Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 6540663, 12 pages
https://doi.org/10.1155/2022/6540663

Research Article

Analysis of Business Environment and Medical Insurance Coverage Rates in the Destination of China's Migrant Population: Based on Geographically and Temporally Weighted Regression Model for Panel Data

Yang Liu, Xiaoyu Chen, and Yujie Zhang

School of Economics and Management, Zhejiang Ocean University, Zhoushan, China

Correspondence should be addressed to Yang Liu; liuyangjg@zjou.edu.cn

Received 19 August 2022; Accepted 26 September 2022; Published 19 November 2022

Academic Editor: Lianhui Li

Health risk is an important issue in the process of population spatial mobility, and it is also an important issue in the process of urbanization in China. Using the dynamic monitoring data of China's migrant population and 31 provincial business environment data from 2011 to 2018, this study systematically investigated the spatial distribution and evolution characteristics of the migrant population's participation in medical insurance in the destination areas and combined it with the Geographically and Temporally Weighted Regression Model for Panel Data (PGTWR) to analyze the impact of the regional business environment on the medical insurance coverage rate of the inflow area. The results are as follows: first, the spatial pattern of the insurance coverage rate is high in Eastern China and low in Western China. The areas with high insurance coverage rates are mainly distributed in the three major economic circles of China and provinces such as Shandong and Xinjiang. Second, there is significant spatial autocorrelation in the coverage rate, which shows that cities with high participation rates tend to form agglomeration areas in geographical space, and so do cities with low participation rates. The coverage rate of the popular areas is scattered, while the unpopular areas are concentrated. Third, the estimation results of the Geographically and Temporally Weighted Regression Model for Panel Data show that the macroeconomic and infrastructure indicators in the business environment have a greater impact on the insurance rate, while the impact of policy environment indicators is relatively weak. However, the overall improvement of the business environment can significantly improve the probability of the migrant population participating in medical insurance in the inflow area.

1. Introduction

China is currently experiencing the largest population movement in its history. Data from the seventh national census show that the total migrant population in China will reach 376 million by 2020, accounting for 26.6% of the total population. With the continuous acceleration of China's urbanization process, the migrant population will still maintain a large scale of growth. In the traditional market economy theory, the micromigration decision of the floating population mainly depends on the net economic benefit difference, i.e., wage difference. However, the labor force will encounter a range of risks and uncertainties in the process of spatial mobility. Therefore, in reality, the factors that determine the inter-regional migration decision of the floating population are more complex: on the one hand, the migration decision takes into account comprehensive income, and family members' consideration of education, medical care, and social security are also important; on the other hand, the migrant population is not highly educated, lacks certain professional labor skills, and is often engaged in jobs with relatively dangerous occupations, long working hours, and poor working environments, which have greater health risks than residents. Therefore, the migrant population has a strong desire to obtain social security and medical benefits in the cities where they lived and worked [1, 2].

In the past 20 years, China's social security system has been continuously improved, and by the end of 2015, before the Chinese government integrated the two systems of new rural cooperative medical care and basic medical care for urban residents, the number of China's medical insurance participants reached 1.336 billion, and the participation rate has been maintained at over 95% since then, basically establishing a "low-level, broad-coverage" medical insurance system. Unfortunately, however, influenced by various factors such as the household registration system, the migrant population, as an important group for economic construction, still cannot have the same social insurance treatment as residents. By 2017, the participation rate in the destination of the migrant population was only 27.6% as per data derived from the China Migrants Dynamic Survey, far below the national average level of medical insurance participation rate, and there are significant geographical differences. Therefore, it is of great significance to deeply analyze the spatial distribution pattern of the migrant population participating in medical insurance in the destination areas and explore the main influencing factors to promote the transformation of medical insurance from full system coverage to full population coverage.

An important difference between the floating and local populations in China's urban labor market is social insurance participation. Existing research has focused on floating access to social insurance and participation rates [3–5], especially for the large rural migrants, the urban-rural divide caused by a rural hukou status is an important source of low access to social insurance opportunities [6–9]. However, it is inaccurate to attribute all the reasons for the exclusion of the migrant population from the urban social security system to policy discrimination and urban-rural differences; as early as the 2000 s, Beijing, Shanghai, and Guangdong (popular destinations for the immigrants) already allowed the migrant population to participate in all five social insurance schemes, including pension, health, work-related injury, unemployment, and maternity insurance. Particularly, in 2016, the Chinese government issued relevant policies to break the original localized management barriers of medical insurance, so that the immigrants can enjoy the same medical insurance treatment as the residents in the destinations, and greatly reduce the cumbersome procedures of the migrant population in the process of insurance participation and reimbursement. Another part of the literature mainly analyzed the influencing factors of migrant population participation in social insurance, such as gender, age, education level, and other demographic characteristics, and analyzed the impact of occupation type, nature of the company, work income, and other occupational characteristics [10–13]. In the existing literature related to this study, metrics are usually measured using econometric models, and commonly used models are discrete choice models such as Logistic and Probit. The above methods are more applicable to the study of individual factors and do not reflect the spatial heterogeneity of the study area and the spillover effects between regions.

However, if the acquisition of social insurance is regarded as an "optional commodity" rather than a legal right, it is not enough to only pay attention to the characteristics of the "demand side," the migrant population. As the "supply side," the policy environment and inclusive culture provided by the destination cities also have an influence that cannot be ignored. Considering the political, economic, social, cultural, ecological, and other factors of the destination government, the business environment is undoubtedly more directly related to labor mobility. In recent years, regions that attach importance to the optimization and improvement of the business environment have gained more migrant population dividends—on the one hand, a good business environment can promote economic development, thereby producing a siphon effect on labor [14]. On the other hand, a high-quality business environment means better financing convenience, better intellectual property protection, more effective government services, and better preferential tax policies. It will mainly serve enterprises and can effectively help enterprises invest in production, thus creating more jobs for labor inflows.

This study expands the relevant literature on China's immigrants participating in medical insurance from the following aspects: the first contribution focuses on "regional differences in policy environment"—the impact of the business environment on migrants' participation in insurance because orderly mobility is an inevitable requirement to promote labor market integration, but migration decisions are largely affected by different regional environments and policies. Therefore, "inter-regional policy environment difference" is the unavoidable problem of labor mobility. But the topic has never received more attention. The second contribution is to examine the temporal and spatial differentiation characteristics of the insurance participation rate from the perspective of economic geography. The distribution of the migrant population is based on regions, and there is spatial distribution heterogeneity that affects each other between regions. The insurance participation data of the migrant population can more objectively show the insurance participation rate. Therefore, this study uses the microdata of the China Migrants Dynamic Monitoring Survey (CMDS) to conduct an exploratory spatial analysis. The third contribution lies in the use of the Geographically and Temporally Weighted Regression Model for Panel Data (PGTWR). Labor mobility and insurance participation in different cities is a spatial phenomenon. The business environment characteristics of city A not only affect the migrant population of itself but also affect the migrant population of neighboring cities and the decision to participate in insurance. The impact will vary with distance, culture, and policy, manifesting as spatial heterogeneity. However, the traditional linear regression model can only estimate the parameters "average" or "global" and cannot take into account the spatial location information contained in the data, while the commonly used Geographical Weighted Regression (GWR) is simply a large cross-sectional treatment of panel data and does not consider the transfer and conduction of spatial spillover effects in time.

The rest of the article is presented below. Section 2 will introduce the data from the China Migrants Dynamic Survey(CMDS) and the PGTWR model. Then, an

exploratory spatial analysis of the health insurance participation rate in the destination cities of China's migrant population is conducted. In Section 4, the impact of the business environment on the participation rate is estimated based on the PGTWR model. The last section summarizes the main findings and related policy recommendations.

