Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

p-ISSN: 0972-6268
e-ISSN: 2395-3454

Vol. 18 | No. 5 (Special Issue) | pp. 1739-1745 | 2019

Original Research Paper | Open Access

Air Treatment Effect Assessment for Improving Vehicle Emission Standards: Counterfactual Analysis Based on Machine Learning

Yunan Zhang*, Bilin Xu† and Haroon Rashid*****

*Tianjin University of Finance and Economics, Complex Network System and Innovation Research Center, Tianjin 300022, China

**Department of Economics and Management, Tianjin Electronic Information Vocational and Technical College, Tianjin 300350, China

***Department of Civil Engineering, Khwaja Fareed University of Engineering and Information Technology, Pakistan

†Corresponding author: Bilin Xu

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com

Received: 30-08-2019
Accepted: 31-10-2019

Key Words:
Air treatment
Motor vehicle emission standards
Counterfactual analysis
Machine learning

ABSTRACT

Automobile exhaust has become an important source of urban air pollution. Improving vehicle emission standards is one of the key measures to control air pollution. This paper takes Tianjin's implementation of the "National V" motor vehicle emission standard as an example. The study found that this policy is an indispensable condition for improving air quality, which helps to reduce carbon monoxide (CO) and nitrogen dioxide (NO_2) in the atmosphere, but the effect on the treatment of atmospheric pollutants such as particulate matter (PM2.5) and respirable particulate matter (PM10) is not significant. It can be seen that while continuously improving the emission standards of motor vehicles, it should also cooperate with the improvement of public transportation systems, the development of new energy vehicles and alternative fuels, and the targeted regulation of air pollution control measures in other high-energy-consuming industries.

INTRODUCTION

Since the reform and opening up, the number of Chinese residents' car ownership has shown a rapid growth. According to data released by the National Bureau of Statistics, China's auto sales have been ranked first in the world for nine consecutive years. It has gradually become the largest sales market for global autos from small auto countries. By the end of 2017, the number of civilian vehicles in the country was 217.43 million, an increase of 11.8%. However, while national travel is becoming more and more convenient, large-scale motor vehicle exhaust emissions have gradually become an important cause of urban air pollution and haze, seriously jeopardizing the health of the people (Song et al. 2007).

Since September 1, 2015, Tianjin has taken the lead in implementing the fifth phase of National Motor Vehicle Air Pollutant Emission Standards (referred to as "Country V"), further reducing the limits of pollutants in motor vehicle exhaust. This paper takes the implementation of the National V Standard as an example. Based on the monthly data of air quality during the period from December 2013 to December 2016 in 21 cities such as Tianjin, using Lasso's counterfactual inference research method to evaluate the improvement of vehicle emission standards for governance of the effect of air pollution.

EARLIER STUDIES

In order to prevent automobile exhaust pollution, governments have introduced many urgent policies and measures. Among them, according to their different mechanisms of action, the existing vehicle exhaust gas treatment measures can be divided into two categories: "market regulation type" and "command control type". Many scholars have also conducted extensive research on this.

Market regulation type: In theory, the "market-regulated" policy can increase the cost of motor vehicles by directly increasing fuel prices or indirectly reducing the price of alternative goods, thereby reducing the use and possession of non-public vehicles to improve the environment. For example, a research studied the impact of gasoline prices on automobile demand based on US motor vehicle registration data from 1997 to 2005. It was found that the increase in fuel cost can significantly reduce the demand for motor vehicles with high fuel consumption (Li & Haefen 2009). In addition, Chinese scholars found that the fuel consumption of private motor vehicles with a large proportion is not affected by price factors (Penghui & Ruobing 2015). In summary, the air control effect of the "market regulation type" policy is not significant, and it is easy to cause consumer welfare losses.

Command control type: Some scholars have pointed out that air governance by the government-led "command-controlled" policy is more effective. For example, Xiaoguang et al. (2010) used Beijing air quality data to evaluate the effect of the Olympic limit policy. During the implementation of the restricted policy, the concentration of nitrogen dioxide (NO_2) and respirable particulate matter in Beijing's atmosphere was significantly reduced. However, other scholars are sceptical, despite the rapid effect of the "command-controlled" policy, the persistence of policy effects is limited and seriously distort consumer buying behaviour (Davis 2008).

