

Simulation analysis of environmental pollution based on parallel data model

Jiangyi Lv¹, Cherry Jiang², Yezun Qu³, and Nan Wang^{1,*}

¹Beijing Polytechnic, Beijing, P.R. China

²Stony Brook University, SUNY, Stony Brook, NY, U.S.A.

³Central University of Finance and Economics, Beijing, P.R. China

Abstract. Through model verification and algorithm correction, many factors unfavorable to environmental protection existing in the production of export products abroad are analyzed. The growth trend of export trade volume of industrial industries and regions with a high degree of openness is generally consistent with the emission trend of environmental pollutants. Environmental pollution restricts China's economic growth and brings certain negative effects on China's economic growth, The relevant data analysis suggestions for improving the impact of export product production on environmental pollution are put forward.

Keywords: Algorithm correction, Model verification, Environmental pollution, Parallel data model.

1 Instructions

This paper uses parallel data, selecting the data of different regions at different times as sample observations, which can not only analyze the impact of national macro policies on environmental pollution, but also analyze the impact of different export trade industrial structure on environmental pollution.

2 Model setting

Assuming that there is a linear relationship between export trade and environmental pollution, the relationship is expressed by the following formula:

$$y_{it} = \alpha_i + x_{it}\beta_i + u_{it} \quad \text{Of which: } y_{it} \text{ Indicates environmental pollution;}$$

x_{it} Means export trade; u_{it} Represents the influence of random variables;

α_i It is the influencing factor of macro policy; β_i It is the influencing factor of regional export trade industrial structure; $i = 1, \Lambda, n$, Number of regions $t = 1, \Lambda, T$, Is the number of time sections. For the above, there are three models:

* Corresponding author: mafannin@126.com

Model 1: if there is no individual influence and structural change on the cross section, the ordinary least squares estimation gives a consistent and effective estimation of sum. It is equivalent to putting the cross-sectional data of multiple periods together as sample data.

here $\alpha_i = \alpha_j$, $\beta_i = \beta_j$.

Model 2: variable intercept model, that is, the individual influence is different on the cross section. The individual influence is the influence of the ignored variables reflecting individual differences in the model, which is divided into fixed influence and random influence. Here $\alpha_i \neq \alpha_j$, $\beta_i = \beta_j$.

Model 3: variable coefficient model, in addition to individual influence, there is also a changing economic structure on the cross section, so the structural parameters are different on different cross sections. here $\alpha_i \neq \alpha_j$, $\beta_i \neq \beta_j$. F check is used for model verification, and there are two assumptions. H1: $y_{it} = \alpha_i + x_{it}\beta + u_{it}$, The slope and time are the same at different cross-section sample points, but the intercept is different. H2: $y_{it} = \alpha + x_{it}\beta + u_{it}$, The intercept and slope are the same at different cross-section sample points and time. If hypothesis 2 is accepted, no further test is necessary, indicating that the sample conforms to model 1. If hypothesis 2 is rejected, hypothesis 1 should be tested to determine whether the slopes are equal. If hypothesis 2 is accepted, it means that the sample conforms to model 2. If hypothesis 1 is rejected, the sample conforms to model 3.

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$$

Various F statistics are calculated as follows:

$$\bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}$$

The least squares parameter estimation of model 3 ($y_{it} = \alpha_i + x_{it}\beta_i + u_{it}$) is:

$$\hat{\beta}_i = W_{xx,i}' W_{xy,i}$$

$$\hat{\alpha}_i = \bar{y}_i - \bar{x}_i \hat{\beta}_i$$

It is called intra group estimation, where

$$W_{xx,i} = \sum_{t=1}^T (x_{it} - \bar{x}_i)' (x_{it} - \bar{x}_i)$$

$$W_{xy,i} = \sum_{t=1}^T (x_{it} - \bar{x}_i)' (y_{it} - \bar{y}_i)$$

$W_{yy,i} = \sum_{t=1}^T (y_{it} - \bar{y}_i)^2$ Then the sum of squares of the residuals of group I is

$RSS_i = W_{yy,i} - W_{xy,i}' W_{xx,i}^{-1} W_{xy,i}$, The sum of squares of residuals of model 3 is

$$S_1 = \sum_{i=1}^n RSS_i$$

Model 2 ($y_{it} = \alpha_i + x_{it}\beta + u_{it}$) The least squares parameter estimation is

$$\hat{\beta} = W'_{xx} W_{xy}$$

$$\hat{\alpha}_i = \bar{y}_i - \bar{x}_i \hat{\beta}$$

Among them, $W_{xx} = \sum_{i=1}^n W_{xx,i}$
 $W_{xy} = \sum_{i=1}^n W_{xy,i}$
 $W_{yy} = \sum_{i=1}^n W_{yy,i}$

Then the sum of squares of residuals of model 2 is $S_2 = W_{yy} - W'_{xy} W_{xx}^{-1} W_{xy}$.