2. Data Sources and Research Methods

2.1. Data Description and Business Environment. The microdata used in this study are derived from the Migrant Population Dynamics Monitoring Survey conducted in China from 2011 to 2018. The survey uses the 2010–2017 migrant population in 31 provinces of China as the basic sampling frame and adopts a stratified, multistage, size-proportional PPS method with a total sample size of 1.648 million people. The respondents were the male and female floating population aged 15 to 59 years old who had lived locally for a month or more, with nonlocal hukou. The tourists and students at stations, docks, and other places were excluded. We define five types of conditions, in which the working population participating in the new rural cooperative medical insurance, cooperative medical insurance for urban and rural residents, medical insurance for urban residents, medical insurance for urban workers, and public medical care in the destination cities as "participating in the medical insurance." After removing missing values and outliers from the data, 1222217 valid samples were obtained.

The definition of the business environment in this study is derived from the World Bank, i.e., the external environmental conditions faced by business activities from the start-up to the end of the process. At the microlevel, the business environment significantly affects the migration and integration intentions of the migrant population in terms of macroeconomic development, government policy competition, and infrastructure construction [15]. Given the availability of data and the homogeneity of the business environment at the provincial level, we constructed a business environment evaluation index system at the provincial level including three primary indicators of the macroeconomic, policy environment, and infrastructure as well as 11 secondary indicators, after that we used the entropy evaluation method to determine the indicator weights, and then measured the business environment of 31 provinces from 2010 to 2017. The data are obtained from the National Bureau of Statistics of China and the China Statistical Yearbook, and the index system and weights are listed in Table 1.

2.2. Geographically and Temporally Weighted Regression Model for Panel Data. The Geographically and Temporally Weighted Regression Model for panel data is a specific form of a geographically weighted regression model based on panel data, which includes cross-sectional data of different regions in different time dimensions. The model is based on the "peer effect" in economics, which emphasizes that the individual's behavior is influenced by the behavior of surrounding individuals, which can be found regularly in a certain number of peer behaviors. The model form of the PGTWR based on the peer effect is shown in the following equation:

$$\mathrm{STW}_{\{\in l\}} y_{\{\in l\}} = \mathrm{STW}_{\{\in l\}} X_{\{\in l\}} \beta l + \varepsilon_{\{\in l\}}. \tag{1}$$

In equation (1), $\{\in l\}$ denotes the set of sample areas in different periods included in the study based on certain rules when l is used as the target analysis area, i.e., in addition to the target area, other neighboring areas that have significant spatio-temporal spillover effects on the target analysis area should also be included in the sample area, and there are $\mathrm{Num}_{\{\in l\}}$ elements in the $\{\in l\}$. $\mathrm{STW}_{\{\in l\}}$ is the spatio-temporal weight matrix of dimension $\mathrm{Num}_{\{\in l\}} \times \mathrm{Num}_{\{\in l\}}$. The y and X denote the matrix of explained variables and explanatory variables of the sample area with dimensions $\mathrm{Num}_{\{\in l\}} \times 1$ and $\mathrm{Num}_{\{\in l\}} \times (k+1)$, respectively, and k is the number of explanatory variables. At this point, y and X contain not only the data information of the target analysis area but also the new information formed by the mapping of the "peers" to the sample area. In this case, the spatio-temporal weight matrix also needs to cover all the information mapped from the sample area to the target analysis area, and its elements are determined as shown in the following equation:

$$\begin{aligned} \Gamma &= \mathrm{STW}'_{\{\in l\}} \mathrm{STW}_{\{\in l\}}, \\ \mathrm{STW}_{\{\in l\}} &= \mathrm{STW}_{l,\mathrm{direct}} + \left[\mathrm{STW}_{l,\mathrm{spillover}} \mathrm{diag}\left(\mathrm{STW}_{l,\mathrm{direct}}\right)\right]. \\ &\quad * \mathrm{I}_{\mathrm{Num}_{\{\in \in l\}}}. \end{aligned} \tag{2}$$

In equation (4), Γ denotes the weighted weight matrix used for parameter estimation of the panel spatio-temporal geographic weighted regression model. $\mathrm{STW}_{\{\in l\}}$ denotes the spatio-temporal weight matrix used in mapping the information of the sample area to the target analysis area, where only the main diagonal has elements and the nonmain diagonal elements are all zero. $\mathrm{STW}_{\{\in l\}}$ is the sum of two spatial effect values, one from the spatial effect of the direct mapping of the sample area to the target analysis area and the other from the spatial effect of the indirect mapping of the sample area to the target analysis area. $\mathrm{STW}_{l,\mathrm{direct}}$ refers to the spatio-temporal weight matrix from the direct spatio-temporal effect of the sample area $\{\in l\}$ to the target analysis area l; $\mathrm{STW}_{l,\mathrm{spillover}}$ refers to the spatio-temporal weight matrix of the spatio-temporal spillover effects between the sample area $\{\in l\}$. $I_{\mathrm{Num}_{\{\in \in l\}}}$ is a $\mathrm{Num}_{\{\in l\}}$ order identity matrix, the symbol . * indicates the dot product between matrices, which takes the elements of the matrix $\mathrm{STW}_{l,\mathrm{spillover}} \mathrm{diag}(\mathrm{STW}_{l,\mathrm{direct}})$ and puts them into the main diagonal elements of $I_{Num_{\{\in \in l\}}}$ to form a new matrix. The spatio-temporal weight matrix $\mathrm{STW}_{l,\mathrm{spillover}}$ of the spatio-temporal spillover effect between the sample area $\{\in l\}$ can be referred to Fan and Hudson (2018) [16].

To ensure that the obtained spatial effect values are comparable, the adaptive bandwidths hd and hl need to be adjusted when determining the elements of the spatial weight matrix, i.e., $h = \mathrm{Max}(hd, hl)$ is used to replace the initial values of the adaptive bandwidths hd and hl, thus ensuring that the total number of sample areas included in

Table 1: Business environment evaluation index system and weights.

Primary indicators	Secondary indicators	Explanation of indicators	Direction	Weight
Macroeconomics	GDP per capita	GDP per capita	Positive	0.2902
	Average wage level	Average wage of employees	Positive	0.2534
	Consumption rate	Final consumption expenditure/GDP	Positive	0.0973
	Fixed asset investment per capita	Fixed assets/resident population	Positive	0.3438
	GDP growth rate	GDP growth rate	Positive	0.0153
Policy environment	Government intervention	Local fiscal expenditure/GDP	Positive	0.0178
	Corporate tax burden	Taxes and surcharges on main business/total profit	Negative	0.9822
Infrastructure	Urban road area per capita	Urban road area/resident population	Positive	0.1323
	Number of beds in health institutions	Number of beds in health institutions	Positive	0.2576
	Power supply capacity	Electricity consumption of the whole society	Positive	0.3177
	Freight capacity	Total cargo volume	Positive	0.2824

the analysis framework is the same for each target area. For the specific programming, h will be calculated based on the following steps: first, based on the threshold of the number of sample areas included in the analysis, to determine the number of sample areas that should be included in the analysis for each target area, and the number of neighboring areas that should be included in each sample area. Second, all distances from the sample areas to the target analysis areas $d_{n_{\{\epsilon\epsilon l\}} \longrightarrow l}$, and all distances $d_{n_o \longrightarrow n_d}$ from each sample area to its immediate neighbors included in the analysis, were calculated. Finally, h can be obtained by the following equation:

$$h = \mathrm{Max}\frac{\left(d_{\{\epsilon l\}\longrightarrow l}, d_{n_o \longrightarrow n_d}\right)}{\sqrt{-(1/\eta)\ln(se\upsilon c)}}. \tag{3}$$

Among them, η is the empirical constant in the Gaussian kernel function, which takes the value of 0.5 in this study. $se\upsilon c$ is the critical value of tolerable spatial effect on the spatial influence boundary, considering the definition domain of spatial effect (0,1) and the way the critical value of significance level is designed, this study sets $sevc = 0.05$, which means the spatial spillover effect does not exist when that between regions is less than 0.05.

