

RESEARCH PAPER



## Regional differences of hepatitis B discrimination in rural China

Jun Li<sup>a</sup>, Jian Wang<sup>b</sup>, Stephen Nicholas<sup>c</sup>, Elizabeth Maitland<sup>d</sup>, and Teng Fei<sup>a</sup>

<sup>a</sup>School of Resource and Environmental Sciences, Wuhan University, Wuhan, China; <sup>b</sup>Dong Fureng Institute of Economic and Social Development; Research Center of Health Economics and Management, Wuhan University, Beijing, China; <sup>c</sup>Newcastle Business School, University of Newcastle, Newcastle, Australia; <sup>d</sup>School of Management, University of Liverpool, Liverpool, England

### ABSTRACT

**Background:** Hepatitis B (HB) is the most serious and common viral hepatitis in China. Previous studies on HB discrimination mainly have focussed on stratified analysis, and there has been no consideration of the impact of geographical-environmental factors on HB discrimination from a spatial perspective.

**Objective:** This study conducted a large nationwide village-based survey to test whether there were significant regional differences in HB discrimination, and to explore the relationship between different socio-economic geographical factors and HB discrimination.

**Methods:** The sample comprised 22618 rural adults, aged over 18 years old, from villages in seven provinces, representing central, southern, and eastern coastal regions of China. Using face-to-face interviews, we surveyed participants' discrimination against HB patients or carriers. Chi-square tests were used to analyze the effects of the region on differences in proportions between three discrimination levels (low, medium, and high). A geographical detector was used to explore the relationship between different socio-economic and geographical-environmental factors and HB discrimination.

**Results:** The distribution of HB discrimination levels across the 42 villages was statistically significant. The level of HB discrimination in the central and southern regions was generally low, and the level of HB discrimination in the eastern coastal regions was higher. Both socio-economic and geographic-environmental factors had a significant relationship with HB discrimination. Most of these relationships were not linear.

**Conclusions:** Developing the economy and accelerating urbanization did not automatically eliminate discrimination against HB. We recommend government HB publicity and education campaigns to inform the population of HB causes and effects, and strengthen the education of students so that they can have a clear and correct understanding of HB from a young age, both of which will address HB discrimination.

### ARTICLE HISTORY

Received 30 July 2020  
Revised 29 October 2020  
Accepted 17 November 2020

### KEYWORDS

Hepatitis B discrimination;  
rural adults; geographical  
detector; nonlinear

## 1. Introduction

Hepatitis B (HB) is one of the most prominent worldwide public health issues, and also the third-largest communicable disease in China. Worldwide, about 350 million people are chronic carriers of HB, and HB alone contributes to about 700,000 cancer deaths each year from hepatocellular carcinoma.<sup>1</sup> In China, there are an estimated 90 million chronic hepatitis B virus (HBV) carriers, accounting for about 7% of the total population and causing 300,000 deaths each year due to HBV infection and HBV-related diseases.<sup>2</sup> HB imposes annual direct medical expenses of RMB900 billion on China's health-care system.<sup>3</sup>

Vaccination is the most effective way to prevent HBV infection. China's HBV vaccination strategy has achieved remarkable results with more than 200 million newborn children being vaccinated. After 2005, China's vaccination policy provided all families with the HBV vaccine free to infants, which has resulted in a high absorption rate. Prior to 2005, families had to pay for HBV vaccination service fees (2002–2004) or both service fees and vaccine dose costs (1992–2001). Due to relatively low uptake rates pre-2001, in 2009 a catch-up HBV vaccination program was implemented for children born between 1994 and 2001.<sup>4</sup>

Nevertheless, general HBV vaccination for adults is not freely available and HB prevalence in rural China among adults remains problematic. Among adults, the incidence of HBV remains high, with the HBV infection rate increasing with age and the adult HBV vaccination rate dropping sharply with age.<sup>5,6</sup> Besides being directly affected by HB, those with HB suffer discrimination in employment and social relationships (including sexual partners), social stigma (including at schools, restaurants, and gyms), and loss of self-esteem.<sup>7</sup> Discrimination and social stigma of HB sufferers contributes to the spread of HB. The stigma of HB causes HBV sufferers to conceal their own health status and reduces their willingness to seek HB medical treatment, limiting the prevention and treatment of the HB population.<sup>8</sup> Since 2007, the Chinese government has taken measures to eliminate the discrimination against HBV carriers in employment and education,<sup>9</sup> but discrimination related to HBV remains widespread, and the comprehensive elimination of discrimination requires large-scale HB information campaigns.

Current research mainly focuses on the cultural values and social cognition factors that affect HB discrimination.<sup>10</sup> Various studies explored the factors that may lead to HB discrimination

based on the personal attributes of the respondents. Surveys of rural adults in three provinces in China found that high levels of discrimination toward HB sufferers were associated with low income and poor education.<sup>11</sup> Surveys of 435 chronic HB patients and 801 controls at Peking University People's Hospital found that higher stigma was associated with older age and lower education.<sup>12</sup> Studies also explored the factors that may lead to discrimination of HB sufferers from the perspective of the respondents' knowledge and perception of HB-related knowledge. Through face-to-face interviews with 903 over 18 years old rural adults who migrated to Beijing, Anli Leng et al. found that a high discrimination score occurred alongside a low mean knowledge of HBV and there was an inverse relationship between discrimination levels and HBV knowledge, especially knowledge about treatment and transmission routes.<sup>13</sup> In addition, it was speculated that the economic development and urbanization can help reduce or even eliminate discrimination against HB patients,<sup>14–16</sup> while the degree of development and urbanization can be represented by indices such as average level of education,<sup>14</sup> transport infrastructures,<sup>15</sup> and household income level.<sup>16</sup>

