluster, road density, whether a city is a municipality directly under the central government, provincially administered city or provincial capital city, fiscal decentralization, and fiscal spending; ε_i is error term for capturing the impact of unmeasurable factors on city size. β measures the impact of dialect diversity on city size, and this paper expects that $\beta<0$, i.e. greater dialects diversity corresponds to smaller cities in a city cluster.

The concept of "city cluster" needs to be specifically explained. This paper primarily references Lu and Chen's (2006) method in creating city clusters with municipal jurisdictions. Each city can be the core city of its city cluster, and whether other municipal jurisdictions are included in the city cluster is solely based on the criterion of whether a city borders the core city. For instance, cities bordering Beijing - such as Zhangjiakou, Baoding, Langfang, Tianjin and Chengde - form a "Beijing city cluster" together with Beijing as the core city, and various cities in each city cluster are equal and not subordinate to one another. This paper creates data indicators for all variables with city cluster as the basic measurement unit. This paper adopts 332 measurement units including prefecture-level cities, autonomous prefectures, municipalities and provincially administered cities and regions, excluding Hong Kong, Macao and cities in Taiwan, Tibet and Hainan.

This paper examines the effects of dialect diversity on the size of cities, and creates "city clusters" rather than traditional cities as the measurement unit for two reasons: First, provincial-level data is inappropriate for measuring the size of cities, and measurement based on provincial jurisdictions will lead to a convergence in the number of dialects in each province considering the large areas and populations of a province, so that variation in the key explanatory variable will be lacking if provincial-level dialect data is adopted. Second, the administrative hierarchy of cities is a key factor that influences the size of cities. Using individual cities as the basic measurement unit will not effectively control for the impact of the administrative hierarchy of cities on the size of cities. With these considerations, this paper creates a "dummy city," i.e. a synthesized city cluster to create an effective dialect diversity indicator and control for the impact of jurisdictions. Meanwhile, this paper incorporates the "number of cities in a city cluster" as a control variable into the regression equation to control for the impact of difference in the number of cities within a city cluster on the size of cities.

2.3 Indicators Definition

(i) Size of city: The explained variable of this paper is the size of city, which is measured by night-time light data for the following reasons: Illuminated urban areas at night have a high concentration of people's workplaces and living quarters (Doll *et al.*, 2006). More vibrant economic activities correspond to a higher concentration of population and night-time illumination (Elvidge *et al.*, 2007), and a higher luminance value of a region represents a higher level of economic prosperity in the region (Chen and Nordhaus, 2011). Using night-time light data, Yang *et al.* (2011) performed a quantitative estimation of urbanization rate, which has verified the reliability and universality of this method. Wu *et al.* (2014) verified the feasibility and credibility of estimating the hierarchical structure and spatial structure of cities using night-time light data. Levin and Zhang (2017) found that NPP-VIIRS urban night light data could provide a high fit for city GDP, population and built area. Using night-time light data, Yang *et al.* (2017) created the city size index to verify the impact of local government competition on change in city size.

This paper references Yang *et al.* (2017) in extracting night-time light data. Specifically, we corresponded NPP/VIIRS night-time light data of 2016 to individual Chinese city jurisdictions under the ArcMap platform, calculated the lattice area of each picture element above a grey value, and selected the grey value having the minimal error with the built area of actual cities as the optimal luminance extraction threshold, believing that the illuminated area and shape at this threshold are the closest to those of actually built cities. Lastly, we extracted the sum of night-time light luminance values of cities above the luminance threshold as the total light luminance for the built area. This paper estimated the optimal light luminance extraction threshold to be 22, i.e. the illuminated area is the closest to the built area of cities when the luminance value is no less than 22, calculated the average night-time luminance value for the built area of each city cluster, and taken the natural logarithm of the result as the explained variable of this paper.

(ii) Dialect diversity: Referencing Xu *et al.* (2015), this paper has included both mandarin dialects and dialects spoken among ethnic minority groups. Data of mandarin dialects diversity and ethnic minority dialects are from the *Great Dictionary of Modern Chinese Dialects* and *China Language Atlas,*[1]

[1] Xu Baohua, Miyata Ichirou. 1999. *Great Dictionary of Modern Chinese Dialects*. Beijing: Zhonghua Book Company; The Chinese Academy of Social Sciences and Australian Academy of the Humanities. 1987. *China Language Atlas*. Hong Kong: Longman Publishing Co., Ltd.

from which city dialect diversity data are collected for this study, and absolute and relative indicators are created to measure dialect diversity.

Absolute indicator: The number of dialects spoken in the city, which can be classified into the following four categories by language hierarchy: Dialect categories depicting the language family of each dialect; dialect subcategories depicting the type of each dialect under the language family; dialect clusters under each dialect subcategory; dialect sub-clusters expressing ethnic minority vernacular and sub-dialects.[2] In the benchmark regression, this paper employs the number of relatively common dialect subcategories to measure dialect diversity. In the subsequent section, we also provide the empirical results of dialect categories, clusters and sub-clusters to verify the robustness of the results.