Under the given spatial bandwidth and temporal bandwidth, the nearest neighboring local points incorporated in the analysis of a single local point will form a new panel, which is a subset of the initial panel containing all spatial points and all periods. At this time, all parameters and statistical properties of a single local point can be determined based on the traditional modeling logic of panel data in econometrics. At the same time, since the calculation of local point parameters and statistical properties depends on the determination of adaptive spatial bandwidth and temporal bandwidth, different adaptive bandwidths will imply different optimal models, so the problem of adaptive bandwidth preference and model selection can be determined based on the overall statistical properties of the model.

2.2.1. Parameter Estimation and Significance Assessment. The parameters of the PGTWR model and their variance estimates are $\widetilde{\beta}_l = \widetilde{S}_{y_{\{\epsilon\epsilon l\}}}$ and $\mathrm{Var}(\widetilde{\beta}_l) = \mathrm{Diag}_l[\widetilde{\Omega}(\widetilde{\beta}_l)]$, where $\widetilde{S} = (X_{\{\epsilon\ l\}}{}'\Gamma X_{\{\epsilon\epsilon l\}})^{-1}X_{\{\epsilon\ l\}}{}'\Gamma$, $\widetilde{\Omega}(\widetilde{\beta}) = \widetilde{S}\Gamma^{-1}\widetilde{S}'\widetilde{\sigma}_1^2$; the random disturbance term variance estimates are calculated by the following equation:

$$\begin{aligned}\widetilde{\sigma}_1^2 &= \left(H_{\{\epsilon\epsilon l\}} - \widetilde{H}_{\{\epsilon\epsilon l\}}\right)'\frac{\left(H_{\{\epsilon l\}} - \overline{H}_{\{\epsilon l\}}\right)}{\left(\widetilde{\upsilon}_{0,l} - 2\widetilde{\upsilon}_{1,l} + \widetilde{\upsilon}_{2,l}\right)},\\ \widetilde{\sigma}^2 &= (H - \widetilde{H})'\frac{(H - \widetilde{H})}{\left(\widetilde{V}_0 - 2\widetilde{V}_1 + \widetilde{V}_2\right)}.\end{aligned} \tag{4}$$

$\widetilde{\sigma}_1^2$ is the variance estimate of the random perturbation term for the local point l, $l = 1, 2, ..., NT$, $H_{\{\epsilon l\}} = \mathrm{STW}_{\{\epsilon l\}}y_{\{\epsilon l\}}$, $\widetilde{H}_{\{\epsilon l\}} = \widetilde{h}_l y_{\{\epsilon l\}}$, $\widetilde{h}_l = \mathrm{STW}_{\{\epsilon l\}}X_{l,\{\epsilon l\}}S$; $\widetilde{\upsilon}_{0,l} = tr(\mathrm{STW}_{\{\epsilon l\}})$, $\widetilde{\upsilon}_{1,l} = tr(\widetilde{h}_l)$, $\upsilon_{2,l} = tr(\widetilde{h}'_l\widetilde{h}_l)$. $\widetilde{\sigma}^2$ is the estimate of the variance of the random perturbation term for the model as a whole; $H = [H1, H2, ..., Hl, ..., HNT]$, $H_l = y_l$; $\widetilde{H} = [\widetilde{H}1t, n\widetilde{H}q2h, ..x, 7\widetilde{H}Cl; , ..., \widetilde{H}NT]$, $\widetilde{H}_l = A\widetilde{H}_{\{\epsilon l\}}$; A are the shares of spatial effects of all local points in the local point l analysis that have an effective effect on the local point l, $A = tr^{-1}\left\{\mathrm{STW}_{\{\epsilon l\}}\right\}\mathrm{Diag}\left\{\mathrm{STW}_{\{\epsilon l\}}\right\}$; $\widetilde{V}_i = \sum_l \widetilde{\upsilon}_{i,l}/\mathrm{NT}$, $i = 0, 1, 2$.

On the basis of determining the variance estimates of the random disturbance terms, it is easy to determine the statistics of T and its distribution for each local point parameter estimation, where $T_{l,k} = \widetilde{\beta}_{l,k}/\mathrm{Se}(\widetilde{\beta}_{l,k}) \sim T(\widetilde{\upsilon}_{0,l} - 2\widetilde{\upsilon}_{1,l} + \widetilde{\upsilon}_{2,l})$ and $\widetilde{\beta}_{l,k}$ is the Kth element of the vector of $\widetilde{\beta}_l$. At this point, because the local point parameter estimation process is still in accordance with the traditional panel data econometric method, the probability value and probability critical value still need to be calculated according to the traditional algorithm, and the same for the determination of the relevant statistical property indicators (including the goodness of fit, AICc criterion value, F distribution value, and the probability value of F distribution). As the parameters of individual explanatory variables become insignificant in the overall analysis of the model, it is no longer necessary to make a separate analysis of the parameters of the explanatory variables and the hypothesis testing process of the overall model.

2.2.2. Overall Statistical Properties and Model Selection. In the PGTWR model, the target analysis region behavior will be explained or predicted by the product of the sample region behavior and its share in the total level of spatio-temporal

spillover effects from all regions to the target analysis region. The inferred or predicted accuracy of the target analysis region behavior, based on the sample region behavior, will be able to determine the most appropriate model and, consequently, the optimal spatial bandwidth [17].

The optimal bandwidth selection for the PGTWR model can be done according to the traditional CV criterion, GCV criterion, and AICc criterion, but the calculation formula needs to be adjusted slightly, as shown in the following equation:

$$\begin{aligned} \mathrm{CV} &= (\tilde{y}_L - y_L)'(\tilde{y}_L - y_L), \\ \mathrm{GCV} &= \frac{\mathrm{CV}}{(NT - k - 1)^2}, \\ \mathrm{AIC}c &= \tilde{V}_0\left[\ln\left(\tilde{\sigma}^2\right) + \ln(2\pi) + \frac{(\tilde{V}_0 + \tilde{V}_1)}{(\tilde{V}_0 - 2 - \tilde{V}_1)}\right]. \end{aligned} \tag{5}$$

Then, the optimal adaptive spatial and temporal bandwidths are determined according to the CV, GCV, and AICc criteria for taking the minimum values. When the values of the different criteria point to a different optimal bandwidth, the overall statistical properties of the model under different spatial and temporal bandwidths can be preliminarily calculated. The selection results of the most suitable bandwidth can be determined based on the optimal results of the statistical properties of the model. The overall merits and disadvantages of the model can be judged by the relevant statistical indicators, as shown in the following equation:

$$\begin{aligned} \tilde{R}^2 &= 1 - \frac{(\tilde{y}_L - y_L)'(\tilde{y}_L - y_L)/(\tilde{V}_0 - 2\tilde{V}_1 + \tilde{V}_2)}{(y_L - \overline{y}_L)'(y_L - \overline{y}_L)/(\tilde{V}_0 - 1)}, \\ F &= \frac{(\tilde{y}_L - \overline{y}_L)'(\tilde{y}_L - \overline{y}_L)/(2\tilde{V}_1 - \tilde{V}_2)}{(\tilde{y}_L - y_L)'(\tilde{y}_L - y_L)/(\tilde{V}_0 - 2\tilde{V}_1 + \tilde{V}_2)}, \\ \tilde{\alpha} &= \alpha\frac{(k+1)}{(2\tilde{V}_1 - \tilde{V}_2)}, \\ \ln L &= -\frac{\tilde{V}_0}{2}\left[\ln\left(\tilde{\sigma}^2\right) + \ln(2\pi) + 1\right]. \end{aligned} \tag{6}$$