The least squares parameter estimation of model 1 ($y_{it} = \alpha + x_{it} \beta + u_{it}$) is:

$$\bar{y} = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T y_{it}$$

$$\bar{x} = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T x_{it}$$

Among them, $\hat{\beta} = T'_{xx} T_{xy}$
 $\hat{\alpha} = \bar{y} - \bar{x} \hat{\beta}$

$$T_{xx} = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x})(x_{it} - \bar{x})$$

$$T_{xy} = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x})(y_{it} - \bar{y})$$

Then the sum of squares of residuals of model 1

$$T_{yy} = \sum_{i=1}^n \sum_{t=1}^T (y_{it} - \bar{y})^2$$

$$(y_{it} = \alpha + x_{it} \beta + u_{it}) \text{ is } (S_3 = T_{yy} - T'_{xy} T_{xx}^{-1} T_{xy})$$

The following conclusions can be drawn from the above:

$$(1) S_1 / \sigma_u^2 \sim \chi^2[n(T - K - 1)] ;$$

$$(2) H2 \text{ lower, } S_3 / \sigma_u^2 \sim \chi^2[nT - (K + 1)] \text{ and } (S_3 - S_1) / \sigma_u^2 \sim \chi^2[(n - 1)(K + 1)];$$

$$(3) S_1 / \sigma_u^2 \text{ and } (S_3 - S_1) / \sigma_u^2 \text{ independent}$$

$$(4) H1, S_2 / \sigma_u^2 \sim \chi^2[n(T - 1) - K] \text{ and } (S_2 - S_1) / \sigma_u^2 \sim \chi^2[(n - 1)K];$$

$$(5) S_2 / \sigma_u^2 \text{ and } (S_2 - S_1) / \sigma_u^2 \text{ independent;}$$

Therefore, the F statistic of test H2 is:

$$F_2 = \frac{(S_3 - S_1) / [(n - 1)(K + 1)]}{S_1 / [nT - n(K + 1)]} \sim F[(n - 1)(K + 1), n(T - K - 1)]$$

Obviously, if it is very small, it is very small, below the critical value, accept hypothesis 2. Where S3 is the sum of squares of residuals of the model with constant intercept and coefficients, and S1 is the sum of squares of staggers of the model with variable intercept and coefficients.

The F statistic of H1 is:

$$F_1 = \frac{(S_2 - S_1) / [(n - 1)K]}{S_1 / [nT - n(K + 1)]} \sim F[(n - 1)K, n(T - K - 1)]$$

Obviously, if it is very small, it is very small, below the critical value, receiving hypothesis 1. Where S2 is the sum of squares of residuals of the model with varying intercept and constant coefficients, and S1 is the sum of squares of residuals of the model with varying intercept and coefficients. Given the significance level, check the F distribution table to obtain the critical value, compare it with the calculated F statistics, and determine whether to accept or reject assumptions H1 and H2.

3 National macro data analysis

3.1 Local data analysis

Taking export trade volume CK and waste gas treatment investment tzgas as influence variables and waste gas emission gas as dependent variables, the impact model of export trade volume and waste gas treatment investment on waste gas emission can be constructed:

$$gas_{it} = \alpha_{gi} + CK_{it}\beta_{gi} + TZgas_{it}\gamma_{gi} + u_{git}$$

Taking export trade volume CK and wastewater treatment investment tzw as influence variables and wastewater discharge was dependent variables, the impact model of export trade volume and wastewater treatment investment on wastewater discharge can be constructed:

$$w_{it} = \alpha_{wi} + CK_{it}\beta_{wi} + TZw_{it}\gamma_{wi} + u_{wit}$$

The data can be obtained by introducing it into Eviews software. In the model, I = 1-31 represents 31 regions, t=2010-2020 represents 10 years, and the influence parameter k = 2.

3.2 Exhaust emission analysis

By using Eviews software, it is divided into three cases: fixed parameters, variable intercept and variable coefficient, and the corresponding residuals of the three exhaust emission models are obtained as follows: S3=8.94E+09, S2=1.51E+09, s1=9.62E+08. Thus, the F statistic is calculated as: F1=1.766, F2=25.71; N=31, T=9, K=2. The degrees of freedom of F statistics are F1 [60,186] and F2 [90,186], respectively. Given the significance level of 0.05, the critical value is: $F_{\alpha_1}(60, 186) = 1.391$, $F_{\alpha_2}(90, 186) = 1.338$, $F1=1.766 > F_{\alpha_1}(60, 186) = 1.391$, so H1 is rejected. 由于 $F2=25.71 > F_{\alpha_2}(90, 186) = 1.338$, H2 is rejected, so the model should adopt the form of variable coefficient.