Relative indicators: Considering the inadequacy of measuring dialect diversity solely by the types of dialects and the correlation of dialect difference with the number of dialect users and dialect similarities, this paper creates the "dialect distance" indicator for measuring the difference of dialects between city clusters referencing Xu *et al.* (2015) and Liu *et al.* (2015). The following equation is thus designed: $div\_s_i=\sum_{j=1}^{J}\sum_{k=1}^{K}S_{ji}\times S_{ki}\times d_{jk}$, where div_s is the dialect distance indicator, S_{ji} and S_{ki} respectively denote the proportions of population speaking dialect subtype j and dialect subtype k within city cluster i, and d_{jk} is the dialect distance between dialect subtypes j and k. This indicator is created with the following method: If j and k are the same dialect subcategory, value 0 is assigned; if they are different dialect subcategories of the same dialect category, value 1 is assigned; if they are different dialect subcategories of different dialect categories, value 2 is assigned. Greater value of div_s indicates a greater difference of dialects within the city cluster, i.e. greater dialect diversity depicted by the relative indicator.

(iii) Other control variables: On the basis of existing literature, this paper further controls for the

Table 1: Control Variables and Methods of Creation

Control variables	Method of creation
Market size	Population per square kilometer*GDP per capita. A region with a denser population and higher income will develop a larger market size and thus form larger cities, making it necessary to control for the impact of market size.
Geographical distance	Calculation of geographical distance between cities: Search for the length (in kilometers) of the most convenient transportation route between two cities in Baidu Map app to calculate the average length between other cities in the cluster and the core city.
Road density	Road length (kilometers) of the city cluster / jurisdictional area of the city cluster (in kilometers). This variable measures the development of transportation in a city cluster.
City administrative hierarchy	Dummy variable for measuring "whether there is any municipality directly under the central government, provincial capital city or provincially administered city." The value is assigned 1 if a city of any of the above levels exists; otherwise, the value is 0.
Fiscal decentralization	Per capita fiscal budgetary spending of a city cluster / national per capita fiscal budgetary spending. This variable controls for the level of a city cluster's fiscal decentralization, i.e. the impact of government fiscal autonomy in the city cluster on the size of the city cluster.
Number of cities in city cluster	Number of cities that form a city cluster. This variable controls for the expansive effect of an excessive number of cities in a city cluster.
Fiscal spending	Total fiscal spending of a city cluster in natural logarithm. This variable controls for the impact of government behaviors and policy-making on the size of cities.
Number of ethnic minority groups	Number of ethnic minority groups with populations above 10,000 in a city cluster.

[2] To create dialect categories of city clusters, this paper matches the dialect data of 2,318 county-level jurisdictions in 1986 to 332 city clusters in 2016.

following variables that may influence the size of cities: Duan (2013) believed that market size and public fiscal spending would significantly influence change in the size of cities; Li *et al.* (2018) proved that a city's administrative hierarchy and access to transportation would significantly influence the city's formation and evolution; Chen *et al.* (2007) found that a city's growth would be significantly influenced by its market potentials and whether it was located in a coastal region; Sun *et al.* (2019) found that policy intervention and the costs of communication and transportation between cities would influence the difference in the size of cities; city clusters with more administrative jurisdictions were more likely to have smaller cities. To control for the above effects, this paper has included the number of prefectural-level cities in each city cluster as a control variable. Please refer to Table 1 for the creation of control variables and *City Statistical Yearbooks* of various provinces, municipalities and cities in 2016 for the relevant statistics of each city.

3. Analysis of Empirical Results

3.1 Benchmark Model

Table 2 is the benchmark regression results of this paper. Estimated results in Column (1) indicate that the regression coefficient of dialect subcategories is significantly negative at 1%, and the estimated results suggested that an increase of each dialect subcategory corresponds to a decrease in the city cluster's night-time light index by 7.1%. Column (2) includes the economic and transportation factors influencing the size of cities, and the regression result remains significantly at 1%. The coefficients of other control variables are more or less consistent with existing studies, and both the expansion of market size and access to transportation will generate positive effects on the expansion of cities. Column (3) includes political factor and the number of ethnic minority groups on the basis of Column (2). The regression coefficient of the key explanatory variable, i.e. dialect subcategories, remains significantly negative at 1%, and the absolute value of the estimated coefficient has spiked to 11.3%, i.e., an increase of a dialect subcategory in the city cluster corresponds to a decrease in the city cluster's night-time light index by 11.3%. The estimated results suggest that city clusters with more diverse dialects have smaller night-time light indexes and smaller cities in the city cluster.

To more intuitively verify the above conclusions, Columns (4) through (6) of Table 2 employ the natural logarithms of the average built area, average permanent population and average GDP of cities within a city cluster as substitutes of night-time light index to participate in regression as the explained variable, respectively. The regression coefficient of the key explanatory variable, i.e. dialect subcategories, in Columns (4) through (6) remains significantly negative: An increase of a dialect subcategory leads to an 8.24% decrease in city built area, a 1.37% decrease in permanent population, and a 5.25% decrease in GDP. The estimated results once again demonstrate a negative correlation between dialect diversity and city size. Results in Table 2 have initially verified our basic hypothesis: The size of cities is smaller in regions with more diverse dialects.

3.2 Robustness Test

In this section, we will perform a series of robustness tests on the basis of the above section.