$\tilde{R}^2$ is the modified goodness-of-fit value in the overall model; F is the F statistic of the overall model significance, whose distribution is of the form $F(2\tilde{V}_1 - \tilde{V}_2, \tilde{V}_0 - 2\tilde{V}_1 + \tilde{V}_2)$, from which the value of the F statistic and its corresponding probability level can be calculated. The probability threshold $\tilde{\alpha}$ of the overall model must be recalculated based on the predetermined significance level α; $\ln L$ is the log-likelihood value of the overall model. $\tilde{y}_L$isthe estimated value of the explained variable,$\overline{y}_L$ isthe mean value of the explained variable, $\tilde{V}i$ isthe estimated value of the variable, and $\tilde{\sigma}^2$ isthe estimated value of the variance of the random disturbance term. In general, the larger the proportion of significant parameter estimates at all local points in the overall model, the larger the value of $\tilde{R}^2$, the more F passes the hypothesis test, and the corresponding model is likely to be the optimal model.

3. Spatial Distribution Characteristics of Health Insurance Coverage Rate of Migrant Population in the Destination

3.1. Analysis of Spatial Pattern. We take the city as the basic unit and divide the number of people participating in medical insurance by the number of surveyed people to measure the insurance participation rate and analyze the spatial distribution pattern of medical insurance participation in China's migrant population based on the natural fracture method. The results are shown in Figure 1. On the whole, the pattern of participation rates in Chinese cities is obvious: the high-value areas are mainly concentrated in the middle reaches of Yangtze River, Yangtze River Delta, Pearl River Delta, core cities in Shandong and Xinjiang, and northeast border cities of China; the rest of the high-value areas are scattered in and around economically strong cities. The participation rate of the migrant population in the above-mentioned areas is above 95%, which is consistent with the overall national medical insurance participation rate. The low-value areas are widely scattered in central, western, southwestern, and localized areas in Inner Mongolia and northeast China, which objectively reflects the reality that the overall level of urban integration willingness of the migrant population is low. No matter in the high-value areas or the low-value areas, the north is relatively concentrated, and the south is relatively scattered, reflecting the overall concentration and partial dispersion of the insurance participation rate in the inflow areas. Further, comparing the changes in the insurance participation rate from 2010 to 2017, it can be found that the willingness of the migrant population to participate in insurance has been increasing, especially in 2016 and 2017, which is closely related to the Chinese government's policy of establishing a unified basic medical insurance system for urban and rural residents introduced in 2016. In 2016, The State Council issued the Opinions on Integrating the Basic Medical Insurance System for Urban and Rural Residents. This policy allows the immigrants to enjoy the same medical insurance treatment as local residents in the destination cities and reduces the cumbersome procedures for the migrant population to enroll in insurance in other places and to return to their hometowns for reimbursement, which substantially reduces the worries of the migrant population to seek medical treatment in the inflow area. Among them, the areas with rapid growth in the participation rate are mainly concentrated in the Yangtze River basin and the more economically developed cities in the Yangtze River Delta, Pearl River Delta, and Shandong Province, indicating a high correlation between the willingness of the migrant population to blend into the city and the level of economic development in the destination area.

3.2. Spatial Association Features. We use spatial autocorrelation to study the spatial agglomeration characteristics of participation in the inflow areas of the migrant population. The calculation results for the global Moran's I are reported in Table 2. It can be seen that Moran's I is positive and highly

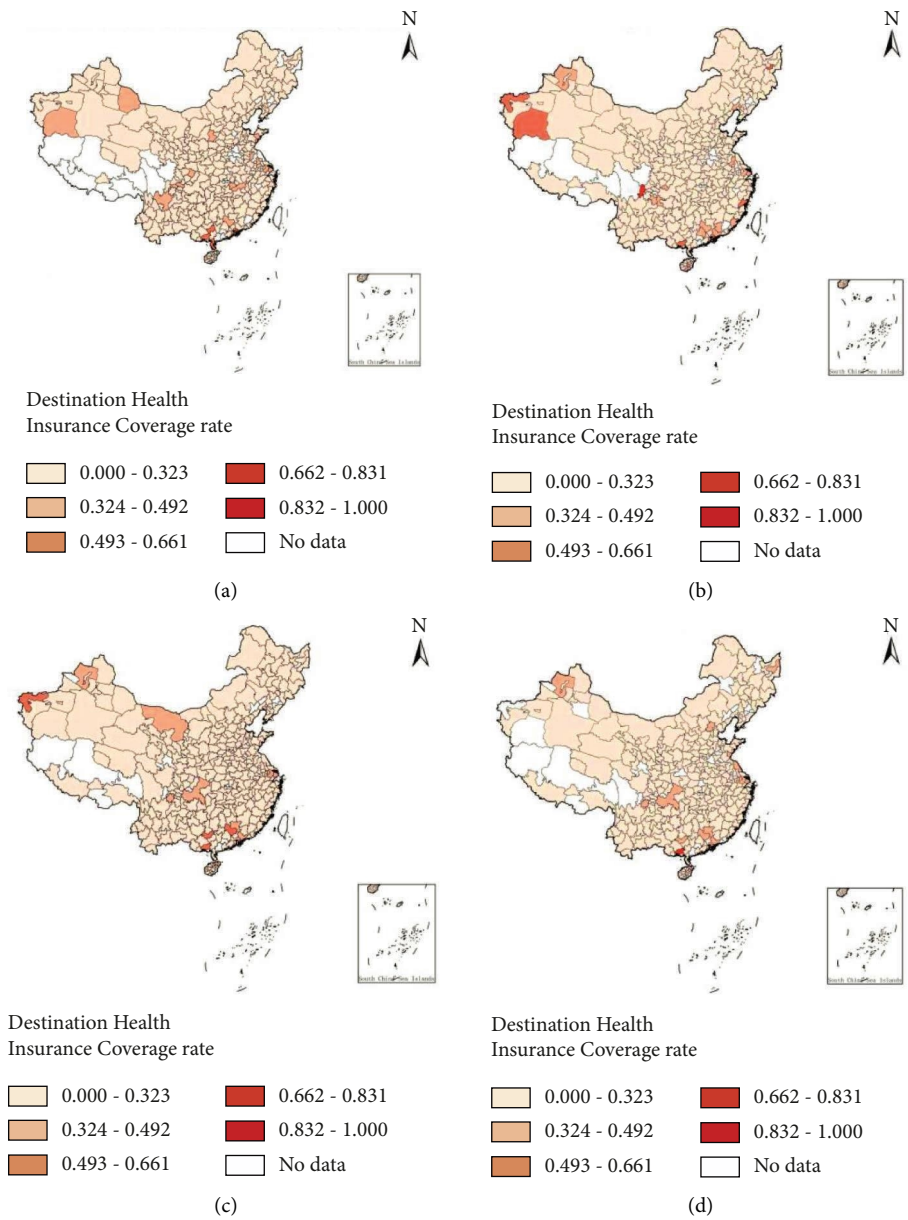

Figure 1: Continued.

