

Article

The STIRPAT Analysis on Carbon Emission in Chinese Cities: An Asymmetric Laplace Distribution Mixture Model [†]

Shanshan Wang , Tianhao Zhao, Haitao Zheng *  and Jie Hu

School of Economics and Management, Beihang University, Beijing 100191, China; sswang@buaa.edu.cn (S.W.); zhaotianhaobuaasem@163.com (T.Z.); hijie_86@163.com (J.H.)

* Correspondence: zhenghaitao@buaa.edu.cn

[†] This paper is an extended version of our paper published in the Second Global Conference on Theory and Applications of OR/OM for Sustainability (GCTAOS 2017) with ID: 106.

Received: 2 October 2017; Accepted: 29 November 2017; Published: 4 December 2017

Abstract: In cities' policy-making, it is a hot issue to grasp the determinants of carbon dioxide emission in Chinese cities. And the common method is to use the STIRPAT model, where its coefficients represent the influence intensity of each determinants of carbon emission. However, less work discusses estimation accuracy, especially in the framework of non-normal distribution and heterogeneity among cities' emission. To improve the estimation accuracy, this paper employs a new method to estimate the STIRPAT model. The method uses a mixture of Asymmetric Laplace distributions (ALDs) to approximate the true distribution of the error term. Meantime, a designed two-layer EM algorithm is used to obtain estimators. We test the robustness via the comparison results of five different models. We find that the ALDs Mixture Model is more reliable the others. Further, a significant Kuznets curve relationship is identified in China.

Keywords: carbon emission; ALDs mixture model; EM algorithm; heterogeneity; Chinese cities

1. Introduction

A bottom-up approach is employed to reduce emission according to The Paris Climate Conference 2015 (COP21) agreement. This approach asks city governments to identify their own environmental issues, to submit their own individualised plans for reducing or limiting emissions and to provide the link with national level environmental policies. To reveal the underlying specific drivers of city-level carbon emissions, studies on the determinants of carbon emissions at city level are rapidly increasing. IPAT (Human Impact, Population, Affluence and Technology) or STIRPAT (Stochastic Impacts by Regression on Population, Affluence and Technology) model is the most popular one to address this issue. The STIRPAT model assumes a linear form, which is easy for estimation and interpretation. It links human activities in the form of driving factors to their impacts on environment and indicates intensity of each factor [1,2]. There are many statistical models and the estimation methods to resolve the IPAT or STIRPAT model, including PLS (Partial Least Square) method [3,4], panel data model [5], linear mixed model [1], LMDI (Log-Mean Divisia Index) method [6], scenario analysis [7,8] and structural decomposition analysis [9]. However, how to choose and improve estimation methods of these models when employing the STIRPAT model? This paper aims at the estimation problem of STIRPAT model. A reliable estimation is vital because it decides targets, path, finance and other key aspects of environmental policies. In this context, we discuss what factors do contribute to the carbon emission and develop models to specify the factors' influencing intensities.

Here we take data of 11-year period from 76 Chinese cities as an empirical example to investigate the above-mentioned concerns. China has been the largest carbon emitter in the world since 2007 [10].

In 2014, China accounted for almost 27% of overall global emissions (IEA world energy statistics 2016). China's carbon mitigation is very important for arriving at the aim of world's carbon reduction. China promises to achieve a peak in carbon emissions and decrease its carbon intensity by 30% compared to 2014 before 2030. In China, cities are the cells of carbon mitigation policy making and execution. The Environmental Protection Law of the People's Republic of China argues that cities should consider environmental protection in their social developments plans. It is noted that there are significant disparities in carbon emission per capita for 76 Chinese cities which may not be subject to normal distribution. Thus, the paper will analyze the data characteristics of 76 Chinese cities' emission and build the best statistical estimation method to analyze STIRPAT model. Many works do efforts to find an optimal solution by obtaining an accurate coefficient estimate.

As mentioned before, policy making usually requires a conditional mean estimate with accuracy. But heterogeneity and non-normality brings challenges to common estimation methods. Fortunately, composite quantile regression (CQR), which is proposed by Zou and Yuan [11], is a useful approach to solve the problem. The CQR estimation not only requires no specific distribution assumption for the error term, but also provides a more efficient estimator. Then in 2017, an improved CQR estimation named as ALDs mixture model is introduced [12]. It assumes that the error term follows a finite mixture of asymmetric Laplace distributions (ALDs), and then develops two-layer EM (Expectation-Maximum) algorithm to obtain the estimators, which shows higher performance. Thus, we will adopt the ALDs mixture model and related estimation procedure to obtain the estimators of interested in this paper.