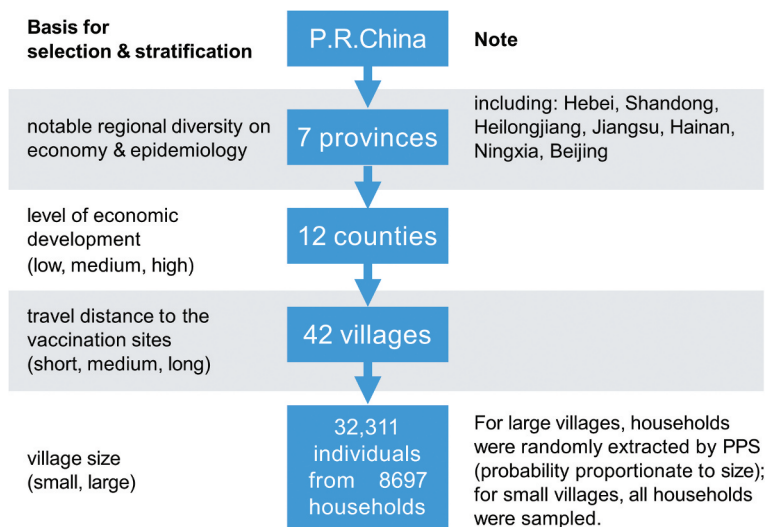
Previous studies on HB discrimination mainly relied on stratified questionnaire-based analysis, identifying influencing factors related to HB discrimination from the perspective of the respondents. These studies mainly used general linear regression to explore the relationship between discrimination factors, ruling out possible nonlinearity in the relationship. More importantly, these studies suffered from limited sample sizes, constrained modeling, and the absence of a spatial perspective. A spatial perspective refers to a research method that takes the size of a specific geographical area as its research granularity and explores the differences between geographical areas and the driving forces within each area behind any spatial differences. Compared with stratified analysis, this method can explore more regional attributes that are shared by respondents that may cause spatial differentiation of discrimination, rather than just the respondent's demographic characteristics gained through questionnaires. Importantly, spatial perspectives provide an opportunity to develop economic and targeted interventions to attenuate HB discrimination from a new standpoint.

In China, the number of HB patients in rural areas is about 11 million, which accounts for more than 50% of the total HB patients in China.<sup>3</sup> Compared to urban areas, the lack of medical and educational resources makes HB distribution in rural areas more severe. From a spatial perspective, in this research we identifies the significant difference in the level of HB discrimination across different rural regions in China; assesses the causes for these regional differences; tests the relationship between HB discrimination and the level of economy and urbanization, and, finally, proposes policy recommendations to address region-based rural HB discrimination.

## 2. Methods

### 2.1. Questionnaire and sampling method

Funded by the Research Council of Norway and implemented by Shandong University and the University of Oslo,<sup>17</sup> our cross-sectional survey collected data on HB awareness, individual vaccination history, and detailed household characteristics in seven provinces in China. Using face-to-face interviews, the survey was conducted in 42 villages from different counties in 7 provinces (Hebei, Shandong, Heilongjiang, Jiangsu, Hainan, Ningxia, and Beijing) between 2011 and 2012, with notable regional, economic, and epidemiological diversity. The counties in each province were stratified according to the level of economic development (low, medium, high), and villages in the selected provinces were stratified according to the travel distance (short, medium, long) to the vaccination sites. In small villages, all households were invited to participate, while in larger villages, households were randomly selected using household size as sampling weights (probability proportionate to size). The process of sample selection is shown in **Figure 1**. The participation rate was approximately 82%, and the predominant reason for nonparticipation was that all household members had moved to industrialized regions.<sup>17</sup> The sample consisted of 32,311 individuals (all ages included) from 8,697 households. After excluding those that had never heard of HBV and were younger than 18, the final sample consisted of 19,719 adults, aged 18–59 years (born in or before 1993), from 6,909



**Figure 1.** The sampling scheme of the cross-sectional survey on HB awareness, individual vaccination history, and detailed household characteristics.

households. Juveniles younger than 18 were not included on the basis that they have not developed stable opinions.<sup>7</sup>

## 2.2. Definition and measurement of dependent variable

The indicators of discrimination were measured by the respondents' attitudes to five related HB events:

- (1) "Are you willing to accept gifts from HB patients or carriers?"
- (2) "Are you willing to shake hands with or hug HB patients or carriers?"
- (3) "Are you willing to have dinner with HB patients or carriers?"
- (4) "Do you think parents should let their children play with HB-infected children?"
- (5) "Do you think parents should accept their child marrying a HB-infected person?"

For each question, respondents selected from three options: "Yes" (0), "Depends on the situation" (1), and "No" (2). The discrimination scores against HBV sufferers were calculated by totaling the scores of each participant from their answer to each of the five questions, which ranged between was 0 to 10 points, with scores closer to 10 points indicating the more severe the discrimination.

## 2.3. Source and definition of independent variables

Based on previous research, we selected the following choice variables that may influence the degree of discrimination against HB patients: socio-economic conditions (lower socio-economic status is linked to lower vaccine uptakes in adults, which in turn leads to more discrimination against HBV<sup>16</sup>); level of education (those with more education reported less discrimination<sup>14</sup>); medical resources (medical resource-related barriers were associated with vaccination behaviors<sup>18</sup>); natural environment (environmental factors are known to affect the immune system and may play a role as triggers of the autoimmune mosaic<sup>19</sup>); land use (migrant populations could be vulnerable to discrimination and stigmatization, in the process of urbanization in China<sup>20</sup>); and transport infrastructures (discrimination against people with infectious diseases are exacerbated in rural areas where transportation is not well-developed<sup>15</sup>).

Appendix Table 1 describes the independent variables used in the above six categories. The definition, source, and year of the socio-economic variables is shown in Appendix Table 2. All variables were collected in a spatial unit of the village, and have been classified into two categories of socio-economic and geographic-environmental data.

The definition, source and year of geographic-environmental data variables is shown in Appendix Table 3 and additional information about the instrument used to collect the data are shown in Appendix Table 4. The geographic-environmental data were mainly derived from open-source maps, meteorological bureaus, and luminous remote sensing, including elevation, road network density, transit time to the provincial capital city

from the surveyed village, distance to the provincial capital city from the surveyed village, number of medical facilities within 1 h walking distance from the surveyed village, number of medical facilities within 1h driving distance from the surveyed village, average temperature, average high temperature, average low temperature, average light intensity, the proportion of construction land and proportion of vegetation land.

There were restrictions on the years that the data on some of the independent variables were available. For example, road network density, transit time to the provincial capital city from the surveyed village, distance to the provincial capital city from the surveyed village, number of medical facilities within 1 h walking distance from the surveyed village, number of medical facilities within 1 h driving distance from the surveyed village can only be obtained to 2020 through web app, and the data on elevation, average temperature, average high temperature, average low temperature, average light intensity, the proportion of construction land, the proportion of vegetation land from open source websites can only be traced back to 2013.

## 2.4. Discrimination score and Chi-square test

Following Leng et al.<sup>13</sup> participants were categorized into three HBV-related discrimination score levels: low discrimination scores 0–3, medium discrimination scores 4–6, and high discrimination scores 7–10. Chi-square tests were used to analyze whether there was a difference between the three discrimination score levels in different regions.