Therefore, in order to scientifically and accurately evaluate the air treatment effect of improving vehicle emission standards, and in view of the fact that the number of individuals in the reference group is far greater than the number of observable pre-existing periods, this paper combines the machine learning Lasso method with the counterfactual inference method, conducted individual selection and counterfactual prediction of control groups.

RESEARCH DESIGN

Principle of counterfactual result estimation: Suppose there are n+1 cities, and each city has a total of T balanced panel observation data for air quality; assume that city 1 is at $T_0(1 < T_0 < T)$. During the period, the implementation of the National V policy began, and the other n regions did not implement the National V policy. So, order $D_{it} = 1$ representing city*i*in*t*. The period was affected by the intervention of the National V policy, $D_{it} = 0$. Indicating individual *i*in*t*. Period is not affected by National V policy intervention.

$$D_{it} = \begin{cases} 1, & i = 1, t \geq T_0 + 1 \\ 0, & else \end{cases}$$

Replace $Y_{it}(1)$ with $Y_{it}(0)$ representing the city*i*in*t*. If you accept policy intervention and if you do not accept the potential consequences of policy intervention, the two cannot be observed at the same time. Y_{it} representing the actual observed results, then the relationship between the potential results and the observations is

$$Y_{it} = \begin{cases} Y_{it}(0) & D_{it} = 0 \\ Y_{it}(1) & D_{it} = 1 \end{cases}$$

Therefore, the treatment effect of the treatment group individuals who accept the policy intervention of the recipient country can be expressed as follows:

$$\tau_{1t} = Y_{1t}(1) - Y_{1t}(0), \ t = T_0 + 1, T_0 + 2, ..., T \tag{1}$$

Since the first city implemented the National V policy, $t \geq T_0$ period, potential results can be observed $Y_{1t}(1)$, but cannot observe the potential consequences if it is not subject to policy intervention $Y_{1t}(0)$. In order to estimate the counterfactual results of City 1, $Y_{1t}(0)$ can be represented by the following model:

$$Y_{it}(0) = \mu_i + b_i' f_t + u_{it} \quad i = 1, ..., J+1; \ t = 1, ..., T \tag{2}$$

Where, μ_i indicates individual *i* fixed effect, f_t is a common factor that represents unobservable, b_i indicates the corresponding factor load, u_{it} is a random error term, and $E(u_{it}) = 0$.

The key to estimating the counterfactual results by regression synthesis is that all individuals are affected by the time-varying common factor, and this relationship will be maintained between the cross-section individuals. The observations of the control group interventions are estimated to be the opposite of the intervention group. The following regression model can be established

$$Y_{it}(0) = \beta_1 + Y_{-1,t}' \beta_{-1} + \xi_{1,t} \tag{3}$$

Where, $Y_{-1,t} = (Y_{2t}(0), ..., Y_{J+1,t}(0))' = (Y_{2t}, ..., Y_{J+1,t})'$,Regression coefficient $\beta_{-1} = (\beta_2, ..., \beta_{J+1})'$ In order to synthesize the weights of the control with the control group individual, specifically, first use before intervention ($t = 1, \ldots T_0$), the data estimate (3) is obtained $\hat{\beta}_1$ with $\hat{\beta}_{-1}$. If the treatment group individual does not receive policy intervention, then after the intervention ($t = T_0 + 1, T_0 + 2, \ldots T$) the relationship should also be maintained. Therefore, the counterfactual results of the individual can be processed using the formula (3) prediction. The prediction model is as follows:

$$\hat{Y}_{1t}(0) = \hat{\beta}_1 + Y_{-1,t}' \hat{\beta}_{-1} \ , \quad t = T_0 + 1, ..., T \tag{4}$$

Control group selection: The key to ensuring that formula (4) can better predict counterfactual results is to select which individuals enter the control group. In fact, the number of individual control groups entering the model is not as good as possible. More control group individuals mean a greater loss of degrees of freedom, which in turn reduces the estimation accuracy. Therefore, this paper uses the machine learning Lasso method to control the selection of individual individuals. Without loss of generality, rewrite equation (3) into a normal regression model, as follows

$$Y_t = X_t' \beta + u_t \quad (t = 1, ..., T_0) \tag{5}$$

Thus, the Lasso method is used to estimate the parameters. First, establish the objective function,

$$\sum_{t=1}^{T_0} (Y_t - X_t' \beta)^2 + \lambda \sum_{i=2}^{N+1} | \beta_i | \tag{6}$$

Where, β_i is the coefficient of all the independent variables in the initial model (Eq. (5)), $\sum_{i=2}^{J+1} | \beta_i |$ for L_1 norm, as a punishment constraint. λ is a non-negative adjustment parameter, i.e. when λ is fully large, it can shrink some coefficients to zero, thus playing the role of controlling individual selection.