In model specification, a squared logarithm GDP per capita is added to identify the existence of an Environmental Kuznets Curve (EKC) relationship at city level in China. EKC curve is firstly proposed by Grossman and Krueger in 1991 [13] when they analyzed the potential environmental benefits of the North American Free Trade Agreement (NAFTA). According to them, economic growth, rather than a threat, will lead to environmental improvement, and the turning point appears in the high stage of income level. Such a non-linear conclusion which is usually in form of a squared logarithm GDP per capita has been largely accepted by early economists [14,15]. Models with squared terms directly show the 'turning point' of income, where emissions or concentrations are at peak [16]. Therefore, it is of vital importance that the coefficient of the quadratic term is estimated accurately. This paper works to obtain more accurate estimators of coefficients. Therefore, it is able to accurately learn whether there exists an EKC relationship at city level in China.

In Section 2, we illustrate models we use in this paper. In the section, the STIRPAT model specification and the ALDs mixture model are introduced. Then data is described with basic statistics in Section 3. Section 4 gives out the empirical result. And to prove the reliability of the ALDs mixture model, we do model comparison in Section 5. At the end, the conclusion and a short discussion about policy making are given.

2. Model

2.1. An STIRPAT Model with EKC Setting

STIRPAT is one of the most popular model in ecology research. It is proposed by Dietz and Rosa [17] and works as a mathematical generalization of traditional IPAT model. The model uses driving factors to measure the human activity's impact (I) on environment. In the field of carbon emission, scholars use it to estimate factor intensities. A basic specification has three ones: population (P), affluence (A) and technology (T). It can be written in a non-logarithm form: $I_i = aP_i^b A_i^c T_i^d e_i$. This makes good on theoretical analysis, but researchers usually use its logarithm form to do estimate.

$$\ln I = a + b(\ln P) + c(\ln A) + d(\ln T) + e \quad (1)$$

Selectin of indicator variables for I , P , A and T is important. In I 's determinant, both total amount of carbon emission and carbon emission per capita are used. It is intuitive to use the former, the total amount of carbon emission, out of its convenience. Wang et al. [18] use total carbon emission to indicate

the impact of I at province level. At city level, the total amount is also used by Zhu and Peng's work [19]. However, in the work by Hao et al. [20], this variable is criticized out of the influence of population scale, and it is substituted by carbon emission per capita. In a recent work by Zheng et al. [1], carbon emission per capita is also used. Following these discussions, we select carbon emission per capita to indicate the I .

In the field, the population factor is usually decomposed into scale and structure. Annual average population and urbanization ratio are commonly used separately for the both [20,21]. Meanwhile, other variables are also tested but they are special for specific issues. For instance, Zhu and Peng [19] add proportion of working age population and household size to the population factor to do further research. In this paper, we follow the common setting and use the annual average population and urbanization ratio. In variable selection of the technology factor, industrialization and energy structure are used by scholars. Yet, they are used separately in different works. Su et al. [22] apply industrialization rate and energy structure when Dong et al. [23] only use the former. Similarly, Li et al. [24] use only the industrialization rate in their study. Referring to Zheng et al. [1], we have the both in our STIRPAT model. As to the affluence factor, GDP per capita is commonly applied [22,24,25]. Then, we add a quadratic term of logarithm GDP per capita to identify a Chinese Kuznets curve at city level. A significant negative estimator of the GDP per capita shows an invert "U" shape curve.

2.2. An ALDs Mixture Model

Here we briefly describe the ALDs mixture model proposed in Wang and Xiang [12], then apply it to perform the STIRPAT analysis on carbon emission via 76 cities in China during 2002–2012. An asymmetric Laplace distribution (ALD) proposed by Koenker and Machado [26] is a generalization of Laplace distribution (LD). Specifically, denote $ALD(\mu, \sigma, \tau)$ as the ALD with the location parameter $\mu \in R$, scale parameter $\sigma > 0$ and asymmetric (skewness) parameter $0 < \tau < 1$, and its probability density function is given by

$$f(\omega; \mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left(-\rho_{\tau}\left(\frac{\omega-\mu}{\sigma}\right)\right), \quad \omega \in R \quad (2)$$