## 2.5. Geographical detector

The geographical detector is a new statistical method to detect spatial-stratified heterogeneity and to identify the driving factors behind any heterogeneity.<sup>21</sup> The core idea is based on the assumption that if an independent variable has an important influence on a dependent variable, and then the spatial distribution of the independent variable and the dependent variable should be similar.<sup>22</sup> There are four functions of the geographical detector: the risk detector indicates where the risk areas are; the factor detector identifies factors that are responsible for the risk; the ecological detector discloses relative importance between the factors, and the interaction detector reveals whether the risk factors interact or lead to discrimination independently. The risk detector was employed to plot the relationship between the independent and dependent variables.

The factor detector was employed to judge the explanatory power of the independent variable to the dependent variable. The factor detector has two parameters,  $p$ -value and  $q$ -statistic. When  $p \leq 0.05$ , the relationship between the independent and dependent variables is significant. The  $q$ -statistic in geographical detector has been applied in many fields of the natural and social sciences to measure spatial-stratified heterogeneity, detect explanatory factors, and analyze the interactive relationship between variables.<sup>21</sup> Strictly within  $[0, 1]$ , the  $q$ -statistic in Equation 1 measures the association between the

independent and dependent variables, both linearly and nonlinearly:

$$q = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma^2} \quad (1)$$

where  $h = 1, \dots, L$  stands for the rank value of a certain independent variable after discretization;  $N$  stands for the number of villages; and  $\sigma^2$  stands for the variance of the dependent variable in all villages. The larger the value of  $q$ , the stronger the ability of the independent variable to explain the dependent variable.

Since the geographical detector requires that the independent variables must be discrete variables and the original independent variables collected are continuous, it is necessary to determine the appropriate discretization method and discrete rank, and perform discrete processing on the original independent variables. Five discretization methods were used 'equal,' 'natural,' 'quantile,' 'geometric,' and 'standard deviation,' with seven discrete ranks from 2 to 8. For each independent variable, a total of  $5 \times 7 = 35$  discretization schemes were generated. The discretization scheme corresponding to the largest  $q$  was regarded as the best discretization scheme for independent variables.<sup>23</sup>

## 2.6. Statistical software

To check for accuracy, all data were double-inputted using Microsoft. Chi-square tests and independent-sample t-tests were performed using SPSS 25.0 (SPSS Inc., Chicago, IL, USA) software, with  $p < .05$  considered statistically significant. The Geographical detector was implemented using R language 3.5.1 on RStudio 1.1.423.0 (RStudio Inc, Boston, MA, USA).

## 3. Results

### 3.1. Descriptive statistics for socio-demographic characteristics of participants

The socio-demographic characteristics of the study population are detailed in Appendix Table 5. The mean age was

42.48 years, with 24.69% in the 18–28 age group, and the respondents were evenly balanced between male and female respondents; 64% of the respondents were at a middle-income level, with 55.80% farmers, 28.71% migratory workers, and 5.1%-unemployed. Roughly 44% of the respondents had between 6 and 9 year of education, and only 20.57% reported that they had been vaccinated against HBV.

### 3.2. Vaccination behavior and willingness in 42 villages

Appendix Table 6 shows the vaccination behavior and willingness to vaccinate by respondents in different villages. In general, the vaccination rate for adults in rural areas in China was low, averaging only 19.62% across the 42 villages. Also, vaccination rates in different regions varied significantly, ranging from 8.72% to 47.4% ( $p \leq 0.05$ ). For people who had not been vaccinated in most villages, their willingness to vaccinate averaged only 29.89%, with the willingness to vaccinate varying significantly between different villages, ranging from 3.06% to 57.14% ( $p \leq 0.05$ ). The independent-sample t-test ( $0.136 > 0.05$ ) indicates that there was no significant difference in the mean value of HB discrimination between people vaccinated and not vaccinated in different villages.

### 3.3. Distribution of discrimination level in 42 villages

Appendix Table 7 displays the average level of HB discrimination and the number in each of the three HB discrimination groups for each village. Using the data in Appendix Table 7, we plotted the spatial distribution of the discrimination level in the 42 villages. In Figure 2, each pie chart represents a village, the size of the pie chart represented the average value of the discrimination level, with the blue, yellow, and red in the pie chart represented the proportion of low, medium, and high discrimination levels of HB in the village.

Figure 2 shows that there were significant differences in the mean value of HB discrimination across different villages. The level of HB discrimination in the central and southern villages was generally low, while the level of HB discrimination in the

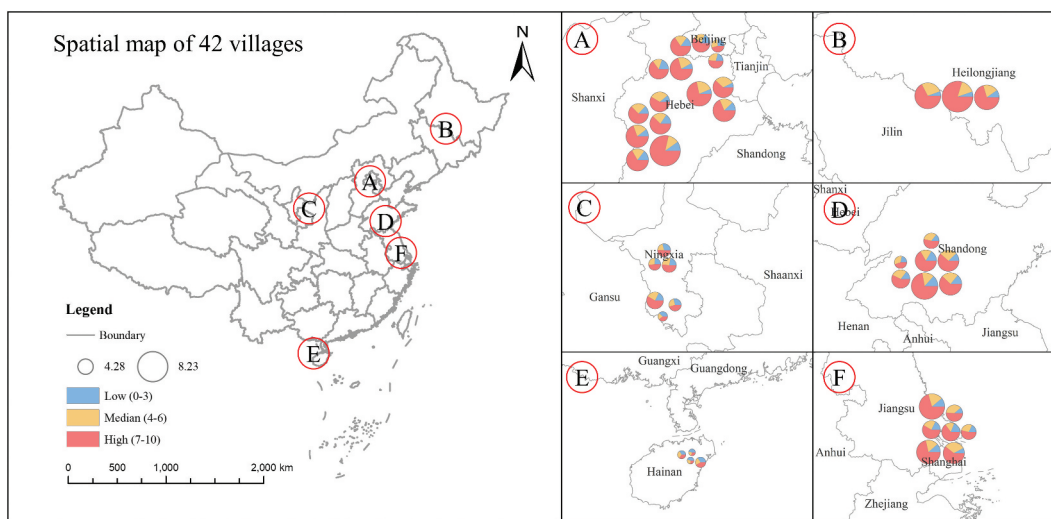


Figure 2. Spatial distribution of discrimination across 42 villages.



eastern coastal villages was high. The high level (discrimination scores: 7–10) of HB discrimination accounted for the largest proportion of all types of discrimination in most villages, indicating that most people in these villages expressed severe discrimination against HB sufferers. The proportions of the three HBV-related discrimination score levels (low, medium, high) varied across different villages.

### 3.4. Statistical significance on the difference of discrimination levels among villages

The Chi-square test, with a probability  $p$ -value less than 0.05, indicates that there was a significant difference in the distribution of HB discrimination levels across the 42 villages.