Second, use the k-fold cross-validation method to determine λ value. The specific steps are as follows: First, the initial sample data is segmented into k sub-samples, one of the sub-samples is selected as the data of the verification model, and the remaining k-1 samples are used as the training set. Second, the cross-validation is repeated k times, and the mean square error of each verification set sample is obtained. Third, the k-group mean square error mean is the smallest λ, the value is what we want.

Then, in the determination of λ, after the value of the formula, let equation (6) reach the minimum value β Lasso estimator $\hat{\beta}$.

Finally, the Lasso estimate $\hat{\beta}$, substituting (4) estimate the counterfactual results of the treatment group $\hat{Y}_{1t}(0)$,

Where, $t = T_0 + 1, T_0 + 2, ..., T$

(3) Evaluation of policy effects

Thus, the counterfactual result estimate $\hat{Y}_{1t}(0)$, bring into the formula (1), we can get the estimated amount of policy processing effect:

$$\hat{\tau}_{1t} = Y_{1t}(1) - \hat{Y}_{1t}(0),\ t = T_0 + 1, T_0 + 2, ..., T \quad (7)$$

Further, the estimated average effect of the intervention period is:

$$\hat{\tau}_1 = \frac{1}{T - T_0} \sum_{t=T_0+1}^{T} \hat{\tau}_{1t} = \frac{1}{T - T_0} \sum_{t=T_0+1}^{T} \left(Y_{1t}(1) - \hat{Y}_{1t}(0) \right) \quad \ldots (8)$$

Empirical Tests

Data source and description: This paper uses empirical data from the panel data of 21 large and medium-sized cities from 2014 to 2016 to evaluate the effect of the "National V" policy on the improvement of air quality in Tianjin. Among

Table 1: Descriptive statistics of 21 cities' air quality indicators.

	CO		NO_2		PM2.5		PM10		AQI	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Tianjin	1.23	0.68	44	15	76	26	123	38	111	28
Chongqing	1.08	0.20	42	7	58	24	87	30	86	27
Harbin	1.04	0.31	49	15	67	44	98	50	97	47
Wuhan	1.13	0.30	51	14	72	36	108	36	106	38
Chengdu	1.13	0.24	53	9	68	32	112	46	103	33
Kunming	1.01	0.22	30	6	30	9	59	15	55	10
Lanzhou	1.39	0.54	50	14	54	17	123	36	93	20
Nanning	0.98	0.21	34	10	44	22	75	29	68	25
Yinchuan	1.18	0.44	38	12	51	19	110	27	89	18
Taiyuan	1.63	0.58	39	11	66	27	123	36	100	28
Changchun	0.91	0.27	43	7	60	31	99	35	95	29
Hefei	1.03	0.23	35	11	71	31	101	30	99	34
Nanchang	1.00	0.17	31	9	47	21	80	26	74	21
Zhengzhou	1.65	0.49	55	13	89	35	154	46	129	35
Changsha	1.02	0.24	39	13	65	28	84	23	93	32
Guiyang	0.74	0.16	28	5	41	16	66	18	64	18
Xi'an	1.76	0.64	48	12	70	38	139	52	110	38
Xining	1.43	0.62	38	12	53	16	111	31	87	18
Hohhot	1.54	0.96	42	10	43	20	106	29	86	15
Lhasa	0.75	0.31	22	8	25	9	65	26	67	14
Urumqi	1.44	0.91	53	16	68	48	135	54	110	48

Note: The data comes from China Air Quality Online Monitoring and Analysis Platform www.aqistudy.cn them, the "National V" policy can be regarded as an air governance policy experiment implemented by the Tianjin Municipal Government, with Tianjin as the processing group individual and the remaining 20 cities that have not implemented the "National V" policy as the control group individual. Therefore, after the implementation of the policy (after September 1, 2015), the air quality status of the "National V" standard was not implemented in Tianjin. After the implementation of the policy, and then the "National V" standard was implemented. After implementing, Tianjin's actual air quality situation is compared to estimate the impact of the "National V" standard on Tianjin air quality. Among them, the main measures of this policy for carbon monoxide (CO), nitrogen dioxide (NO_2), particulate matter (PM2.5), respirable particulate matter (PM10) and other air pollutants and air quality (AQI) Governance effect. Table 1 is a descriptive statistical analysis of five air quality indicators for 21 cities, where Mean represents the mean and Sd represents the standard deviation.