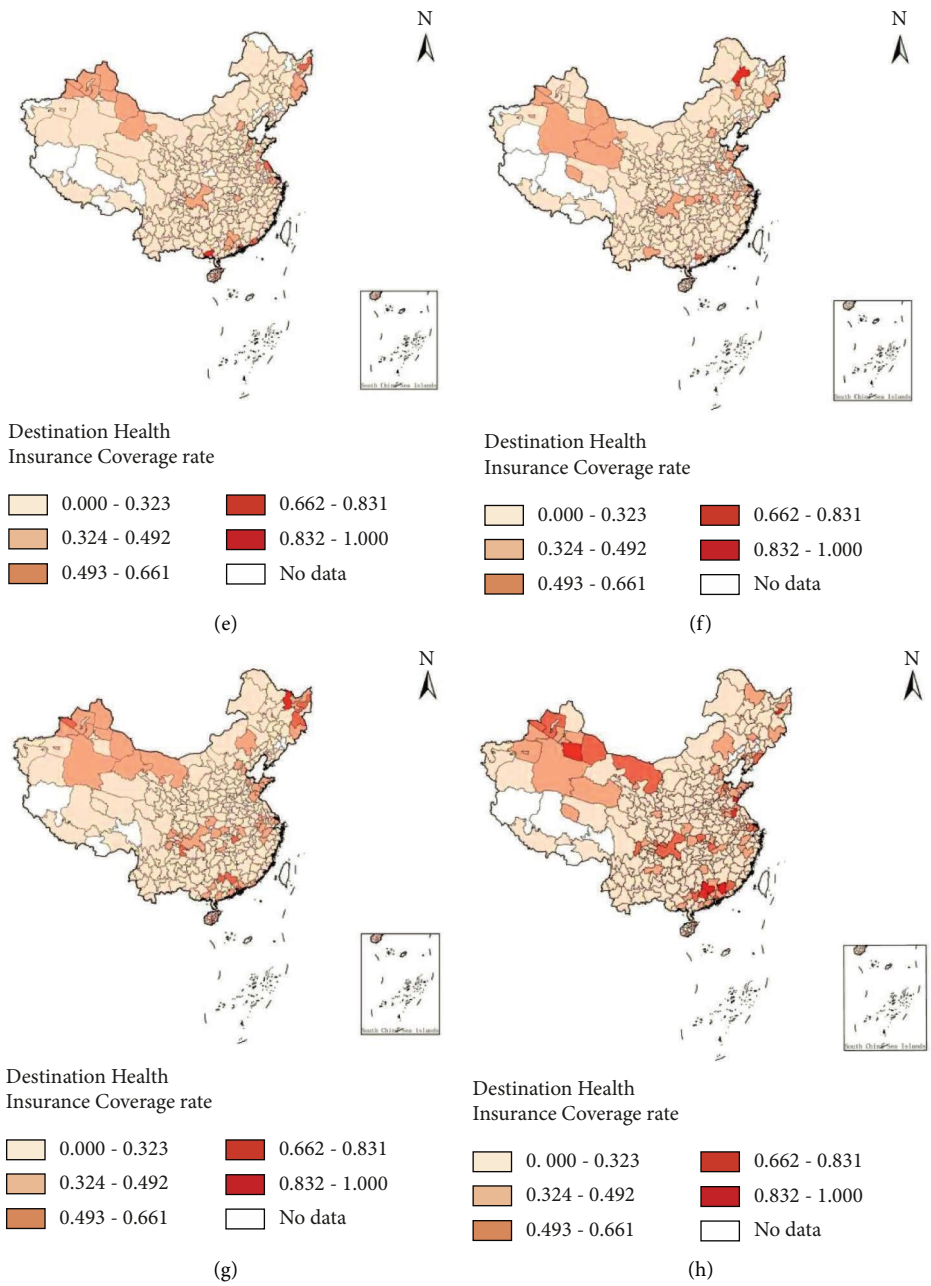

Figure 1: Spatial pattern of insurance participation rates 2010–2017. (a) Distribution map of destination health insurance coverage rate of floating population in 2010. (b) Distribution map of destination health insurance coverage rate for urban employees of floating population in 2011. (c) Distribution map of destination health insurance coverage rate of floating population in 2012. (d) Distribution map of destination health insurance coverage rate of floating population in 2013. (e) Distribution map of destination health insurance coverage rate of floating population in 2014. (f) Distribution map of destination health insurance coverage rate of floating population in 2015. (g) Distribution map of destination health insurance coverage rate of floating population in 2016. (h) Distribution map of destination health insurance coverage rate of floating population in 2017.

significant in all years, indicating that there is a significant spatial autocorrelation characteristic of the participation rate of migrant population inflow places among Chinese cities, and the areas with high participation rate have a positive impact on the increase of the participation rate in neighboring areas, thus making them neighboring. Further, it can be found that Moran`s I increased from 0.090 in 2010 to 0.138 in 2017, indicating that the spatial concentration trend of regions with similar levels of insurance participation rates is gradually increasing, and the development of insurance participation rates among cities tends to be agglomerated, showing two trends of high-speed development and low-speed development, which makes the difference in the local participation rate show a trend of magnified.

From the local spatial *LISA* agglomeration map in Figure 2, it can be seen that the spatial agglomeration characteristics of

the local area of the participation rate of the migrant population inflow are more obvious, mainly high-high agglomeration and low-low agglomeration, compared with the 2010, the cities with a high-high agglomeration of the participation rate increased significantly in 2017, and mainly concentrated in the Yangtze River Delta, Pearl River Delta, Shandong Province, and other regions, due to their high level of economic development, driving the development of surrounding cities and forming "rapid growth zone," showing new high-high agglomeration characteristics. Low-low agglomeration mainly appears along the Longhai Railway from Shangqiu to Lanzhou and along with the border cities in Tibet, forming a "lagging agglomeration zone." It can be seen that the participation rate has a more obvious spatial correlation, by this we can further analyze the spatial clustering law of the participation rate.

3.3. Spatial Clustering Characteristics. The spatial agglomeration distribution characteristics of the participation rate are identified using $Getis - Ord\ G^*$, as shown in Figure 3. It can be seen that the hotspot areas of the participation rate become scattered, while the coldspot areas become clustered: compared with 2010, the hotspot areas are more widely distributed in the Pearl River Delta, Circum-Bohai-sea Region, Circum-Yellow-Sea Region, and most of Shandong and Xinjiang. In addition, the number of extremely hot spot cities has increased significantly. The hotspot and subhotspot areas are distributed at the outer edge of the hotspot area, such as southern Hunan, southern Fujian, Anhui, and western Jiangsu. As for the cold spot area, its scope spreads westward along the Longhai Railway, including Shaanxi, Henan, Ningxia, southeastern Gansu, northeastern Qinghai, and other regions, and new cold spot areas appear in southern Tibet and Yunnan in southwest China. The cold spot area originally in the central northeast gradually disappeared.

In general, the participation rate shows a clear spatial distribution pattern, showing differences between the eastern and western regions, i.e., the participation rate in the eastern cities and coastal regions is higher, while the participation rate in the vast central and western regions and the northeastern region is lower, and this pattern maintains a high level of evolutionary stability.

4. Business Environment and Migrant Population in Insurance at the Destination—Estimates Based on PGTWR

To examine the spatial and temporal differences in the business environment affecting the participation rate, we use the PGTWR model to conduct the analysis. Traditional OLS estimation tends to ignore the spatial spillover effects, which leads to biased estimation results. Therefore, an embedded spatial econometric model is needed to ensure that the estimation process can scientifically resolve the spatial spillover effects of the business environment and labor mobility. However, the spatial econometric global model needs to incorporate data from all regions for all periods to obtain the regularity of spatial dependence, which will ignore the heterogeneity of spatial dependence among factors related to the business environment at an individual local point. Given this, this study combines the Geographically and Temporally Weighted Regression Model for Panel Data (PGTWR) with a self-coded program of MATLAB R2019a to estimate the problem of our concern.