2.3 Wastewater discharge analysis

Using the same method, the residual in the wastewater discharge model can be obtained:

S3=9.73E+09, S2=1.49E+09, s1=9.77E+08, Thus, the F statistic is calculated as:

F1=1.90, F2=21.6, Since n = 31, t = 10 and K = 2, the degrees of freedom of F statistics are F1 [60,217] and F2 [90,217], respectively. Given the significance level of 0.05, the critical value is: $F_{\alpha_1}(60, 217) = 1.38$, $F_{\alpha_2}(90, 217) = 1.327$

F1=1.90) $F_{\alpha_1}(60, 217) = 1.327$, so H1 is rejected. Since $F2=21.6 > F_{\alpha_2}(90, 217) = 1.38$, H2 is rejected, so the model should adopt the form of variable coefficient.

The parameters in the variable coefficient model are as follows :

region	C (α)	CK (β)	TZ (γ)
_BEIJING	2521.360	0.000390	9.42E-05
_TIANJIN	1895.141	0.001025	-2.55E-05
_HEBEI	12324.03	0.025342	-0.268558
_SHANXI	7343.075	0.009690	0.062775
_NEIMG	766.1471	0.046806	0.158908
_LIAONING	-341.5261	0.008205	0.007034

_JILIN	2345.642	0.014138	-0.049779
_HEILONGJ	3527.101	0.003009	0.010703
_SHANGHAI	6825.319	0.000239	-0.025852
_JIANGSU	10858.32	0.000746	-0.008144
_ZHEJIANG	8561.762	0.000853	-0.032516
_ANHUI	1740.920	0.014200	-0.039805
_FUJIAN	829.9739	0.001689	-0.005507
_JIANGXI	1642.829	0.007700	0.019794
_SHANDONG	9223.640	0.003081	-0.004900
_HENAN	7023.826	0.014026	0.008068
_HUBEI	6022.952	0.007847	-0.033478
_HUNAN	1465.758	0.003857	0.091125
_GUANGDONG	7244.503	0.000380	-0.018633
_GUANGXI	4712.906	0.017559	-0.059533
_HAINAN	-485.1364	0.012252	-0.003381
_CHONGQING	-107.2641	0.013079	0.083548
_SICHUAN	6071.921	0.008335	-0.020031
_GUIZHOU	2178.321	0.045941	-0.076856
_YUNNAN	2042.451	0.015949	-0.051322
_XIZANG	21.47971	-0.000225	-0.008724
_SHANXI1	473.7128	0.013478	0.029950
_GANSU	3055.328	-0.013287	0.081271
_QINGHAI	327.2994	0.027809	0.486375
_NINGXIA	568.7707	0.037604	-0.020456
_XINJIANG	2471.959	0.002402	0.006825

Therefore, our government should pay more attention to environmental governance, increase investment in environmental pollution governance, improve the operation and management level of environmental pollution governance facilities, standardize the operation market order of environmental pollution governance facilities, and ensure the sustainable development strategy. The specific measures may include: establishing a high-level environmental pollution monitoring network covering the whole country to monitor the degree of environmental pollution of production-oriented enterprises in real time. Establish environmental protection funds at the central and local levels. Special funds for environmental protection should also be set up separately in the financial budgets of all localities, cities and counties. The growth of the scale of environmental protection funds should be linked to the growth rate of industrial production; The use of environmental protection funds should be widely supervised by the society. The environmental protection department should regularly report the degree of environmental pollution and treatment effect in the region to the people's Congress at the same level and the public, and put forward specific rectification suggestions.

This work is supported by Ministry of Education Foundation (Grant NO. SGXC03080201) and Horizontal project (Grant NO. 2019H014-001).

References

1. National Bureau of statistics of the people's Republic of China, China Statistical Yearbook, China Statistics Press, 2000-2020

2. Yongguang Yu. The Analysis of Model on Population Growth with Stage-structure in the Polluted Environment [J]. *Acta Mathematicae Applicatae Sinica(English Series)*,2006(5)
3. Feng Xiao and Huang Bo and Li Yuyu. R&D investment in new energy vehicles with purchase subsidy based on technology adoption life cycle and customers' choice behaviour[J]. *IET Intelligent Transport Systems*, 2020, 14(11) : 1371-1377.
4. Matarazzo Michela et al. Digital transformation and customer value creation in Made in Italy SMEs: A dynamic capabilities perspective[J]. *Journal of Business Research*, 2021, 123 : 642-656.
5. Udokporo Chinonso and Anosike Anthony and Lim Ming. A decision-support framework for Lean, Agile and Green practices in product life cycle stages[J]. *Production Planning & Control*, 2021, 32(10) : 789-810.
6. Tarhini Ali et al. Investigating the Situated Culture of Multi-Channel Customer Management: A Case Study in Egypt[J]. *Journal of Global Information Management (JGIM)*, 2021, 29(3) : 46-74.