### 3.5. The relationship between the mean value of HB discrimination in different villages and the geographic-environmental and socio-economic factors

The  $p$  value,  $q$  value, optimal discretization method, and discrete rank of each independent variable calculated by the factor detector is displayed in Appendix Table 8. Appendix Table 8 shows that there was a significant relationship between certain independent variables and the mean value of HB discrimination. For the independent variables with  $p \leq 0.05$ , their relationship with the mean value of HB discrimination was plotted by the risk detector in Figure 3.

According to the curve change trend in Figure 3, the independent variables were classified into four groups. The four groups were classified arbitrarily, and the grouping was for the convenience of describing the results.

**Group 1:** The general trend of the image saw the independent variable increase, the mean value of HB discrimination decreased rapidly, but rebounded suddenly, before gradually decreasing. This group contained GDP per capita, road

network density, average light intensity, and the percentage of middle school students in 10000 people.

**Group 2:** The general trend of the image saw the independent variable increase; then, the mean value of HB discrimination slowly decreases. This group comprised average temperature, average high temperature, and average low temperature.

**Group 3:** The general trend of the image saw the independent variable increase, then the mean value of HB discrimination rises slowly before declining rapidly. This group contained the proportion of vegetation land.

**Group 4:** The general trend of the image saw the independent variable increase, with the mean value of HB discrimination rising rapidly and then slowly decreasing. This group contained the number of beds in medical institutions.

## 4. Discussion

Previous studies on the influencing factors of HB discrimination have relied on hierarchical studies based on questionnaire data. Our study used village-specific data as the research scale to explore the regional differentiation of HB discrimination and identify the relationship between different socio-economic and geographic-environmental factors and HB discrimination.

The independent variables in Group 1 reflected the degree of urbanization and development. The geographical detector shows an overall weak negative correlation trend between education level and HB discrimination, which is consistent with previous research,<sup>11</sup> but our model shows a novel and interesting nonlinear pattern. Not only one indicator displayed a nonlinear pattern but all four indicators (GDP per capita, road network density, average light intensity, and the percentage of middle school students in 10000 people) show a similar pattern. This nonlinear pattern exactly reflected the process of urbanization: from accepting rumors (high discrimination), to

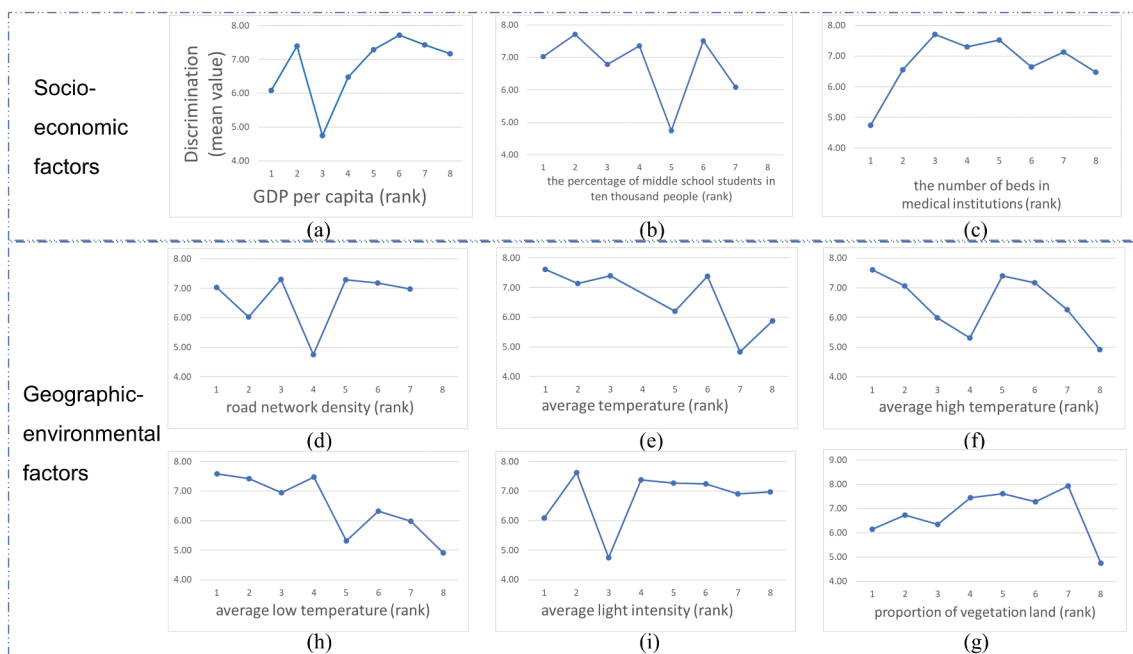


Figure 3. Diagram of the relationship between independent variables and the mean value of HB discrimination.

accepting governmental information (moderated discrimination), to exposure to commercial advertising (elevated discrimination), to finally have access to scientific knowledge (moderate discrimination). Before urbanization, people in rural poverty regions had limited access to sources of news and knew little about HB, so the level of discrimination was low. In the process of urbanization, people in remote areas suffering extreme poverty were exposed to information from outside for the first time. However, this information may not always be true. Some rumors about the hepatitis B virus have been distorted many times in the process of spreading which led to excessive speculation and inexplicable fear of HB. This argument is consistent with a study by Lijie et al. that the fear of HBV infection predominantly led to HBV-related discrimination.<sup>7</sup> The higher the region's levels of development, the greater the number, and scale of supporting medical and testing facilities and the more professional publicity and guidance, which aids people to understand HB and gradually reduce the level of discrimination. This argument is consistent with Yang and Wu who found that public education is the key to eliminating HBV discrimination and that China's health authorities should play a crucial part in HB educational programs.<sup>9</sup> There is also evidence that some hospitals and unscrupulous drug providers have exaggerated the harm of HB in pursuit of economic benefits,<sup>24</sup> which encouraged HB sufferers to seek unnecessary medical examinations and treatment procedures. Influenced by this, the level of discrimination has risen again as a result of the over-promotion of the dangers of HB. As a region achieves a higher degree of economic development, people's education level generally increases, with a better understanding of HB,<sup>12</sup> and the level of discrimination gradually decreased.

The independent variables in Group 2 showed that the higher the annual mean ambient temperature, the less people discriminated against HB carriers. This may be due to the lower incidence of HB in places where the annual mean ambient temperature is higher, such as southeastern China.<sup>25</sup>

The independent variable in Group 3 reflected that the relationship between vegetation cover and HB discrimination was not linear. A lower level of HB discrimination was usually associated with regions that have either a low proportion or extremely high proportion of vegetation. A lower proportion of vegetation represents higher urbanization, which was found to have a relationship with lower levels of discrimination against HB. On the other hand, extremely high vegetation cover represented highly inaccessible places where little knowledge about HB; thus, a low level of HB discrimination can be observed.

