**Data Sources**

The data of this study were extracted from the China Health Statistical Yearbook (2010–2018) and China Statistical Yearbook (2010-2018), which covered 31 provinces, autonomous regions, and municipalities. Those Yearbooks have collected extensive city-level information on health revenue and expenses, health service utilization, and medical facilities since 2003. Population aging was coming from the Chinese National Bureau of Statistics, whose statistical data covered various fields, from demographic, health, education, economic to industrial.

**Variables selection**

Through the literature review and consider the availability of data, we have chosen the following variables. The definition of variables and their expected signs are in Table 1. For economic variables, we chose the GDP per capita and THE as a share of GDP to examine their relationship with THE. And we chose the out-of-pocket (OOP) payments as a percentage of THE as a financial protection factor. The demographic factor is the proportion of the population aged 65 and over (POP65), which is an indication of population aging. Epidemiological variables include mortality (MOR), which represents the "baseline" health level, and the incidence of infectious diseases (DIS) which is the incidence of Class A and Class B infectious diseases. We use the number of beds per 10,000 population (BEDS) as the health service supply, for the hospital in China always good at expanding, which performed in beds expansion. And it is the bed in all kinds of health care facilities, not only the number of beds in hospitals.

**The Model**

In the literature, the dynamic panel model is frequently used to describe the longitudinal dependence of the response variable[25]. To examine the factors associated with the PTHE per capita, we consider in a ﬁrst step: a basic model that takes the following form:

 yit = yi,t-1α+ xit,1β1 + xit,2β2 + …+ xit,pβp+γi + eit （1）

i = 1,…,N, t =1,…,T

where yit is the response variable, yi,t-1 is the lagged value of the dependent variable. Xit,j is the j th covariates, γi denotes unobserved time-invariant heterogeneity, and eitis the noise term. In practice, this model could be efficiently tackled through GMM (Generalized Method of Moments) with an equivalent model:

yit = yi,t-1α+xit,1β1 +xit,2β2 + …+xit,pβp + eit  (2)

i= 1,…,N, t=1,… ,T

For () denotes the first difference. Where yit is the response variable, yi,t-1 is the 1-order lagged term, Xit,j is the j th difference of covariates[26].

Alternatively, some time-invariant country characteristics (geography, demographic) might correlate with the dependent variables, and the presence of the lagged explanatory variable yit produced auto-correlation. Hence, to cope with this kind of problem, the following difference model is usually considered in practice,

yit =yi,t-1α+xit,1β1 +xit,2β2 + …+xit,pβp+γi + eit （3）

i= 1,…,N, t=1,…,T

where γi denotes unobserved time-invariant heterogeneity, similarly[27]. This model manages to describe the longitudinal dependence of the response through the 1-order lagged terms . However, it fails to directly specify the spatial dependence of responseyi,t which could lead to confusing statistical interpretation.

In the paper, we employ the spatiotemporal panel model to analyze the complex data that simultaneously has longitudinal and spatial dependencies. Specifically, we consider the following spatiotemporal panel model:

yit = β0+x1,itβ1 +x2,itβ2 + …+xp,itβp+ eit （4）

i= 1,…,N, t=1,… ,T

where yitis the response variable, Xj,it is the j th covariates, and eit is the noise term.

In particular, we assume:

et = (eit,…,eNt), ei = (ei1,…,eiT), and , .

where is the variance of noise term, is a (N\*N) compound-symmetry structured correlation matrix with correlation coefficient , and is a (T\*T) order-1 autoregressive structured correlation matrix with correlation coefficient . That is, the noise term eit is allowed to simultaneously have cross-sectionally and serially correlation. The correlation between provinces is assumed to be equal, while the correlation in time is assumed to be order-1 autoregressive.

In practice, we adopt the generalized estimating equation (GEE) method to estimate β[29],where the associated estimating equation is

where y = (y11,…,y1T,…,yN1,…,yNT), = (1,…,p), j = (11,j,…, 1T,j,…,N1,j,,…,NT,j), and , and AB is the Kronecker product of matrices A and B. The correlation coefficients and are artificially pre-determined before we solve estimating equation. , because a misspecified correlation structure of W does not influence the consistency of coefficient estimation[28,29].

In this paper, we use PTHE per capita as dependent variables, for it offset the influence of demographics varied across provinces. PTHE per capita can be expressed as the product of these factors: POP65, MOR, GDP per capita, OOP/THE, THE/GDP, BEDS and DIS. Finial, we formalize the following dynamic regression model:

lnPTHE= β0 + β1lnGDPit + β2lnOOP/THEit + β3lnTHE/GDPit + β4lnPOP65it + β5lnMORit + β6lnDISit +β7lnBEDit + eit (5)

To offset the influence of inflation, all expenditure variables in the study were converted to 2009 yuan using the Consumer Price Index (CPI). Besides, all variables used in this analysis were log-transformed to deal with the heteroscedasticity between different variables. We assigned values to rhot and rhos with -0.2 and 0.05 respectively, which maximizes the significance of the regression coefficient. Indeed, the consistency of the regression will not be affected if the choices of these two parameters are in appropriate intervals. The estimation procedures have been performed using the software R version 3.5.3. A two-sided P value of <0.05 was considered statistically signiﬁcant.

**Sensitivity analysis**

To examine the stability of the regression results, a sensitivity analysis was undertaken to gain insight into the effect in the model of the level of uncertainty. We undertake sensitivity analysis by using variables before their log-transformed and employing a Generalized Method of Moments (GMM). We have reported the estimates and test results in appendix Tables 1 and Tables 2.