Model selection is an important part of the analysis using PGTWR. The optimal spatial and temporal bandwidths are 31 and 4 for both the GCV-based and RSS-based criteria, and 30 and 8 for the AICc-based criterion. Since the optimal bandwidth selection results based on the GCV and CV criteria are equivalent, the optimal spatial or temporal bandwidths based on the CV criterion will not be considered in this study. Given the fact that the GCV criterion, RSS criterion, and AICc criterion point to different optimal spatial bandwidths and temporal bandwidths, we will report the overall statistical properties of the Geographically and Temporally Weighted Regression Model for Panel Data under the two optimal spatial and temporal bandwidths respectively, as listed in Table 3. Because the effective nearest neighbor local points incorporated under different bandwidth dimensions constitute the new panel data, this study also tries to calculate the overall statistical properties of the mixed effects, individual fixed effects, period fixed effects, and individual-period double fixed-effects PGTWR under the two optimal bandwidth dimensions, respectively. Since this study considers the Chinese provincial-level problem, in fact, all 31 spatial units are included and no further estimation of stochastic effect is required.

Comparing the overall statistical properties of the various possible models in Table 3, it can be found that the individual-period double fixed-effects PGTWR with an optimal spatial bandwidth of 31 and an optimal temporal bandwidth of 4 reflects relatively better statistical properties, where the significance ratio of the estimated values of the local coefficients reaches 77.85%, the modified goodness of fit reaches 0.9994, and the F-statistic can pass the hypothesis test with a significance level of 0.01. Meanwhile, the individual-period double fixed effects PGTWR with an optimal spatial bandwidth of 30 and an optimal temporal bandwidth of 8 also exhibits better overall statistical properties than the other three. In comparison, the model with an optimal spatial bandwidth of 30 and an optimal temporal bandwidth of 8 has smaller variance estimates of the stochastic disturbance term, and smaller values of the CV criterion, GCV criterion, and AICc criterion, reflecting better overall statistical properties. Therefore, in this study, the optimal spatial bandwidth and temporal bandwidth of 30 and 8, respectively, will be selected for the interpretation of the impact of the business environment on the health insurance participation rate in the destination of the migrant population.

After the optimal spatial and temporal bandwidths are selected, the relevant parameters and their statistical properties can be estimated for all local points, and the validity of the local point parameters can be diagnosed, as shown in Figure 4. Compared to the analysis based on the spatial metric global model, PGTWR gives each local point a separate analysis and incorporates only the near-neighboring local points that have an effective effect on the local point in the respective analysis process, thus making the estimation more accurate.

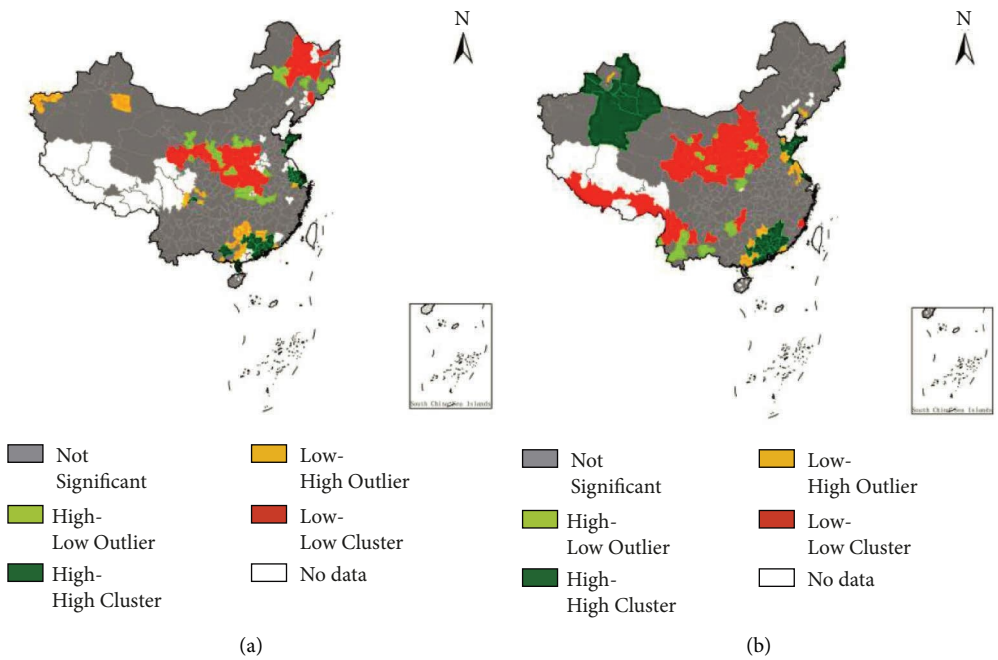

Figure 2: LISA agglomeration map. (a) LISA agglomeration map in 2010. (b) LISA agglomeration map in 2017.

Table 2: Moran's I estimated value.

Year	2010	2011	2012	2013	2014	2015	2016	2017
Moran's I	0.090	0.106	0.226	0.253	0.175	0.129	0.135	0.138
Z value	7.150	7.731	16.739	18.538	13.259	10.057	9.583	10.567
P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

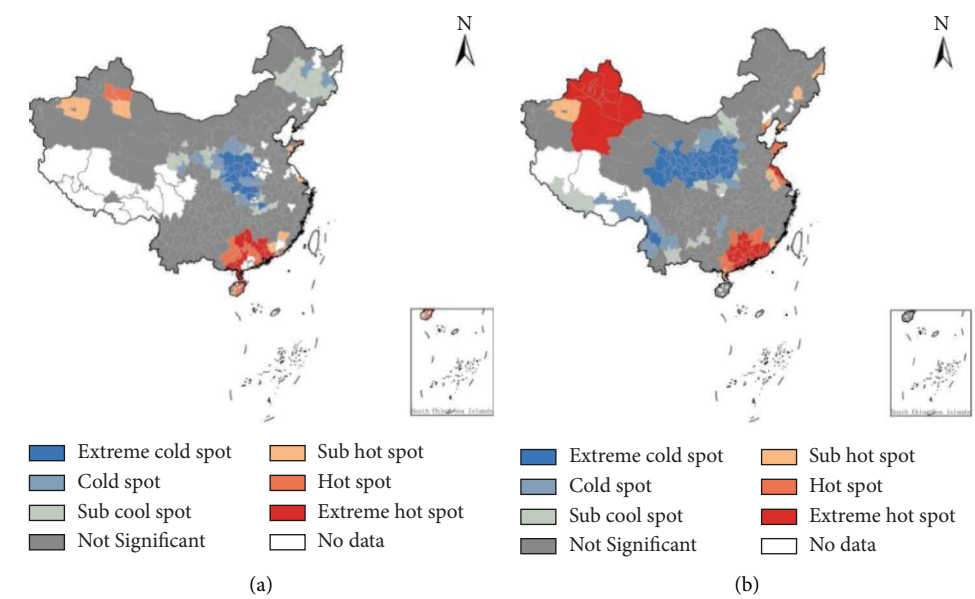

Figure 3: Distribution of hot and cold regions. (a) Distribution of hot and cold regions in 2010. (b) Distribution of hot and cold regions in 2017.