(i) Language hierarchy: There is some subjectivity in language classification, which is based on vocabulary, pronunciation, intonation and other linguistic attributes. Quantitative research on dialects aims to conduct a quantitative classification of dialects into different language hierarchies, i.e. dialect categories, subcategories, clusters and sub-clusters, according to those linguistic attributes. In the benchmark regression of this paper, we only used dialect subcategories as the indicator for measuring dialect diversity. To test the robustness of the benchmark regression results, Columns (1)-(3) in Panel A of Table 3 further use such indicators as "dialect categories," "dialect clusters" and "dialect sub-clusters"

Table 2: Benchmark Regression Results

	Night-time light index			City built area	Permanent population	GDP
	(1)	(2)	(3)	(4)	(5)	(6)
Dialect subcategories	-0.071*** (0.018)	-0.054*** (0.018)	-0.113*** (0.025)	-0.0824*** (0.013)	-0.0137* (0.007)	-0.0525*** (0.014)
Market size		0.000176*** (0.000)	0.000134*** (0.000)	0.0000835*** (0.000)	0.0000269*** (0.000)	0.000100*** (0.000)
Geographical distance		0.269** (0.106)	0.219** (0.110)	0.0232 (0.060)	-0.0378 (0.033)	0.0628 (0.061)
Road density		-0.231 (0.176)	-0.212 (0.187)	0.270*** (0.102)	0.649*** (0.057)	0.633*** (0.104)
Administrative hierarchy of cities			0.0421 (0.165)	0.0719 (0.090)	0.0231 (0.050)	0.100 (0.091)
Fiscal decentralization			-0.00179 (0.075)	-0.0110 (0.041)	0.0172 (0.023)	-0.0101 (0.041)
Number of cities in the city cluster			-0.0419 (0.037)	-0.0273 (0.020)	-0.0909*** (0.011)	-0.0410** (0.021)
Fiscal spending			0.147 (0.124)	0.346*** (0.068)	0.637*** (0.038)	0.404*** (0.069)
Number of ethnic minorities			0.0741*** (0.019)	0.0332*** (0.011)	0.00858 (0.006)	0.0188* (0.011)
N	300	300	300	300	300	300
R^2	0.047	0.275	0.324	0.569	0.806	0.651

Notes: ***, ** and * respectively denote significance at 1%, 5% and 10%, and numbers in parentheses are robust standard errors. The same below.

to substitute the robustness test results of "dialect subcategories" in the benchmark regression. The key explanatory variable in Column (4) is replaced with "dialect distance" as the relative indicator for dialect diversity. Regression results once again indicate that dialect diversity, no matter measured by which method, has a significantly negative impact on the size of cities.

(ii). Remeasurement of night-time light index: The night-time light indexes of city clusters are the arithmetic average values of the sums of the luminance values of all cities in each city cluster. Considering the potential impact of the luminance index measurement method on the results, this paper further conducted a weighted average treatment of the luminance values of cities in each city cluster using GDP, administrative area and permanent population. Regression results are shown in Columns (1)-(3) of Panel B in Table 3. The regression coefficient of the core explanatory variable remains significantly negative and is close to the regression coefficient (11.3%) of dialect subcategories in Column (3) of Table 2, which indicates that difference in the light index measurement methods does not affect this paper's basic conclusions.

Since the night-time light data provides a good fit of such attributes as shape, layout and area of the city built area (Levin, 2017), the night-time light data employed in this paper's benchmark regression as the explained variable is obtained based on the actual built area of each city. With the night-time light data extraction method provided in Section 2 of this paper, the extracted lattice fitted area with luminance no less than 22 has an 85% similarity with the built area of actual cities. That is to say, there is still a certain measurement error in fitting the size of cities with night-time light data. In Column (4) of Panel B in Table 3, this paper further replaced the original data in the benchmark regression with the night-time light data extracted under the "optimal goodness of fit" principle for regression. The fitted

Table 3: Robustness Test: Different Measurement Methods Based on Dialects and Night-Time Light Index

Panel A	**Night-time light index**			
	(1)	(2)	(3)	(4)
Dialect categories	-0.246*** (0.062)			
Dialect clusters		-0.0892*** (0.017)		
Dialect sub-clusters			-0.0442*** (0.015)	
Dialect distance				-0.416** (0.189)
Control variables	Yes	Yes	Yes	Yes
N	300	300	300	300
R^2	0.312	0.341	0.297	0.287
Panel B	**Explained variable: night-time light index under different measurement methods**			
	GDP weight	Area weight	Population weight	Night-time light index
Dialect subcategories	-0.132*** (0.030)	-0.132*** (0.025)	-0.131*** (0.028)	-0.0988*** (0.018)
Control variables	Yes	Yes	Yes	Yes
N	300	300	300	300
R^2	0.273	0.334	0.312	0.428

area under the "optimal goodness of fit" principle has a 90% goodness of fit with the actual built area of cities,[3] which has to some extent reduced data bias arising from the goodness of fit. Results of regression have once again verified the robustness of the benchmark regression results.

4. Causality Identification

This paper will further discuss the potential problem of estimation bias by further controlling for more control variables and estimating the instrumental variable.

4.1 Problem of Missing Variables

For a cleaner identification of the impact of dialect diversity on the size of cities, this paper further introduces six factors: the potential attributes of city cluster, other cultural factors, the level of government intervention, the level of external openness, the layout of cities in a city cluster and industrial structure to avoid potential missing variables.