Inspection analysis: This paper draws on the similar displacement test method proposed by Abadie et al. (2010), that is, the individual placebo test, to test the average causal effect to determine whether there are other cities with no V standard in the country; the situation, and how likely it is to appear. The test ideas and implementation steps are as follows:

First of all, the original hypothesis is that the air governance effect of the National V standard is not significant, so that Tianjin can be regarded as a control group individual. Then, a city is randomly selected from the remaining 20 cities as the processing group individual, and the policy effect of the city is estimated. In addition, 20 cities outside Tianjin were selected as treatment group individuals, and the synthetic objects based on Lasso were used to construct the synthetic objects of the corresponding cities, thus obtaining 20 policy effect estimates. Finally, compare the policy effects of Tianjin's actual production and the policy effects of the control group's urban assumptions. If the policy effect gap between the two is large enough, then there is reason to believe that the policy effect of the National V standard is significant.

Placebo tests for CO, NO_2, PM2.5, PM10, and AQI were performed in this paper. The results are shown in Figs. 1-5. The left side of each figure shows the placebo test results of the variables of interest, the solid line shows the counterfactual results obtained by Tianjin as the treatment group, and the dotted line shows the estimated results obtained by other cities as treatment groups.

In Fig. 1, before June 2015, the gap between the CO-content changes of Tianjin and other cities was not large, but after the implementation of the National V standard, the gap between the CO-content changes of Tianjin and other cities gradually widened, and its distribution was located outside the other cities. This indicates that the National V policy significantly reduces the CO-content in the atmosphere. Similarly, as can be seen from Fig. 2, before June 2015, the difference in NO_2 content between Tianjin and other cities was not large, but after the implementation of the National V standard, the NO_2 content of Tianjin and other cities changed. The gap has gradually widened and its distribution is also outside the other cities. To a certain extent, it is shown that the reduction of CO and NO_2 content in the atmosphere is significant after the implementation of the National V standard in Tianjin.

However, as shown in Figs. 3-5, after the implementation of the National V standard, the gap between PM2.5, PM10

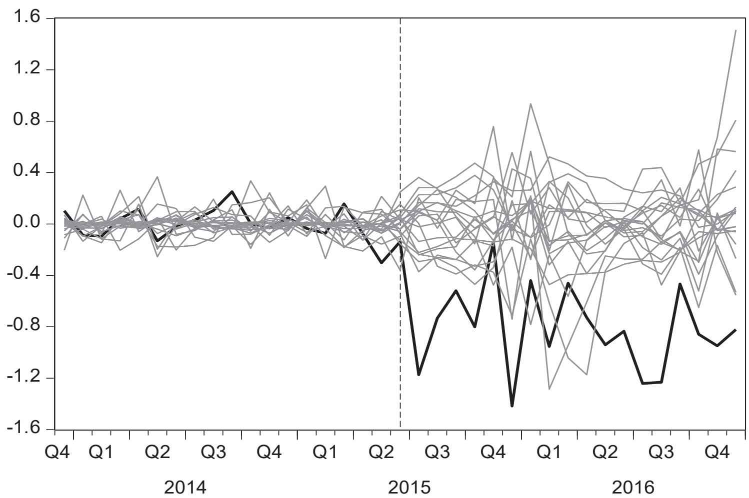

Fig. 1: Difference distribution of cities in CO.

and AQI in Tianjin and other cities is not significant, and its distribution is located in other cities. To a certain extent, it supports the above conclusions. The policy has no obvious effect on the treatment of pollutants such as urinary particulate matter (PM2.5) and respirable particulate matter (PM10) and integrated air quality (AQI).