The parameter estimation process at local points still follows the traditional econometric approach for panel data, but the individual explanatory variable parameters become insignificant in the overall analysis of the model, and we will no longer need to make a separate analysis of the explanatory variable parameters and hypothesis testing process for the

Table 3: Overall statistical properties of the model.

	GCV/RSS criterion (Optimal spatial bandwidth = 31; optimal temporal bandwidth = 4)				AICc criterion (Optimal spatial bandwidth = 30; optimal temporal bandwidth = 8)			
	Mixing effect	Individual fixed effects	Period fixed effects	Individual-period fixed effects	Mixing effect	Individual fixed effects	Period fixed effects	Individual-period fixed effects
Proportion of significance of local coefficient estimates	45.2%	55.78%	57.3%	77.85%	45.7%	77.02%	55.1%	78.66%
Sample size	248	248	248	248	248	248	248	248
Degree of freedom	86	95	87	95	116	119	116	120
Estimates of the variance of the random disturbance term	0.0345	0.0048	0.0269	0.0166	0.0280	0.0059	0.0203	0.0129
CV guidelines	2.9701	0.4531	2.3417	1.5783	3.2496	0.6975	2.3490	1.5526
GCV guidelines	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AICc guidelines	−118.58	−611.97	−183.00	−302.25	−170.19	−560.37	−253.31	−364.02
Modified goodness of fit	0.9845	0.6789	−0.2482	0.9994	0.9697	−0.1195	−2.3029	0.9994
F-statistic	7551.0	518.6	92.90	27845.0	16805.0	186.8	687.2	29235.0
Probability of the F-statistic	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***
Modified probability threshold ($\alpha = 0.01, 0.05, 0.1$)	0.019 0.095 0.190	0.024 0.121 0.242	0.026 0.128 0.256	0.028 0.141 0.282	0.083 0.416 0.833	0.023 0.113 0.226	0.157 0.787 1.574	0.027 0.136 0.273
Log-likelihood value	65.46	310.96	96.37	156.19	91.41	285.38	131.65	187.20

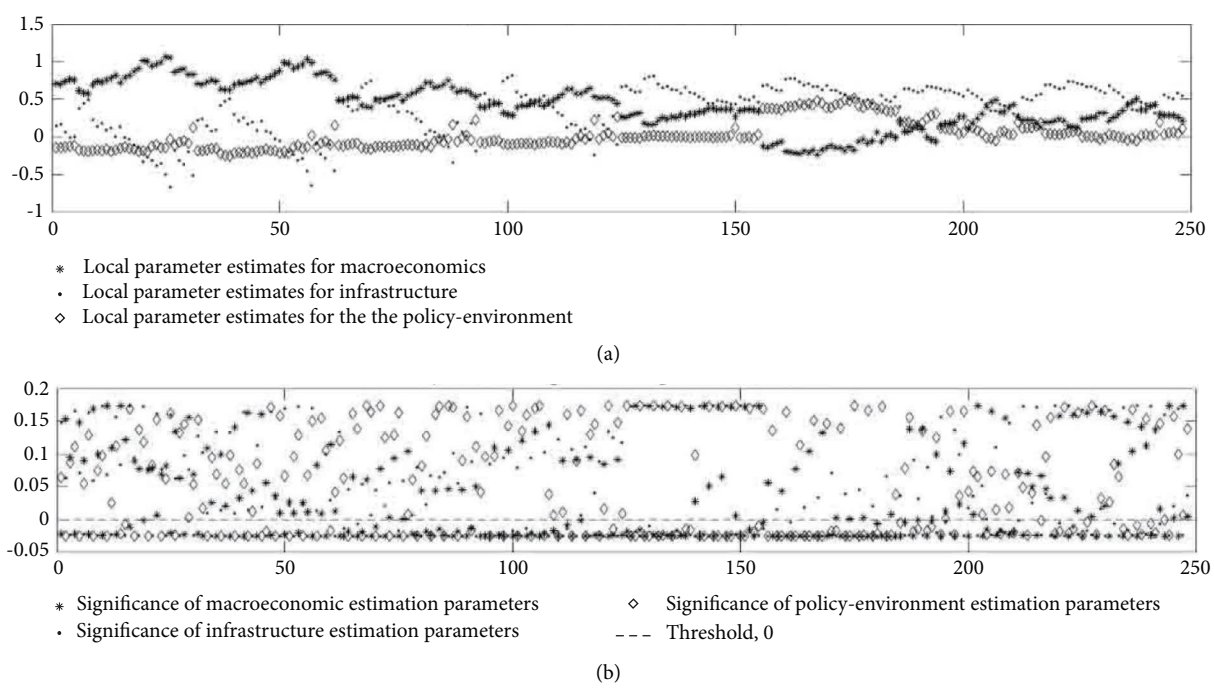

Figure 4: Parameter estimation results and validity diagnosis, (a) PGTWR parameter estimation results, (b) parameter significance diagnosis of the PGTWR.

overall model. To better explore the impact of the business environment on the participation rate, Table 4 reports the statistical characteristics of the regression coefficients of the PGTWR model and compares them with the global OLS estimation results. From the global linear regression results, the macro environment and facility environment have a positive impact on the increase of the participation rate in the inflow of migrant population, while the policy

Table 4: Estimated results.

Variables	OLS estimation	Distribution of PGTWR estimation results						
		Average value	Minimum value	1/4 quartile	Median	3/4 quartile	Maximum value	Standard deviation
Macro environment	0.618 *** (0.052)	0.4170	−0.2319	0.2279	0.4022	0.6556	1.0674	0.3111
Policy environment	0.064 * (0.069)	0.0169	−0.2563	−0.1139	−D0.0106	0.0762	0.5065	0.1820
Facility environment	−0.282 *** (0.037)	0.3530	−0.6677	0.1498	0.4585	0.5999	0.8144	0.3238
Adj. R^2	0.4385	0.9994						

environment has a negative impact, which may be related to the presence of the secondary indicator of the corporate tax burden. But of course, the linear estimation results are only the overall "average" and do not reflect the spatial and temporal variability of the impact of the business environment. The R^2 of this model is 0.4385, indicating that the model only explains less than half of the information, and its explanatory power needs to be further enhanced by embedding a spatial econometric model.

By comparing the adjusted R^2 it can be found that PGTWR has better explanatory power compared to OLS. First, the effect of macroeconomic indicators has a positive effect on the participation of the migrant population in the destination cities in the vast majority of provinces and autonomous regions in most years, except for a part of negative values in 2015. This situation accounted for 87.5% of the study period. On average, every 1% increase in the macroeconomic indicator increases the insurance participation rate by about 0.42%. Compared to the other two first-level indicators, the macroeconomic environment has a stronger positive impact, reflecting the important role of local economic development on the integration and willingness of the migrant population to participate in insurance. From a national perspective, for Jiangxi, Hubei, and Hunan in the central region; Guangxi, Chongqing, Sichuan, Guizhou, and Yunnan in the southwestern region; and Shaanxi, Gansu, Qinghai, Xinjiang, Tibet and other provinces in the western region, the average value of the estimated coefficients is higher than the national average (the national average is 0.4170). It indicates that in relatively backward regions, the marginal impact of macroeconomic development on the attractiveness of the floating population is stronger. In terms of time horizon, the estimated coefficients are *U*-shaped in all provinces, with the lowest value appearing in 2014 or 2015, which may be related to the implementation of the "National Insurance Registration Program" by the Chinese government in 2014 to integrate urban and rural social insurance information.