(i) Impact of potential city cluster attributes: Potential city cluster attributes refer to the intrinsic factors that do not change with time in a certain period and may also influence the size of cities in the city cluster. Based on the research background, this paper divides potential city cluster attributes into the following categories: First, geographical factors, including whether the city cluster is located in a coastal region and the average slope of the city cluster. The second factor is administrative jurisdiction, i.e. the

[3] The "optimal goodness of fit" extraction method is similar to the night-time light data extraction method explained in Section 3 of this paper. The difference is that this method no longer requires the fitted area to be similar to the actual built area of cities and instead requires the highest goodness of fit between the two in regression without the intercept term. The light luminance value extracted with this method is 9, and the actual built area of cities is about 0.45 times the fitted area. This paper has also employed this method for estimation and supplemented the missing values for the actual built areas of some cities.

jurisdictional area of each city cluster is included on the basis of the benchmark regression to further control for the impact of jurisdiction. Column (1) of Table 4 reports the regression results. It can be seen that with geographical and jurisdictional factors further taken into account, this paper's basic conclusions remain robust.

(ii) Other cultural factors: Dialects as the vehicle of regional culture influence the size of cities probably due to the impact of other cultural factors rather than dialects *per se*. To exclude this possibility, this paper controls for the "number of cross-cultural zones of a city cluster" as the variable for other cultural factor to identify whether the size of cities is influenced by dialects *per se* or other cultural factors. The division and distribution of cultural zones are from Wu (1996). Column (2) of Table 4 displays the results of controlling for other cultural factors. The regression coefficient of "number of cross-cultural zones in a city cluster" is insignificant, and the regression coefficient of dialect subcategories as the key variable is significantly negative, which indicates that the size of cities is influenced by dialects rather than other cultural factors.

(iii) Government intervention: Cities are a manifestation of the agglomeration effect, which has a lot to do with the size of cities. Given the potential impact of government intervention on the size of cities, we define the ratio of government fiscal spending in a city cluster / GDP of the city cluster as the indicator for measuring "level of government intervention in the city cluster," which is included into the regression equation as a control variable with regression results shown in Column (3) of Table 4. Compared with the regression results of the first two columns of Table 4, the absolute value of the regression coefficient for dialect categories as the key explanatory variable has decreased by about 2% but is still significantly negative.

(iv) Level of external openness. Li *et al.* (2017) found that dialect diversity has a significantly negative impact on a city's external openness. Fan (2019) demonstrated that an increase in the level of external openness would promote a city's development. This paper measures a city cluster's level of external openness by the ratio of the total import and export volume of a city cluster / GDP of the city cluster, and Column (4) of Table 4 displays the results of controlling for this variable. Regression results indicate that the level of external openness has a significantly positive effect on the size of cities. Meanwhile, the regression coefficient of dialect diversity as the key explanatory variable remains significantly negative with an estimated coefficient of -7.9%. After the level of the city's economic openness is taken into account, the negative impact of dialect diversity on the size of cities is free from interference.

(v) Distribution of cities: The distribution of cities is either concentrated or scattered. When the distribution of cities in a city cluster is concentrated, there will be a concentration of human, financial and material factors in one city or a few leading cities within the city cluster. On the contrary, when the distribution of human, financial and material factors is scattered in various cities of a city cluster, the distribution of cities in the city cluster is considered as scattered. The distribution of cities is measured by population concentration and economic concentration, respectively. A higher indicator of concentration suggests greater concentration in the distribution of cities, and vice versa. After controlling for the distribution of cities, the regression coefficient of the core explanatory variable remains negative (see Column (5) of Table 4).

(vi) Level of industrial structure: Referencing Zhao and Shen (2019), we have created the industrial structure upgrade index (β) to measure the status of industrial development in a city cluster. Specifically, $\beta=\sum_{i=1}^{3} y_i*i$, where yi is industry i's output value as a share of GDP with a value range of $1\leq\beta\leq3$, the closer β is to 1, the less sophisticated the city's industrial structure is, and the closer it is to 3, the more sophisticated the city's industrial structure is. With the industrial structural upgrade index β incorporated as a control variable for regression, Column (6) of Table 4 reveals that after the industrial structure of a city cluster is controlled for, dialect diversity still has a significantly negative impact on the size of cities.

Table 4: Missing Variables

	Night-time light index						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dialect subcategories	-0.0867*** (0.025)	-0.0883*** (0.025)	-0.0667*** (0.023)	-0.0799*** (0.024)	-0.0763*** (0.025)	-0.0701*** (0.023)	-0.0455** (0.022)
Average slope	-0.108*** (0.036)						-0.0834** (0.033)
Location in a coastal region	0.437** (0.170)						0.120 (0.150)
Jurisdictional area of the city cluster	0.132 (0.243)						0.301 (0.216)
Number of cross cultural zones		0.0258 (0.066)					-0.00307 (0.057)
Level of government intervention			-4.644*** (0.654)				-3.746*** (0.623)
Level of external openness				1.424*** (0.362)			0.921*** (0.329)
Population concentration					0.154*** (0.050)		0.121*** (0.045)
Economic concentration					0.0456 (0.048)		0.133*** (0.042)
Level of industrial structure						4.424*** (0.671)	3.827*** (0.632)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample size	300	300	300	300	300	300	300
R^2	0.473	0.359	0.455	0.392	0.382	0.444	0.554

Column (7) of Table 4 shows the regression results after incorporating all the above missing variables, and the regression coefficient of dialect diversity remains significantly negative. Compared with the benchmark regression, an inclusion of each dialect subcategory will reduce the decrease in the size of cities from 11.3% to 4.55%. This explains that after the impact of missing variables is taken into account, the estimated results of this paper have become cleaner.