where $\rho_{\tau}(t) = t(\tau - I(t < 0))$ is the check function with $I(\cdot)$ denoting the usual indicator function. ALD is common in quantile regressions, and this three-parameter distribution is useful for estimating quantile and quantile regression. See Yu and Zhang [27] for a summary of its properties and extensions. Using the unifying bridge existing between the CQR estimation and the ALD, Wang and Xiang [12] proposed the ALDs mixture model, which could be formulated as:

$$Y = X^T \beta + \varepsilon, \quad \varepsilon \sim \sum_{k=1}^K \pi_k ALD(b_k, \frac{\sigma}{\psi_k}, \tau_k), \quad (3)$$

where Y is the response, X is an p -vector of independent variables, $\beta = (\beta_1, \dots, \beta_p)^T$ is a p -vector of unknown parameters, and ε is the error term following K -component mixture of the ALDs, with the notations being $\tau_k = \frac{k}{K+1}$, and $\psi_k^2 = \kappa_k^2 + \nu_k^2$ with $\kappa_k = \frac{1-2\tau_k}{\tau_k(1-\tau_k)}$ and $\nu_k^2 = \frac{2}{\tau_k(1-\tau_k)}$. Here, K is the number of ALD components, π_k s are weights of each ALD distribution and summarized to 1, and τ_k is the skewness defined as $\tau_k = \frac{k}{K+1}$. That is, τ_k s are actually equally spaced quantiles as recommended in Zou and Yuan [11]. The ALDs mixture model asks for centralized or normalized variables, which means there is no intercept in model (3).

Specifically, in model (3), let the response Y be *ln carp*, and the independent variables $X = (\text{lngdpp}, \text{lnapop}, \text{lnurb}, \text{lnind}, \text{lnenc}, (\text{lngdpp})^2)^T$ (Refer to the definitions of each variable in Table 1), then the model proposed in this paper is formulated as

$$\text{ln carp} = \beta_1 \text{lngdpp} + \beta_2 \text{lnapop} + \beta_3 \text{lnurb} + \beta_4 \text{lnind} + \beta_5 \text{lnenc} + \beta_6 (\text{lngdpp})^2 + \varepsilon, \quad (4)$$

where the error term $\varepsilon \sim \sum_{k=1}^K \pi_k ALD(b_k, \frac{\sigma}{\psi_k}, \tau_k)$. Here we adopt the two-layer EM algorithm proposed in Wang and Xiang [12] to obtain the estimators for these unknown parameters. For more details of the proposed ALDs mixture model and its implementation algorithm, please refer to the work of Wang and Xiang [12].

3. Data

Data in this work is a balanced panel setting, which covers 76 cities in China and has an 11-year period. All are acquired from yearbooks of the 76 cities. Further, the amount of data, 836 observations, provides enough freedom degree for at most 20 parameters. Definitions of each variable could be found in Table 1.

Table 1. Variable Definitions.

Variables	Definition	Measurement
<i>ln carp</i>	Logarithm of per capita carbon emissions	Tons per capita
<i>ln gdpp</i>	Logarithm of GDP divided by population	Yuan per capita (2002 Beijing price)
<i>ln apop</i>	Logarithm of average annual resident population	Ten thousand people
<i>ln urb</i>	Logarithm of percent of the urban population in resident population	-
<i>ln enc</i>	Logarithm of share of coal consumption in energy consumption	-
<i>ln ind</i>	Logarithm of share of valued added in secondary industry	-
$(\ln gdpp)^2$	Quadratic term of <i>ln gdpp</i>	Yuan per capita squared

In the model, *ln carp* works as the dependent variable. It indicates the carbon emission per capita accounted by special procedures [1]. we calculate the carbon emission of each city by adding the carbon emission embodied in energy consumption, major industrial productions, urban garbage disposal and subtracting the carbon emission sinked by grass land [1,28]. For more details, descriptive statistics are tabulated in Table 2. Here *ln gdpp* bases on the constant price of Beijing in 2002. And *ln apop* uses average annual resident population which follows the new official statistical caliber after 2008. The new one considers more about population mobility. In the setting of *ln urb*, the ratio of urban population is selected as an descriptor. The other two variables, *ln enc* and *ln ind* are also ratios. The former is accounted with a conversion [1] and it reflects the use of coal in energy consumption. Considering the role of second industry in carbon emission, *ln ind* is designed as the logarithm percentage of secondary industry.

When doing estimate by the ALDs Mixture Model, each variable is demeaned. However, in quantile regressions, this limit does not exist. The ALDs method is a conditional mean estimation. Then such a setting does not matter in the condition that we focus most on the estimator of β . Moreover, a density histogram of *ln carp* is demonstrated in Figure 1, from which we can see that it is non-normal distributed, and we will verify it further via normality tests in Section .