The independent variable in Group 4 reflected the level of regional medical standards. As the accessibility of HB testing increased, people have learnt about the level of HB infection, especially people close to them in small villages. Knowing that more people have HB leads to excessive fear of HB infection, raising the level of HB discrimination. Government information campaigns and correct guidance about the spread of HB by medical institutions reduce discrimination against HB sufferers.<sup>9</sup>

The nonlinear relationship between HB discrimination and regional development indicators showed that economic

development and accelerated urbanization cannot automatically eliminate HB discrimination. In addition to information campaigns by medical institutions, it is necessary to strengthen the education of students so that they can have a clear and correct understanding of HB from a young age. The education of adults should occur when newborns are given their compulsory HBV vaccination in the hospital. To prevent people from being misled on the causes and consequences of HBV, government regulatory agencies need to regulate more strictly illegal medical institutions and false-exaggerated information about HB. Finally, the village-level education on personal privacy needs to be strengthened, especially the protection of personal medical records to ensure that HBV sufferers remain anonymous and confidential, which limits the probability of discrimination.

## 5. Limitations

The study has a number of limitations. The real-world measurements of HB discrimination include, but are not limited to, people's attitudes to the five questions evaluated in our questionnaire. Further HB discrimination questions should be considered in the future research. The second limitation is data collection. The dependent variable derived from the questionnaire data was collected in 2011–2012, but data limitations meant some of the independent variables were measured in 2012–2019 period. For most independent variables, these differences were small and nonconsequential. Some variables (road network density and the proportion of vegetation land) could only be collected in 2019 and may have been larger. One possible solution might be constructing new data, such as extracting the road network based on remote sensing images in 2012 and calculating the actual road network density for that year on this basis. The choice of independent variables is arbitrary, considering variables in the existing literature and data availability. The causal relationship between the independent variables and the response variable needs to be investigated further in future studies. Finally, only the trend of the curve was explained. Based on the trend description of the curve, this study analyzed the reasons for the curve change. However, further research should investigate fluctuations in the curve and changes in the slope of the curve.

## 6. Conclusion

This study utilized a large nationwide HB awareness survey and granular village-specific data to reveal significant regional differences in HB discrimination and to explore the relationship between different socio-economic and geographical-environmental factors and HB discrimination. The results showed that there was a significant difference in the level of discrimination in different villages. Both socio-economic and geographic-environmental factors had a significant relationship with HB discrimination. We found that most of these relationships are nonlinear, which indicated that simply developing the economy and accelerating urbanization cannot attenuate discrimination against HB sufferers. Besides proper HB information and education campaigns, it is also necessary to eliminate false advertising and information around HB. In addition to information campaigns aimed at adults, it is

necessary to strengthen the education of students so that they can have a clear and correct understanding of HB from a young age.

## Acknowledgments

The authors thank the health workers in all the participating provinces for their support and assistance.

## Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.




## Ethics

Participants were informed that they could refuse to answer any question. The questionnaire did not ask about infection status, and no biological samples were collected. The project was approved by the Medical Ethics Committee at the Shandong University School of Medicine (Grant No. 201001052).

## Funding

This work was supported by the Norwegian Research Council (Project no. 196400/S50).

## ORCID

Jian Wang  <http://orcid.org/0000-0001-9539-2016>  
Elizabeth Maitland  <http://orcid.org/0000-0003-1551-4787>  
Teng Fei  <http://orcid.org/0000-0002-3415-1654>