4.2 Estimation of Instrumental Variable (IV)

Tests with the above methods have shown the robustness of the estimated results in this paper, but the following simultaneity problem cannot be theoretically excluded: Did dialect diversity impede the expansion of cities or did the historically sporadic layout of cities lead to a greater variety of dialects? With this question in mind, this paper conducts a further analysis with an instrumental variable. Referencing Ding *et al.* (2018), this paper employs local opera types as the instrumental variable of dialect diversity. Normally, the local operas of a region are sung in the local dialect for the target audience of local residents, and the diversity of local operas results from the diversity of local dialects (You and Zhou, 1985). Hence, the relevance of local opera types as the instrumental variable for dialect diversity is satisfied. In this paper, we have calculated the number of local operas in various city clusters based on the distribution of local operas in the *China Opera Classification Handbook*.

As far as exogeneity is concerned, we have little *a priori* reason to believe that opera types influence the size of cities. Theoretically, there could be some other factors that influence both opera types and city size. The greatest possibility is that cultural diversity as a factor other than dialects influences both opera types and the expansion of cities. To exclude the above possibility, Table 5 has controlled for the number of cultural zones in a city cluster in the empirical regression, and the specific control variables are the

Table 5: Instrumental Variable Estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	Second-stage estimation results			Plausible exogeneity of the instrumental variable: LTZ method		
	Full sample	Full sample	Excluding samples containing municipalities	Excluding ethnic minority languages zones	Full sample	Excluding samples containing municipalities	Excluding ethnic minority languages zones
Dialect subcategories	-0.0530** (0.023)	-0.238* (0.138)	-0.260** (0.123)	-0.241* (0.135)	-0.239* (0.134)	-0.259** (0.123)	-0.240* (0.132)
Local opera types	-0.0130 (0.009)						
Number of cross-cultural zones	0.0483 (0.060)	0.118 (0.085)	0.128 (0.080)	0.101 (0.083)			
Control variables	Controlled	Controlled	Controlled	Controlled	Controlled	Controlled	Controlled
Sample size	300	300	277	278	300	277	278
		First-stage estimation results Explained variable: Number of dialect subcategories in a city cluster					
Local opera types		0.0702*** (0.022)	0.0863*** (0.025)	0.0729*** (0.025)			
Control variables	Controlled	Controlled	Controlled	Controlled			
Sample size	300	300	277	278			
R^2	0.499	0.655	0.644	0.632			
F-statistic	—	25.12	25.3	21.89			

same with Column (2) of Table 4. On the other hand, we retested the robustness of the IV's estimation results. Conley *et al.* (2012) assumed that the instrumental variable was close to plausible exogeneity, and thus observed the trend of change in the IV's estimation results at different levels of plausible exogeneity. The two-stage estimated results of robustness test of the instrumental variable "local opera types" under the condition of plausible exogeneity are shown in Columns (5) through (7) of Table 5.

Table 5 reports the results of instrumental variable regression. In Column (1), we also included the instrumental variable into the regression equation for a simple display of the instrumental variable's exogeneity, and the result shows that the estimated coefficient of opera types is insignificant, which indirectly verifies that the instrumental variable had no direct impact on the size of cities. Columns (2) through (4) report the estimated results using the two-stage least square method. The first-stage regression results are significantly positive at 1%, which verifies a high degree of correlation between opera types and dialect diversity, and F statistics are all greater than 10. According to experience, the assumption of weak instrumental variable is rejected. The second-stage estimation coefficient is negative and has passed the significance test at 10%. Overall, the estimated coefficient is consistent with expectations. The second-stage estimation coefficient has expanded by about five times compared with the regression coefficient (4.55%) of OLS regression results. The expansion of coefficient may derive from the regional average effect captured by the instrumental variable, and OLS estimation results primarily capture the overall average treatment effect, which led to an expansion in the estimated coefficient. Columns (3) and (4) are samples after excluding those containing municipalities and ethnic minority language zones, and the estimated coefficients have little difference with those in Column (2). Empirical results of the local to zero (LTZ) approach are shown in Columns (5) through (7) of Table 5. Under the LTD approach, the estimated coefficient of the endogenous variable (number of dialect subcategories) remains significantly negative, which supports the robustness of the estimated results of the instrumental variable in this paper. Of course, this paper still employed the estimated results in Column (7) of Table 5 for explanation considering the expansiveness of the instrumental variable's estimated coefficient.

