**Appendix A** Terms and Concepts

13th Five-Year Plan: One of the most important planning programs in China, mainly planning major construction projects (including environmental governance goals), productivity distribution and important proportional relationships of the national economy for the next five years.

The eastern region of China consists of the following provinces: Hebei, Liaoning, Beijing, Tianjin, Shandong, Jilin, Heilongjiang, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, Hainan.

The central region of China consists of the following provinces: Shanxi, Henan, Anhui, Hubei, Jiangxi, and Hunan.

**Appendix B** Selection of Spatial Model

We chose a spatial error model instead of a spatial autoregressive (SAR) model for the following reasons:

(1) The SAR model considers a spatially lagged dependent variable. In our case, it assumes that a city’s PMV is influenced by the PMV of its neighboring cities. According to the definition of PMV, a city's PMV is determined only by the use of fossil fuel vehicles but is not affected by the PMV of the surrounding cities. In addition, PMV is not as highly transmissible as PM2.5 because PMV’s diameter and mass are much larger than PM2.5. As a result, there should not be a spatial autoregressive effect for PMV.

(2) The results of the LM-test and Robust LM-test (Anselin, 1995) reported in **Table 3** are also biased towards the spatial error model.

**Table 3** LM and Robust LM test

|  |  |  |
| --- | --- | --- |
| **Test** | **Spatial Lag** | **Spatial Error** |
| LM-test | 3.501\* | 4.616\*\* |
| Robust LM-test | 0.048 | 1.163 |

Note: \*\*p < 0.05, \*p < 0.1

For the same reason, the spatial Durbin model (SDM) and spatial autoregressive combined (SAC) model that contain spatially lagged dependent variables were excluded.

**Appendix C** Instrumental Variable Estimation

An instrumental variable should be (1) correlated with the endogenous explanatory variable and (2) exogenous to the stochastic disturbances term. The number of rivers is regarded as an essential factor that could significantly influence the shape of a city (Bosker and Buringh, 2017), so it can assume to be correlated with UP. Rivers are (predominantly) determined by natural geographic factors and, therefore, strictly exogenous. However, the number of rivers is a cross-sectional variable invariant across time and hence cannot be (directly) used in panel models. To transform it into panel form, we multiply the number of rivers by the exchange rate, which is exogenous and correlated with UP: a lower exchange rate will expand the export trade, leading to a further concentration of resources and thus changing the spatial structure of the city (Chen et al., 2021). Consequently, we use the interaction term of the number of rivers in a city and the exchange rate as the instrumental variable.

Due to the incompatibility of instrumental variable and spatial panel models, the regression with an instrumental variable is based on the panel model without spatial effects.

**Table 4** shows the test results of the instrumental variable analysis. The coefficient of the instrumental variable is significant at the 1% level, indicating a strong correlation between the instrumental and endogenous variables. The F-test of excluded instruments is greater than 10, the Cragg-Donald Wald F statistic is larger than 16.38, suggesting that our instrumental variable is not a weak instrumental variable.

**Table 4** Test results of the instrumental variable analysis

|  |  |
| --- | --- |
| **First-stage regression** |  |
| The coefficient of instrumental variable | -0.147\*\*\* |
| F-test of excluded instruments: | 11.08\*\*\* |
| **Underidentification test** |  |
| Anderson canon. corr. LM statistic | 8.956\*\*\* |
| **Weak identification test** |  |
| Cragg-Donald Wald F statistic | 18.947 |
| Stock-Yogo critical values: 10% maximal IV size | 16.38 |
| **Overidentification test** |  |
| Sargan statistic | P=0.000 |

Note: \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.

**Table 5** shows the results of instrumental variable regression. When the instrumental variable is included, the significance of 3 variables increases and the direction (positive or negative) of the coefficients of all variables remains unchanged. This result supports the reasonableness and robustness of our panel model (Equation 3).

**Table 5** Panel Model with instrumental variable

|  |  |  |
| --- | --- | --- |
| **Dependent variable** | **PMV** | **PMV** |
| **Instrumental variable** | **NO** | **YES** |
| *UP* | -0.850\*\*\* | -4.153\*\* |
| *GDP* | 4.912\*\*\* | 5.531\*\*\* |
| *GDP2* | -0.216\*\*\* | -0.244\*\*\* |
| *Density* | -0.715\* | -1.203\*\* |
| *Industry* | -0.013 | -0.028\*\* |
| *Car* | 0.193\*\* | 0.354\*\*\* |
| *FDI* | 0.033 | 0.065 |
| *Constant* | -23.311\*\*\* | -23.510\*\* |
| Time FE | YES | YES |
| Individual FE | YES | YES |
| Observations | 510 | 510 |

Note: \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.