CONCLUSIONS AND RECOMMENDATIONS

In recent years, with the continuous increase of the number of motor vehicles in China, automobile exhaust has become the main source of urban air pollution. In response to this problem, tightening vehicle emission standards has become an important measure to reduce vehicle emissions. Taking the implementation of the Tianjin National V standard as an example, based on the monthly data of Tianjin air quality, the Lasso method is used to select individual control groups and regression synthesis methods for counterfactual analysis, and the evaluation of motor vehicle emission standards

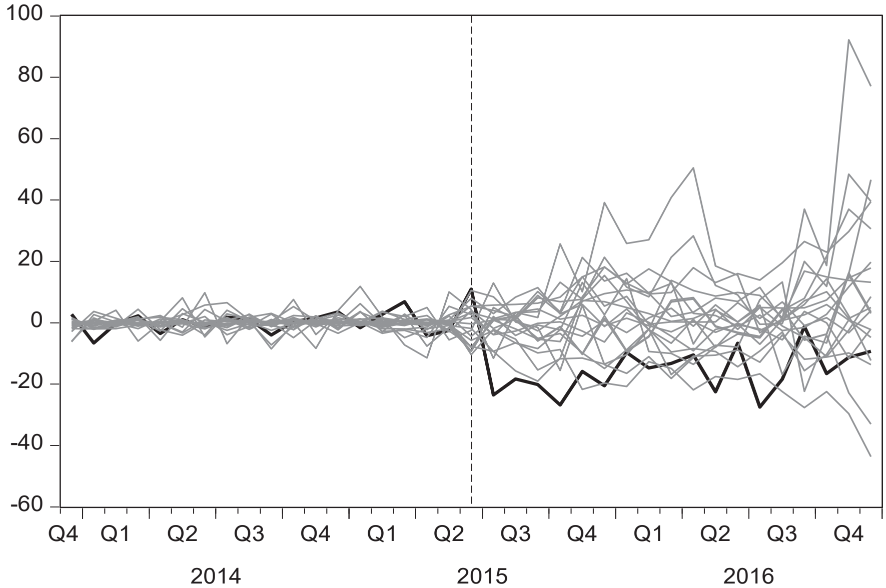

Fig. 2: Difference distribution of NO_2 in each city.

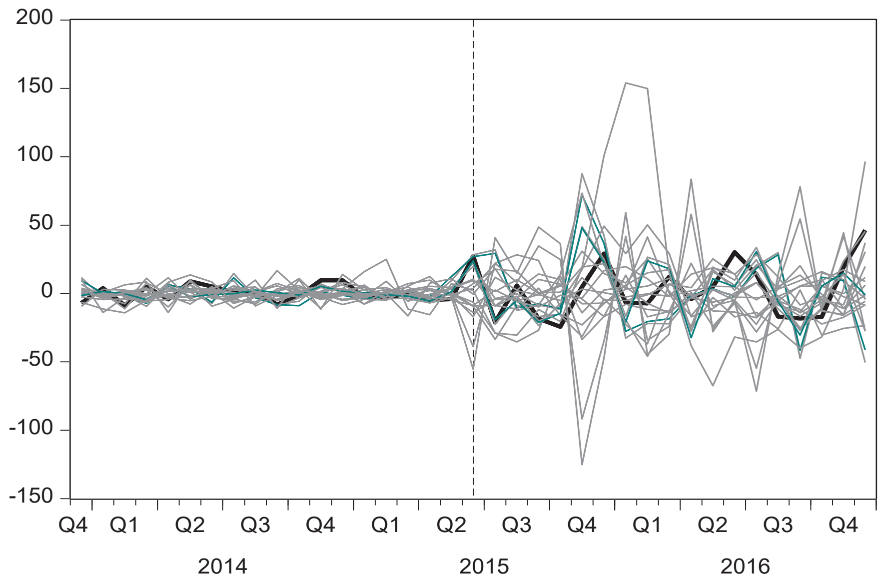

Fig. 3: PM2.5 difference distribution in each city.