Second, the effect of the policy environment on the participation rate of the migrant population is relatively weak, with each 1% increase in this indicator during the research period increasing the participation rate in the destination city by about 0.017%, mainly because the policy environment for doing business is mainly expressed as the degree of "friendly" to enterprises. In the secondary indicators we selected, the weight of the corporate tax burden by the entropy method is 98%. Therefore, the tax burden of enterprises is the main component explaining the insurance participation rate. Nationally, the provinces with a negative mean estimated coefficients include Shanghai, Jiangsu, Zhejiang, Fujian, Guangdong, Guangxi, and Hainan in the coastal region and Anhui, Jiangxi, Hubei, and Hunan in the central region, which is also the main regions for undertaking manufacturing and processing industries in China, and a high enterprise tax burden will reduce the incentive of enterprises to invest and produce, which in turn affects the inflow of migrant population. The provinces with a positive mean estimated coefficients and higher than the national average of 0.016 are concentrated in western China, including Chongqing, Sichuan, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang, which indicates that enterprises in the western region still have a high potential in absorbing the migrant population.

Third, the infrastructure environment has a positive effect on the participation rate for the vast majority of the sample, with a share of 85.08% over the research period. On average, a 1% increase in infrastructure environment indicators will increase the migrants' insurance participation rate by 0.35%, indicating that public services such as education, health care, and infrastructure development have a significant impact on the mobility decisions of the migrant population in the current market-oriented labor mobility process in China. The estimated coefficients are higher than the national average of 0.3530 for provinces in Northeast China, North China, East China, and Shaanxi, Anhui, Fujian, and Jiangxi in the central region. The sample with negative effects is small and has a spatial and temporal concentration trend, with only four provinces having negative effects for more than three years, namely Sichuan (2010–2012), Yunnan (2010–2013), Tibet (2010–2013), and Xinjiang (2010–2013).

5. Conclusion

The issue of medical insurance coverage for migrants is important in the urbanization process of China. This study systematically examines the spatial distribution and evolutionary characteristics of medical insurance coverage in the destination areas of the migrants using dynamic monitoring data of China's migrant population and business environment data of 31 provinces from 2011 to 2018 and analyzes the impact of regional business environment on health insurance coverage rates in the inflow areas of the mobile population using the Geographically and Temporally Weighted Regression Model for Panel Data. The results are as follows: first, the spatial pattern of health insurance coverage in the inflow areas of the migrants in China presents a pattern of high in the east and low in the west, and the areas with high coverage rates are mainly distributed in the three major economic zones of China or provinces such as Shandong and Xinjiang. Second,

there is a significant spatial autocorrelation in the participation rate, which is manifested by the formation of clusters in cities with high and low participation rates each form agglomeration areas in the geographical space; the hotspot areas of the participation rate become dispersed, while the coldspot areas become clustered. Third, the use of panel spatio-temporal geographically weighted regression models examined not only the peer effects in the spatial neighbors but also the transfer and conduction effects of the spatial spillover effects in the sample regions, which better fits the reality of population mobility. The estimation results show that macroeconomic and infrastructure indicators in the business environment have a greater impact on the participation rate, while policy environment indicators are relatively weak, but overall the improvement of the business environment can significantly increase the probability of the migrants participating in health insurance in the destination area.

This study has important policy implications as follows: first, for a long time, the huge floating population is not only an important force for economic growth but also the main body to promote "100 million people to settle in cities and towns." The spatial distribution and evolution characteristics of the medical insurance participation rate can reflect the medical insurance policy of the floating population in China. The regional characteristics of the implementation are to provide a reference for the rational allocation of regional health service resources and the formulation of regionally differentiated medical insurance policies. Second, the medical security system plays an important role in alleviating the migration worries of the floating population and enhancing the urban identity. After clarifying the role of the business environment in promoting the medical insurance participation of the floating population, it can guide the local government to improve the business environment as an important measure to promote the urban integration of the migrant population. At the same time, it can also continuously remove the management barriers to the social security of the floating population and improved and expanded the medical insurance participation rate and coverage of the migrants.

Data Availability

The data used to support the findings of this study may be released upon application to the China Migrants Dynamic Survey, who can be contacted at National Health Commission P.R.China.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant no. 71863032), Zhejiang Soft Science Research Program (Grant no. 2022C35032), and Zhejiang Basic Research Funds (Grant no. 2021R002).

References

[1] K. Levecque, I. Lodewyckx, and J. Vranken, "Depression and generalised anxiety in the general population in Belgium: a comparison between native and immigrant groups," *Journal of Affective Disorders*, vol. 97, no. 1-3, pp. 229–239, 2007.

[2] J. Lindert, M. Schouler-Ocak, A. Heinz, and S. Priebe, "Mental health, health care utilisation of migrants in Europe," *European Psychiatry*, vol. 23, no. S1, pp. s114–s20, 2008.

[3] B. Li and X. Z. Lu, "[Recent progress on angiotensin III]," *Zhonghua Xinxueguanbing Zazhi*, vol. 36, no. 2, pp. 184–186, 2008.

[4] G. Nelson, K. Hayward, and S. M. Pancer, *Partnerships for Prevention: The story of the Highfield Community Enrichment Project*, University of Toronto Press, Toronto, Canada, 2005.

[5] L. Song and S. Appleton, "Social protection and migration in China: what can protect migrants from economic uncertainty," *Migration and social protection in China*, vol. 14, pp. 138–152, 2008.

[6] W. Feng, X. Zuo, and D. Ruan, "Rural migrants in Shanghai: living under the shadow of socialism," *International Migration Review*, vol. 36, no. 2, pp. 520–545, 2002.

[7] C. C. Fan, "The elite, the natives, and the outsiders: migration and labor market segmentation in urban China," *Annals of the Association of American Geographers*, vol. 92, no. 1, pp. 103–124, 2002.

[8] X. Wu and D. J. Treiman, "The household registration system and social stratification in China1955-1996," *Demography*, vol. 41, no. 2, pp. 363–384, 2004.

[9] Q. Fu and Q. Ren, "Educational inequality under China's rural-urban divide: the hukou system and return to education," *Environment & Planning A*, vol. 42, no. 3, pp. 592–610, 2010.

[10] A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, John Wiley & Sons, Hoboken, New Jersey, United States, 2003.

[11] B. Huang, B. Wu, and M. Barry, "Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices," *International Journal of Geographical Information Science*, vol. 24, no. 3, pp. 383–401, 2010.

[12] Z. Cheng, I. Nielsen, and R. Smyth, "Access to social insurance in urban China: a comparative study of rural–urban and urban–urban migrants in Beijing," *Habitat International*, vol. 41, pp. 243–252, 2014.

[13] L. Bursztyn, F. Ederer, B. Ferman, and N. Yuchtman, "Understanding mechanisms underlying peer effects: evidence from a field experiment on financial decisions," *Econometrica*, vol. 82, no. 4, pp. 1273–1301, 2014.

[14] J. I. Haidar, "The impact of business regulatory reforms on economic growth," *Journal of the Japanese and International Economies*, vol. 26, no. 3, pp. 285–307, 2012.

[15] S. Yao and K. Wei, "Economic growth in the presence of FDI: the perspective of newly industrialising economies," *Journal of Comparative Economics*, vol. 35, no. 1, pp. 211–234, 2007.

[16] Q. Fan and D. Hudson, "A new endogenous spatial temporal weight matrix based on ratios of Global Moran's I," *J. Quant. Tech. Econ*, vol. 35, pp. 131–149, 2018.

[17] Q. Fan and A. J. Guo, "A new geographically and temporally weighted regression model for panel data based on holographic mapping," *J. Quant. Tech. Econ*, vol. 38, no. 4, pp. 120–138, 2021.