## References

- Brill H, Azzopardi P, Roy M, Azzopardi P, DiGravio B, Ly L. 23 preventing perinatal transmission of hepatitis B: a quality assurance review. *Paediatrics & Child Health*. 2019. doi:10.1093/pch/pxz066.022.
- Lan Z, Zhang W, Ruan S. Modeling the transmission dynamics and control of hepatitis B virus in China. *Journal of Theoretical Biology*. 2017;262:330–38.
- WHO. Guidelines for the prevention, care and treatment of persons with chronic hepatitis B infection. 2015.
- Jia Y, Li L, Cui F, Zhang D, Zhang G, Wang F, Gong X, Zheng H, Wu Z, Miao N. Cost-effectiveness analysis of a hepatitis B vaccination catch-up program among children in Shandong Province, China. *Human Vaccines & Immunotherapeutics*. 2014;10:2983–91. doi:10.4161/hv.29944.
- Zhang L, Yan BY, Li MS, Song LZ, Lu JJ, Xu Q, Xu AQ. Preliminary analysis on the prevalence and causes of breakthrough hepatitis B virus infection among children in Shandong province, China. *Zhonghua Yu Fang Yi Xue Za Zhi*. 2013;47:933–39.
- Zhang M, Yuan Y, Mao P, Zhuang Y. Analysis on morbidity and mortality of viral hepatitis in China, 2004–2013. *Zhonghua Liu Xing Bing Xue Za Zhi*. 2015;36:144–47.
- Yu L, Wang J, Zhu D, Leng A, Wangen KR. Hepatitis B-related knowledge and vaccination in association with discrimination against Hepatitis B in rural China. *Human Vaccines & Immunotherapeutics*. 2015;12(1):70–76.
- Wu H, Yim C, Chan A, Ho M, Heathcote J. Sociocultural factors that potentially affect the institution of prevention and treatment strategies for hepatitis B in Chinese Canadians. *Canadian Journal of Gastroenterology = Journal Canadien De Gastroenterologie*. 2009;23:31–36. doi:10.1155/2009/608352.
- Yang T, Wu MC. Discrimination against hepatitis B carriers in China. *Lancet*. 2011;378:1059. doi:10.1016/S0140-6736(11)61460-8.
- Yj VDV, Van EP, Looman CW, Richardus JH. Social-cognitive and socio-cultural predictors of hepatitis B virus-screening in Turkish migrants, the Netherlands. *Journal of Immigrant & Minority Health*. 2014;16:811–21. doi:10.1007/s10903-013-9872-y.
- Yu L. Discrimination against hepatitis B in rural China: can HBV vaccination reduce the discrimination? 2015.
- Huang J, Guan ML, Balch J, Wu E, Rao H, Lin A, Wei L, Lok AS. Survey of hepatitis B Knowledge and stigma among chronically infected patients and uninfected persons in Beijing, China. *Liver International*. 2016;36(11):1595–1603.
- Leng A, Li Y, Wangen KR, Nicholas S, Maitland E. Hepatitis B discrimination in everyday life by rural migrant workers in Beijing. *Human Vaccines & Immunotherapeutics*. 2016. doi:10.1080/21645515.2015.1131883.
- Han B, Yuan Q, Shi Y, Lai W, Hou J, Jia S, Ying H, Jin C, Chan PL, Hui Z. The experience of discrimination of individuals living with chronic hepatitis B in four provinces of China. *Plos One*. 2018;13: e0195455. doi:10.1371/journal.pone.0195455.
- Qian HZ, Wang N, Dong S, Chen H, Zhang Y, Chamot E, Shi X, Gao J, Vermund SH, Shao Y. Association of misconceptions about HIV transmission and discriminatory attitudes in rural China. *AIDS Care*. 2007;19:1283–87. doi:10.1080/09540120701402814.
- Zhu D, Guo N, Wang J, Nicholas S, Wang Z, Zhang G, Shi L, Wangen KR. Socioeconomic inequality in hepatitis B vaccination of rural adults in China. *Human Vaccines & Immunotherapeutics*. 2018;14:464–70. doi:10.1080/21645515.2017.1396401.
- Zhu D, Wang J, Wangen KR. Hepatitis B vaccination coverage rates among adults in rural China: are economic barriers relevant? *Vaccine*. 2014;32:6705–10. doi:10.1016/j.vaccine.2013.06.095.
- Strong C, Lee S, Tanaka M, Juon HS. Ethnic differences in prevalence and barriers of HBV screening and vaccination among Asian Americans. *Journal of Community Health*. 2012;37:1071–80. doi:10.1007/s10900-012-9541-4.
- Molina V, Shoenfeld Y. Infection, vaccines and other environmental triggers of autoimmunity. *Autoimmunity*. 2005;38:235–45. doi:10.1080/08916930500050277.
- Gong P, Liang S, Carlton EJ, Jiang Q, Wu J. Urbanisation and health in China. *Lancet London*. 2012. doi:10.1016/S0140-6736(11)61878-3.
- Wang JF, Li XH, Christakos G, Liao YL, Zhang T, Gu X, Zheng XY. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *International Journal of Geographical Information Science*. 2010;24:107–27. doi:10.1080/13658810802443457.
- Wang J-F, Hu Y. Environmental health risk detection with GeogDetector. *Environmental Modelling & Software*. 2012;33: p.114–5. doi:10.1016/j.envsoft.2012.01.015.
- Li H, Jia P, Fei T. Geographical association between dietary tastes and chronic diseases in China: An exploratory study using crowdsourcing data mining techniques. *Acta Geographica Sinica*. 2019;74:1637–1649.
- Pan L, Pan R. Analysis on discrimination against hepatitis B virus carriers. *Medicine and Philosophy (A)*. 2005;026:77–79.
- Chun-Rong H, Han-Sheng L. Clustering analysis of hepatitis B incidence rate from 2004 to 2007 in different regions of China. *Modern Preventive Medicine*. 2009;36(20):3831–3832.

**Appendix Table 1.** Indices of six variable categories.

Variable Category	Indices
Economy	GDP per capita
Education	The percentage of middle school students in ten thousand people the average level of education
Medical condition	The number of beds in medical institutions number of medical facilities within 1 h walking distance from surveyed village number of medical facilities within 1h driving distance from surveyed village
Natural geographic factors	Elevation average (/high/low) temperature average light intensity
Transport infrastructure	Road network density transit time to provincial capital city from surveyed village distance to provincial capital city from surveyed village
Land use	Proportion of construction land proportion of vegetation land

**Appendix Table 2.** Details of socio-economic data.

Independent variable	Definition	Source	Year
GDP per capita	$\frac{\text{regional GDP (ten thousand RMB)}}{\text{regional population (ten thousand people)}}$	Statistical yearbooks	2012
The percentage of middle school students in ten thousand people	$\frac{\text{number of middle school students}}{\text{total population}} * 10^4$	Statistical yearbooks	2012
The number of beds in medical institutions	Number of beds in hospitals and health centers in the region	Statistical yearbooks	2012
The average level of education	The average education level of the surveyed subjects ((1) illiterate and semi-illiterate; (2) preschool children; (3) elementary school; (4) junior high school; (5) high school or technical secondary school; (6) university college; Undergraduate; (8) Master and above)	Questionnaires	2012

**Appendix Table 3.** Details of geographical environment data.

Independent variable	Definition	Source	Year
Elevation	Village center elevation	Google Earth	2013
Road network density	$\frac{\text{total length of regional roads}(m)}{\text{total area of area}(m^2)}$	Gaode Map	2019
Transit time to provincial capital city from surveyed village	Driving time from village center to provincial capital	Baidu Map	2019
Distance to provincial capital city from surveyed village	Driving distance from village center to provincial capital	Baidu Map	2019
Number of medical facilities within 1-h walking distance from surveyed village	Retrieve the number of medical facilities in a circular area (suppose the average person's pace is 1.25 m/s, the search radius is 4.5 km, and the search origin is village center.)	Baidu Map	2019
Number of medical facilities within 1-h driving distance from surveyed village	Retrieve the number of medical facilities in a circular area (suppose the average driving speed is 40 km/h, the search radius is 40 km, and the search origin is village center.)	Baidu Map	2019
Average temperature	The annual average temperature of the meteorological station with the shortest distance from the village	China Meteorological Administration	2013
Average high temperature	The annual average high temperature of the meteorological station with the shortest distance from the village	China Meteorological Administration	2013
Average low temperature	The annual average low temperature of the meteorological station with the shortest distance from the village	China Meteorological Administration	2013
Average light intensity	$\frac{\text{sum of pixel light values in the grid area}}{\text{total number of pixels}}$	DMSP/OLS luminous data	2013
Proportion of construction land	$\frac{\text{construction land area}}{\text{total area}}$	Google Earth	2013
Proportion of vegetation land	$\frac{\text{vegetation land area}}{\text{total area}}$	Google Earth	2013