5. Impact Channels

Section 2 of this paper identified the following possible channels through which dialect diversity could influence the size of cities: Under the group and cultural identity effects, dialects form a trust segmentation among groups speaking different dialects, which impedes the flow of urban labor, capital and technology factors, prevents the spatial agglomeration effect from being brought into full play, and thus inhibit the expansion of cities. This paper tests the above transmission mechanism and creates the following variables to measure the flow and agglomeration of labor, capital and technology factors. (i) Labor flow variable: Referencing Shao *et al.* (2016) and Xu *et al.* (2019), this variable measures urban labor flow by the ratio of urban permanent population / registered population (Lflow). When this ratio is greater than 1, the locality has an inflow of labor force; or the ratio is smaller than 1, the locality has an outflow of labor force. (ii) Capital variable: Referencing Wang and Zhao (2020), capital stock is measured by the fixed asset investment volume of each city (InFA). (iii) Technology variable: Technology variable is created referencing Xu *et al.*'s (2015) method, taking into account the neoclassical economic growth model under the steady state: $Y=K^{\alpha}(ALh)^{1-\alpha}$, where Y, K, L and h are substituted with GDP, fixed asset investment volume, permanent population and the logarithm of average employee wage of each city, and α's value is 1/3 for simplicity (Xu *et al.*, 2015). After substituting the above data into the neoclassical

Table 6: Indicator Explanations

Labor force variable	Cluster Lflow:	Total permanent population of a city cluster / total registered population: measures the overall labor flow status of the city cluster. The greater this value, the more workforce would migrate into the city cluster, and the greater labor force agglomeration effect. The smaller this value is, the poorer workforce agglomeration effect this region has.
	STD-Lflow:	Standard deviation of each city's Lflow in the city cluster: measures the difference in the workforce migration between various member cities of a city cluster. The greater this value is, the greater the flow of labor force is between cities of the city cluster, and the greater workforce agglomeration effect. The smaller this value is, the less flow of labor force is between cities of the city cluster, and the poorer workforce agglomeration effect this region has.
Capital variable	Cluster InFA:	Total fixed asset investment volume of the city cluster (logarithm): measures the overall capital stock of a city cluster, and a greater value suggests a higher capital stock of the city cluster.
	STD-InFA:	Difference in the fixed asset investment volumes of various cities in a city cluster: measures the difference in the capital stock of various cities in a city cluster, and a greater value suggests that the difference is greater between cities in the city cluster, that capital inside the city cluster concentrates in a few cities, and that the spatial agglomeration effect of capital is stronger.
Technology variable	Cluster A:	Technology level in a city cluster: obtained by substituting the aggregate GDP of a city cluster, aggregate fixed asset investment volume, total permanent population and the logarithm of average employee compensation into the neoclassical economic growth model. This indicator denotes a city cluster's overall level of technology, and a greater value means a higher level of technology.
	STD-A:	Standard deviation A of technology level between cities in a city cluster: denotes difference in the technology level between cities in the city cluster, and a greater value means more disparate levels of technology and a stronger agglomeration effect of technology factor.

economic growth model, the level of technology development for each city (A) can be calculated. The following treatment is performed for the above variables to better measure the flow and concentration of the three factors:

Table 7 shows the results of regression using the above indicators. It can be learned from the results of regression using labor, capital and technology factors in Panel A of Table 7 that an increase in dialect diversity in a city cluster has an adverse impact on the labor inflow of the city cluster (Cluster Lflow) and has no significant influence on the labor agglomeration effect (STD-Lflow) within a city cluster; an increase in dialect diversity in a city cluster will significantly worsen the spatial agglomeration (STD-InFA) of capital within a city cluster and has no significant effect on the flow of capital stock between city clusters (Cluster InFA); at the level of technology flow, an increase in dialect diversity in a city cluster is not only unfavorable to the overall technology level of a city cluster (Cluster A) and the free flow of technology factor between city clusters, but impedes the spatial agglomeration effect of technology (STD-A) within each city cluster. In Panel B of Table 7, we conducted a regression analysis of the factor flow and the factor agglomeration indicators with the night-time light index for measuring the size of cities. Results indicate that labor inflow (Column (1) in Panel B of Table 7), capital flow and agglomeration (Columns (3) and (4) in Panel B of Table 7) and the flow and agglomeration of technology factor (Columns (5) and (6) in Panel B of Table 7) have a significantly positive correlation with the expansion of cities.

Table 7: Test of Transmission Mechanism

	Labor flow		Capital flow		Technology flow	
Panel A	(1)	(2)	(3)	(4)	(5)	(6)
	Cluster Lflow	STD-Lflow	Cluster InFA	STD-InFA	Cluster A	STD-A
Dialect subcategories	-0.00895*** (0.002)	-0.00313 (0.003)	-0.0117 (0.012)	-0.0538*** (0.019)	-0.0821*** (0.015)	-0.0369*** (0.013)
Control variable	Yes	Yes	Yes	Yes	Yes	Yes
N	300	300	300	300	300	300
R^2	0.842	0.678	0.702	0.394	0.644	0.305
Panel B	**Explained variable: night-time light index**					
Cluster Lflow	2.687*** (0.649)					
STD-Lflow		-0.644 (0.513)				
Cluster InFA			1.034*** (0.105)			
STD-InFA				0.642*** (0.069)		
Cluster A					0.711*** (0.086)	
STD-A						0.656*** (0.110)
Control variable	Yes	Yes	Yes	Yes	Yes	Yes
N	300	300	300	300	300	300
R^2	0.315	0.279	0.441	0.440	0.414	0.354

Notes: Labor, capital and technology factor statistics are from the *China City Statistical Yearbook 2016* and the *Statistical Yearbooks* of various provinces and autonomous regions.