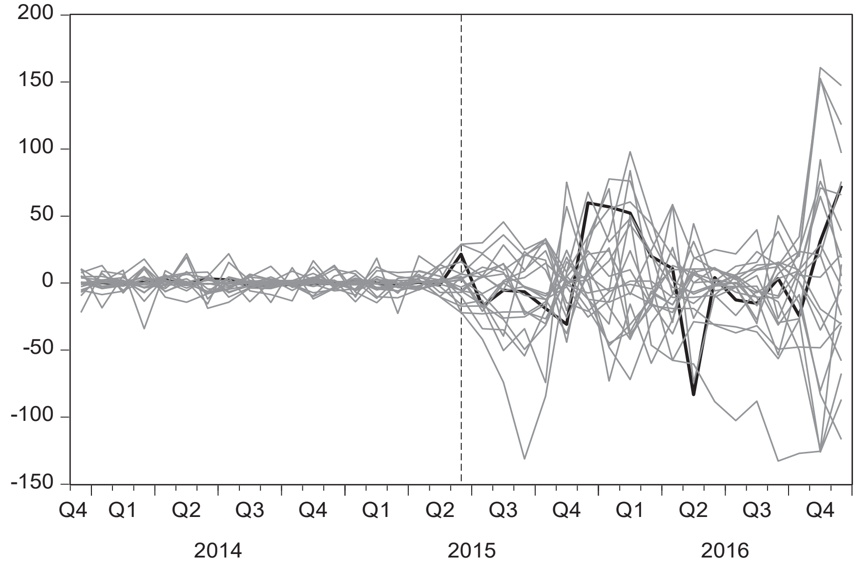

Fig. 4: PM10 difference distribution in each city.

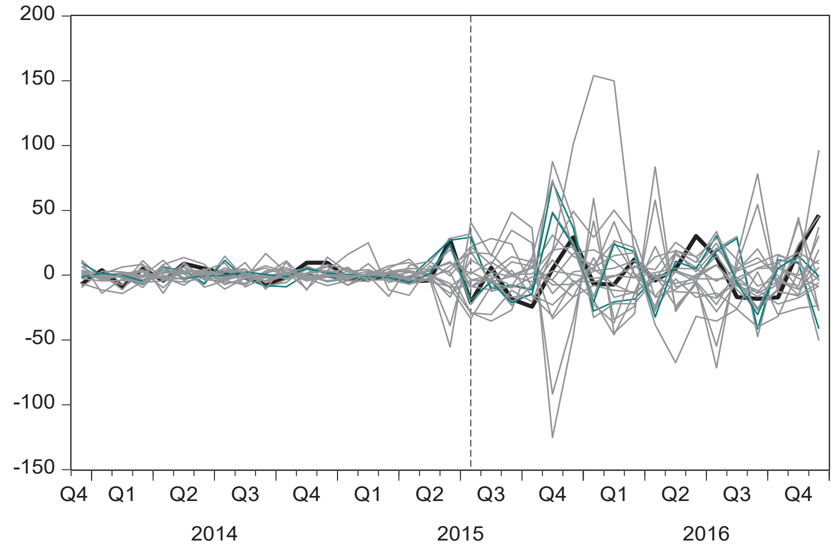

Fig. 5: Distribution of AQI difference in each city.

is evaluated for Tianjin effect of air treatment. The study found that improving vehicle emission standards is an indispensable condition for improving air quality. The long-term implementation of National V standards helps to reduce the concentrations of carbon monoxide (CO) and nitrogen dioxide (NO_2) in the atmosphere. The effects of pollutants such as particulate matter (PM2.5) and respirable particulate matter (PM10) are not obvious, and the effect on overall air treatment is not significant.

Therefore, while further regulating motor vehicle pollutant emission standards, other management and control policies should be implemented to control air quality. First, further improve the urban comprehensive transportation system and explore the development model of green public transportation; second, develop new energy vehicles and alternative fuels, reduce the dependence of road traffic on fossil fuels; and third, strengthen other pollution industries outside motor vehicle exhaust. The supervision of pollutant discharge reduces the pollutant emission intensity of high energy-consuming industries.

ACKNOWLEDGEMENTS

Project supported by National Natural Science Foundation of China (Grant No. 71502121)

REFERENCES

Davis, L.W. 2008. The effect of driving restrictions on air quality in Mexico City. Journal of Political Economy, pp. 116.

Jing, C., Xin, W. and Xiaohan, Z. 2014. Does the restriction policy improve the air quality in Beijing? Economics 3: 1091-1126.

Li, S. and Haefen, T.R.H.V. 2009. How do gasoline prices affect fleet fuel economy? American Economic Journal: Economic Policy, 1(2): 113-137.

Penghui, X. and Ruobing, L. 2015 Impact of oil price changes on air pollution: using motor vehicles as a transmission path. China Industrial Economy, 10: 100-114.

Song, Y., Shao, M. and Liu, Y. 2007. Source apportionment of ambient volatile organic compounds in Beijing. Environmental Science &Technology, 41(12): 4348-4353.

Wu, Y., Wang, R. and Zhou, Y. 2011. On road vehicle emission control in Beijing: past, present, and future. Environmental Science & Technology, 45(1): 147-53.

Xiaoguang, Z., Zhencheng, X. and Xuan, W. 2010. Analysis of the impact of Beijing motor vehicle restrictions on air quality. Journal of Safety and Environment, 4: 82-87.