**Appendix Table 4.** Information about the instrument used to collect the data.

Independent variable	Instrument information
GDP per capita	Web Site: statistical yearbooks URL: <a href="https://www.yearbookchina.com/index.aspx">https://www.yearbookchina.com/index.aspx</a>
The percentage of middle school students in ten thousand people	Web Site: statistical yearbooks URL: <a href="https://www.yearbookchina.com/index.aspx">https://www.yearbookchina.com/index.aspx</a>
The number of beds in medical institutions	Web Site: statistical yearbooks URL: <a href="https://www.yearbookchina.com/index.aspx">https://www.yearbookchina.com/index.aspx</a>
The average level of education	Questionnaires
Elevation	Software: LocaSpaceViewer (Windows V3.2.9) Download url: <a href="http://www.locaspace.cn/LSV.jsp">http://www.locaspace.cn/LSV.jsp</a>
Road network density	Web API: Gaode Map API URL: <a href="https://lbs.amap.com/api/webservice/summary/">https://lbs.amap.com/api/webservice/summary/</a>
Transit time to provincial capital city from surveyed village	Web API: Baidu Map URL: <a href="http://lbsyun.baidu.com/index.php?title=webapi">http://lbsyun.baidu.com/index.php?title=webapi</a>
Distance to provincial capital city from surveyed village	Web API: Baidu Map URL: <a href="http://lbsyun.baidu.com/index.php?title=webapi">http://lbsyun.baidu.com/index.php?title=webapi</a>
Number of medical facilities within 1-h walking distance from surveyed village	Web API: Baidu Map URL: <a href="http://lbsyun.baidu.com/index.php?title=webapi">http://lbsyun.baidu.com/index.php?title=webapi</a>
Number of medical facilities within 1-h driving distance from surveyed village	Web API: Baidu Map URL: <a href="http://lbsyun.baidu.com/index.php?title=webapi">http://lbsyun.baidu.com/index.php?title=webapi</a>
Average temperature	Web Site: China Meteorological Administration URL: <a href="http://data.cma.cn/">http://data.cma.cn/</a>
Average high temperature	Web Site: China Meteorological Administration URL: <a href="http://data.cma.cn/">http://data.cma.cn/</a>
Average low temperature	Web Site: China Meteorological Administration URL: <a href="http://data.cma.cn/">http://data.cma.cn/</a>
Average light intensity	Web Site: DMSP/OLS luminous data URL: <a href="https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html">https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html</a>
Proportion of construction land	Software: LocaSpaceViewer (Windows V3.2.9) Download url: <a href="http://www.locaspace.cn/LSV.jsp">http://www.locaspace.cn/LSV.jsp</a>
Proportion of vegetation land	Software: LocaSpaceViewer (Windows V3.2.9) Download url: <a href="http://www.locaspace.cn/LSV.jsp">http://www.locaspace.cn/LSV.jsp</a>

**Appendix Table 5.** Descriptive statistics for socio-demographic characteristics of participants.

Socio-demographic characteristics	Characteristic definition	n	%
Male	1 if male; 0 otherwise	9876	50.04
Female	1 if female; 0 otherwise	9852	49.96
Age 18–28	1 if aged 18–28; 0 otherwise	4868	24.69
Age 28–38	1 if aged 28–38; 0 otherwise	3852	19.53
Age 38–48	1 if aged 38–48; 0 otherwise	4146	21.03
Age 48–58	1 if aged 48–58; 0 otherwise	3415	17.32
Age 58–	1 if aged 58–; 0 otherwise	3438	17.43
Low education	1 if years of schooling ≤ 6 years; 0 otherwise	7402	37.53
Medium education	1 if 6 < years of schooling ≤ 9 years; 0 otherwise	8731	44.28
High education	1 if years of schooling > 9 years; 0 otherwise	3571	18.11
Farmer	1 if farmer; 0 otherwise	11003	55.80
Unemployed	1 if unemployed; 0 otherwise	1005	5.10
Migratory workers	1 if migratory workers; 0 otherwise	5662	28.71
Public officers	1 if public civil servants, doctors or teachers; 0 otherwise	923	4.68
Other occupations	1 if not farmer, unemployed, migratory, workers, or public officers;	1126	5.71
Without vaccination history	1 if received 0 dose of HBV vaccine; 0 otherwise	12351	62.64
Received vaccination	1 if received at least 1 dose of HBV vaccine; 0 otherwise	4056	20.57
Household income group 1	1 if household income in the bottom quintile; 0 otherwise	755	3.83
Household income group 2	1 if household income in the second lowest quintile; 0 otherwise	4272	21.66
Household income group 3	1 if household income in the middle quintile; 0 otherwise	12620	64.00
Household income group 4	1 if household income in the second highest quintile; 0 otherwise	1647	8.35
Household income group 5	1 if household income in the top quintile; 0 otherwise	234	1.19
Knowledge of symptoms 1	1 if identify 0 true symptom; 0 otherwise	7128	36.15
Knowledge of symptoms 2	1 if identify 1/2 true symptoms; 0 otherwise	5605	28.42
Knowledge of symptoms 3	1 if identify ≥ 3 true symptoms; 0 otherwise	6986	35.43
Knowledge of consequence 1	1 if identify 0 true knowledge of consequence; 0 otherwise	4697	23.82
Knowledge of consequence 2	1 if identify 1/2 true knowledge of consequences; 0 otherwise	12966	65.75
Knowledge of consequence 3	1 if identify ≥ 3 true knowledge of consequences; 0 otherwise	2056	10.43
Knowledge of treatment. route 1	1 if identify 0 true route of treatment; 0 otherwise	3211	16.28
Knowledge of treatment. route 2	1 if identify 1/2 true routes of treatment; 0 otherwise	7439	37.73
Knowledge of treatment. route 3	1 if identify ≥ 3 true routes of treatment; 0 otherwise	9069	45.99
Knowledge of transm. route 1	1 if identify 0 true route of transmission; 0 otherwise	3394	17.21
Knowledge of transm. route 2	1 if identify 1/2 true routes of transmission; 0 otherwise	4059	20.58
Knowledge of transm. route 3	1 if identify ≥ 3 true routes of transmission; 0 otherwise	12266	62.20
No fear of being infected with HBV	1 if no fear of being infected; 0 otherwise	9979	50.61
Fear of being infected with HBV	1 if fear of being infected; 0 otherwise	9740	49.39

Appendix Table 6. Vaccination behavior and willingness in 42 villages.