In summary of the regression results of Table 7, the transmission mechanism through which dialect diversity influences the size of cities is not hard to verify: An increase in dialect diversity is unfavorable to the flow and agglomeration of labor and technology factors and will significantly worsen the spatial agglomeration effect of capital within a city cluster, resulting in an adverse effect on the growth of cities. At a deeper level, this transmission mechanism stems from barriers to the migration of labor force from a cultural diversity perspective: City clusters with more diverse dialects have a more significant trust segmentation due to the identity effect of dialects and a lower level of trust among people, who tend to distrust strangers and trust those from more familiar and similar cultural and geographical backgrounds. Without a doubt, those tendencies will influence people's choice of jobs, investment decisions and technology cooperation. Cliques based on dialects and cultural backgrounds impede the flow of workforce, capital and technology factors between different groups of people, posing a barrier to the flow and agglomeration of various factors in cities and affecting the spatial agglomeration effect of cities to the detriment of the growth of cities.

6. Concluding Comments

This paper examined the effects of dialect diversity on the size of cities from cultural and economic perspectives. Since the reform and opening-up program was initiated in 1978, China's urbanization rate has vastly increased, but the spatial development of cities remains unbalanced and insufficient. Tremendous gaps exist in the size of cities between China's eastern and central and western regions and between coastal and inland regions. Differences in the size of cities in China have been extensively discussed among academics. Based on the existing literature, this paper explores the reasons behind the differences in the size of cities in China from a cultural perspective. China is a large country with diverse cultures. The cultural zones divided by dialects historically comprise a structural basis of regional economies. Through the identity effect, dialects form a trust segmentation between people speaking different dialects, thus affecting the general level of social trust and the flow and agglomeration of factors between and within cities and creating an adverse impact on the size of cities.

To exclude the interference of jurisdictional factor to dialects and the size of cities in our empirical study, this paper divided virtual city clusters as the basic measurement unit by the criterion of adjacency to core cities. To measure the quantitative and spatial attributes of city size, this paper employed night-time light data as the proxy variable for measuring the size of cities. The estimated results of empirical regression in this paper indicate that dialect diversity has a significantly negative impact on the size of cities, and the robustness of empirical results in this paper is proven by a series of robustness test results. An increase of each dialect subcategory leads to a decrease in the size of cities by 4.55% measured by night-time light index. To test the causality in the above relationship, this paper employs the "local opera types" as the instrumental variable of dialect diversity, and the estimated results of the two-stage least square method and the LTZ have verified the causality effect of dialect diversity on the size of cities. A further analysis of the transmission mechanism found that dialect diversity would influence the size of cities through its effects on the flow and agglomeration of labor, capital and technology factors.

This study shows that in the context of China's transition period and new-type urbanization, cultural diversity and the associated dialect diversity also influence difference in the size of cities at a deeper level. While striving to protect cultural and language diversity, policymakers should also raise the level of social trust, mitigate the adverse effects of trust segmentation arising from the identity and screening mechanisms of dialects, remove cultural barriers, and thus give play to the complementary effect of cultural diversity and build more diverse and inclusive modern cities.

References:

[1] Chen, Jianjun, and Jie Huang. 2008. "China's Industries, Cities and Regions from the Perspective of Agglomeration: A Summary of the Latest Developments in Domestic Spatial Economics." *Journal of Zhejiang University* (Humanities and Social Sciences Edition), 4: 12-21.

[2] Chen, Liangwen, Kaizhong Yang, and Jiao Wu. 2007. "An Empirical Study on the Evolution of China's Urban System." *Jiangsu Social Sciences*, 1: 81-88.

[3] Chen, X., and W. Nordhaus. 2011 . "Using Luminosity Data as a Proxy for Economic Statistics." *Proceedings of the National Academy of Science of the United States of America*, 108(21): 8589-8594.

[4] Chen, Z., M. Lu, and L. Xu. 2014. "Returns to Dialect: Identity Exposure Through Language in the Chinese Labor Market." *China Economic Review*, 30: 27-43.

[5] Conley, T. G., C. B. Hansen, and P. E. Rossi. 2012. "Plausibly Exogenous." *Review of Economics and Statistics*. 94(1): 260-272.

[6] Dai, Yiyi, Jinli Xiao, and Yue Pan. 2016. "Can 'Xianyin' Reduce Company Agency Costs?-A Study Based on the Perspective of Dialects." *Economic Research*, 51(12): 147-160+186.

[7] Ding, Congming, Zhenlin Ji, Yu Lei, and Zhenqiao Liang. 2018. "Diversity of Dialects and Market Integration: Based on the Perspective of Urban Circles." *Economic Research*, 11: 148-164.

[8] Doll, C. N. H., J. P. Muller, and J. G. Morley. 2006. "Mapping Regional Economic Activity from Night-time Light Satellite Imagery." *Ecological Economics*, 57(1): 75-92.

[9] Duan, Ruijun. 2013. "Study on the Size of Chinese Cities and Its Influencing Factors: Empirical Evidence from 284 Cities at the Prefecture Level and Above." *Finance Research*, 39(9): 82-94.

[10] Elvidge, D., P. Cinzano, and D. Pettit. 2007. "The Nightsat Mission Concept." *International Journal of Remote Sensing*, 28(12): 2645-2670.