Village number	Vaccination behavior (N = 4056)			Not Vaccinated (N = 12351)		
	Vaccinated (%)	Not vaccinated (%)	Not know	Having vaccination willingness (%)	Having no vaccination willingness (%)	Not know
1	36(13.14)	164(59.85)	74(27.01)	38(23.17)	67(40.85)	59(35.98)
2	54(19.78)	183(67.03)	36(13.19)	55(30.05)	73(39.89)	55(30.05)
3	59(14.53)	227(55.91)	120(29.56)	49(21.59)	120(52.86)	58(25.55)
4	61(13.86)	322(73.18)	57(12.95)	59(18.32)	194(60.25)	69(21.43)
5	50(15.2)	245(74.47)	34(10.33)	140(57.14)	65(26.53)	40(16.33)
6	42(10.29)	289(70.83)	77(18.87)	104(35.99)	91(31.49)	94(32.53)
7	120(25.1)	294(61.51)	64(13.39)	88(29.93)	176(59.86)	30(10.2)
8	67(14.05)	371(77.78)	39(8.18)	115(31.0)	209(56.33)	47(12.67)
9	106(23.3)	205(45.05)	144(31.65)	49(23.9)	91(44.39)	65(31.71)
10	52(14.17)	261(71.12)	54(14.71)	103(39.46)	116(44.44)	42(16.09)
11	73(16.98)	288(66.98)	69(16.05)	42(14.58)	136(47.22)	110(38.19)
12	120(23.9)	321(63.94)	61(12.15)	108(33.64)	152(47.35)	61(19.0)
13	53(21.29)	137(55.02)	59(23.69)	63(45.99)	56(40.88)	18(13.14)
14	95(19.83)	301(62.84)	83(17.33)	118(39.2)	126(41.86)	57(18.94)
15	45(11.87)	289(76.25)	45(11.87)	77(26.64)	184(63.67)	28(9.69)
16	70(12.32)	460(80.99)	38(6.69)	209(45.43)	232(50.43)	19(4.13)
17	85(20.99)	266(65.68)	54(13.33)	73(27.44)	140(52.63)	53(19.92)
18	111(28.17)	219(55.58)	64(16.24)	57(26.03)	114(52.05)	48(21.92)
19	39(8.84)	349(79.14)	53(12.02)	118(33.81)	198(56.73)	33(9.46)
20	86(12.86)	515(76.98)	68(10.16)	115(22.33)	328(63.69)	72(13.98)
21	97(16.75)	389(67.18)	93(16.06)	74(19.02)	290(74.55)	25(6.43)
22	64(13.39)	307(64.23)	107(22.38)	85(27.69)	160(52.12)	62(20.2)
23	50(16.45)	210(69.08)	44(14.47)	21(10.0)	135(64.29)	54(25.71)
24	70(15.56)	298(66.22)	82(18.22)	39(13.09)	233(78.19)	26(8.72)
25	74(17.09)	218(50.35)	141(32.56)	11(5.05)	135(61.93)	72(33.03)
26	62(14.22)	360(82.57)	14(3.21)	11(3.06)	309(85.83)	40(11.11)
27	36(11.36)	238(75.08)	43(13.56)	22(9.24)	146(61.34)	70(29.41)
28	57(17.22)	212(64.05)	62(18.73)	40(18.87)	122(57.55)	50(23.58)
29	39(8.72)	359(80.31)	49(10.96)	179(49.86)	140(39.0)	40(11.14)
30	27(16.46)	78(47.56)	59(35.98)	26(33.33)	38(48.72)	14(17.95)
31	77(25.41)	163(53.8)	63(20.79)	38(23.31)	93(57.06)	32(19.63)
32	80(21.33)	166(44.27)	129(34.4)	39(23.49)	48(28.92)	78(46.99)
33	182(20.02)	373(41.03)	354(38.94)	72(19.3)	154(41.29)	147(39.41)
34	60(10.42)	478(82.99)	38(6.6)	239(50.0)	134(28.03)	104(21.76)
35	209(24.68)	599(70.72)	39(4.6)	278(46.41)	295(49.25)	26(4.34)
36	151(19.38)	511(65.6)	117(15.02)	139(27.2)	255(49.9)	116(22.7)
37	214(40.15)	293(54.97)	26(4.88)	111(37.88)	134(45.73)	47(16.04)
38	337(47.4)	343(48.24)	31(4.36)	98(28.57)	159(46.36)	86(25.07)
39	204(38.35)	302(56.77)	26(4.89)	139(46.03)	102(33.77)	55(18.21)
40	121(25.58)	242(51.16)	110(23.26)	117(48.35)	61(25.21)	64(26.45)
41	225(33.58)	259(38.66)	186(27.76)	124(47.88)	69(26.64)	66(25.48)
42	196(30.2)	247(38.06)	206(31.74)	104(42.11)	75(30.36)	68(27.53)

**Appendix Table 7.** HB discrimination score of each village.

Village number	Average discrimination	Low discrimination	Medium discrimination	High discrimination
1	7.50	33	53	188
2	7.27	20	84	169
3	7.70	26	89	291
4	7.25	66	86	288
5	7.18	65	50	214
6	7.50	33	87	288
7	7.45	70	89	319
8	8.20	47	55	375
9	7.50	48	91	316
10	7.19	41	111	215
11	7.21	57	104	269
12	7.32	78	110	314
13	7.48	35	59	155
14	7.00	70	134	275
15	7.34	46	100	233
16	7.89	76	82	410
17	7.36	67	68	270
18	6.58	57	126	211
19	5.99	96	148	197
20	7.83	37	181	451
21	7.72	57	114	408
22	8.23	25	72	381
23	7.63	35	52	217
24	7.80	50	83	317
25	6.95	76	103	254
26	6.75	50	163	223
27	7.37	25	100	192
28	6.50	59	94	178
29	6.97	85	71	291
30	4.88	55	51	58
31	5.53	104	72	127
32	4.30	169	110	96
33	4.28	420	219	270
34	5.98	167	148	261
35	5.31	328	183	336
36	6.79	136	188	455
37	5.94	132	144	257
38	6.13	207	143	361
39	6.36	117	140	275
40	6.94	89	94	290
41	6.46	139	185	346
42	6.05	169	184	296

**Appendix Table 8.** The results by the geographical detector.

Independent variable	P	Q	Discretization method	Discrete rank
GDP per capita	0.000	0.77	Geometric	8
The percentage of middle school students in ten thousand people	0.000	0.80	Quantile	7
The number of beds in medical institutions	0.011	0.72	Natural	8
The average level of education	0.566	0.46	Standard deviation	8
Elevation	0.732	0.60	Natural	8
Road network density	0.007	0.62	Geometric	7
Transit time to provincial capital city from surveyed village	0.999	0.26	Standard deviation	8
Distance to provincial capital city from surveyed village	0.497	0.59	Geometric	8
Number of medical facilities within 1h walking distance from surveyed village	0.987	0.18	Quantile	8
Number of medical facilities within 1 h driving distance from surveyed village	0.060	0.66	Geometric	8
Average temperature	0.000	0.77	Quantile	8
Average high temperature	0.003	0.67	Natural	8
Average low temperature	0.007	0.74	Standard deviation	8
Average light intensity	0.000	0.75	Quantile	8
Proportion of construction land	0.984	0.28	Natural	8
Proportion of vegetation land	0.000	0.83	Natural	8