[11] Fan, Yi. 2019. "The Turning Point and Transformation of China's Urban Development: Based on the Perspective of Dynamic Mechanism." *Economic Aspects*, 8: 49-60+2.

[12] Huang, Jiuli, and Chang Liu. 2017. "Dialects and Social Trust." *Financial Research*, 43(7): 83-94.

[13] Levin N., and Q. L. Zhang. 2017. "A Global Analysis of Factors Controlling VIIRS Night-time Light Levels from Densely Populated Areas." *Remote Sensing of Environment*, 190: 366-382.

[14] Li, Guangqin, Jianhua Cao, and Shuai Shao. 2017. "Language Diversity and Regional Differences in China's Opening to the Outside World." *World Economy*, 40(3): 144-168.

[15] Li, Qin, and Lingsheng Meng. 2014. "Dialects, Mandarin and Regional Mobility of Labor in China." *Journal of Economics*, 1(4): 68-84.

[16] Li, Wan, Bindong Sun, Qianqian Liu, and Tinglin Zhang. 2018. "Characteristics and Influencing Factors of China's Urban Spatial Structure." *Geographical Science*, 38(5): 672-680.

[17] Lin, Jianhao, and Zile Zhao. 2017. "Invisible Barriers to Balanced Development: Dialects, Institutions and Technology Diffusion." *Economic Research*, 52(09): 182-197.

[18] Liu, Yuyun, Xianxiang Xu, and Ze Xiao. 2015. "The Inverted U-shaped Pattern of Labor Force Flow across Dialects." *Economic Research*, 50(10): 134-146+162.

[19] Lu, Ming, and Zhao Chen. 2006. *Market Integration and Industrial Agglomeration in China's Regional Economic Development*. Shanghai: People's Publishing House.

[20] Pendakur, K., and R. Pendakur. 2002. "Language as Both Human Capital and Ethnicity." *International Migration Review*, 36(1): 147-177.

[21] Ruan, Jianqing, and Ling Wang. 2017. "Language Differences and Market System Development." *Management World*, 4: 80-91+187-188.

[22] Shao, Yihang, Yujuan Wang, and Yanan Liu. 2016. "Labor Mobility and the Evolution of Income Gap: Theory and Empirical Research Based on Chinese Cities." *The Economist*, 1: 33-41.

[23] Sun, Bindong, Xiaoxi Jin, and Jie Lin. 2019. "China's New Urbanization Pattern towards the Coordinated Development of Large, Medium and Small Cities: the Evolution and Influencing Factors of China's City Size Distribution since 1952." *Geographical Research*, 38(1): 75-84.

[24] Wang, Linhui, and Xing Zhao. 2020. "Spatial Flow of Elements, Types of Heterogeneous Industrial Agglomeration and Regional Economic Growth-Based on the Analysis of the Yangtze River Delta and Northeast Region." *Learning and Exploration*, 1: 116-122.

[25] Wei, Xiahai, Siyu Chen, and Jiahui Li. 2016. "Dialect Skills and Floating Population's Entrepreneurship Choices." *Chinese Population Science*, 6: 36-46+126-127.

[26] Wu, Bihu.1996. "The Formation and Division of Chinese Cultural Districts." *Academic Monthly*, 3: 10-15.

[27] Wu, Jiansheng, Hao Liu, Jian Peng, and Lin Ma. 2014. "The Hierarchical Structure and Spatial Pattern of China's Urban System:

Empirical Based on DMSP/OLS Night Light Data." *Acta Geographica Sinica*, 69(6): 759-770.

[28] Xu, Xianxiang, Yuyun Liu, and Zekai Xiao. 2015. "Dialects and Economic Growth." *Journal of Economics*, 2(2): 1-32.

[29] Xu, Yan, Xiao Yin, and Xiguo Yin. 2018. "Labor Mobility and Wage Changes of Chinese Urban Residents." *Population and Economy*, 4: 47-64.

[30] Yang, Mei, Shixin Wang, Yi Zhou, and Litao Wang. 2011. "Remote Sensing Estimation Method of Urbanization Level Based on DMSP/OLS Image." *Remote Sensing Information*, 4: 100-106.

[31] Yang, Mengyu, Zhibing Cai, and Keyun Zhang. 2017. "The Measurement of the Size of Chinese Cities and the Source of Spatial Competition-Research Based on Global Night Light Data." *Financial and Trade Economics*, 38(3): 38-51.

[32] Yang, Renfei, Hongxia Luo, Sheng Zhou, Yusi Cheng, Jingyi Chen, Haiyan Xiang, and Qian Lei. 2017. "Reconstruction and Analysis of the Spatial Formation Process of Chengyu Urban Agglomeration Driven by Night Light Data." *Journal of Geo-Information Science*, 5: 653-661.

[33] You, Rujie, and Zhenhe Zhou. 1985. "Dialects and Chinese Culture." *Fudan Journal* (Social Science Edition), 3: 232-237.

[34] Zhao, Ranran, and Chunmiao Shen. 2019. "Capital Flow, Industrial Agglomeration and Industrial Structure Upgrading: An Empirical Analysis Based on Panel Data of 16 Central Cities in the Yangtze River Delta." *Exploration of Economic Issues*, 6: 135-142.