Table 2. Descriptive Statistics.

Items	<i>ln gdpp</i>	<i>ln apop</i>	<i>ln urb</i>	<i>ln enc</i>	<i>ln ind</i>	$(\ln gdpp)^2$
Minimum	7.990	3.900	-3.270	-9.130	-1.590	4.319
25% percentile	9.220	5.800	-1.270	-0.590	-0.810	4.935
Median	9.680	6.200	-0.820	-0.340	-0.670	5.153
Mean	9.691	6.175	-0.891	-0.425	-0.706	5.153
Standard deviation	0.653	0.653	0.653	0.653	0.653	0.653
75% percentile	10.180	6.590	-0.530	-0.160	-0.570	5.384
Maximum	11.240	7.980	0.000	1.860	-0.030	5.854

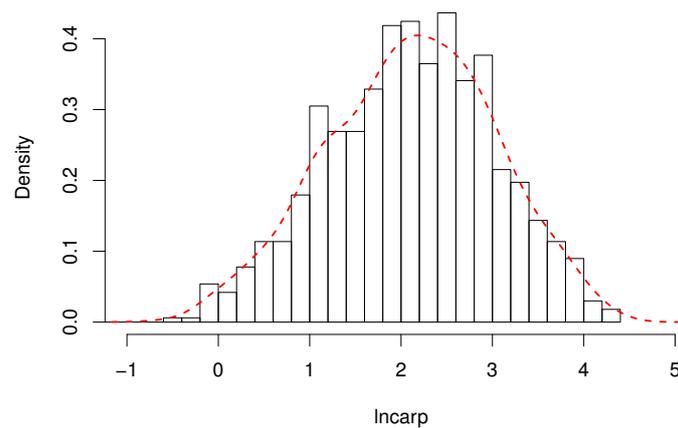


Figure 1. Distribution of *Incarp*.

4. Empirical Results

4.1. Normality Test

Here we illustrate the motivation of the proposed method. As we know, normality assumption is important for the ordinary least square (OLS) method. If it violated, the results obtained by OLS, i.e., the t test and F tests, will be unreliable. First, we show the distribution of *Incarp* via normal Q-Q (quantile-quantile) plot in Figure 2, where the black circles mark relative positions of sample quantiles to theoretical ones and points on the red line indicate an equal of the both. There exists obvious deviation of the normal Q-Q plot from a straight line, which shows abnormal behaviors at tails.

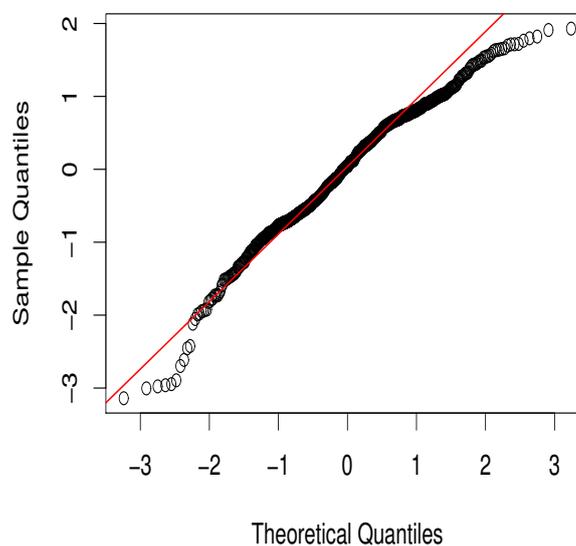


Figure 2. Normal Q-Q Plot of *Incarp*.

Moreover, we perform 5 popular normality tests described in [29]: Anderson-Darling test, Cramer-Von Mises test, Lilliefors (Kolmogorov-Smirnov) test, Pearson Chi-Square test and Shapiro-Francia test (see in Table 3). All tests have a p -value fast smaller than 0.05, which indicates an abnormal distribution of data. Thus, the results based on the normality assumption, including OLS and panel methods, will be unreliable. This motivates us to consider new estimation method.

Table 3. Normality tests on *Incarp*.

Test	Statistics	p-Value
Anderson-Darling Test	4.851514	0.0000
Cramer-Von Mises Test	0.621861	0.0000
Lilliefors (Kolmogorov-Smirnov) Test	0.049977	0.0000
Pearson Chi-Square Test	57.32536	0.0001
Shapiro-Francia Test	0.931959	0.0000

4.2. Result of the ALD Method

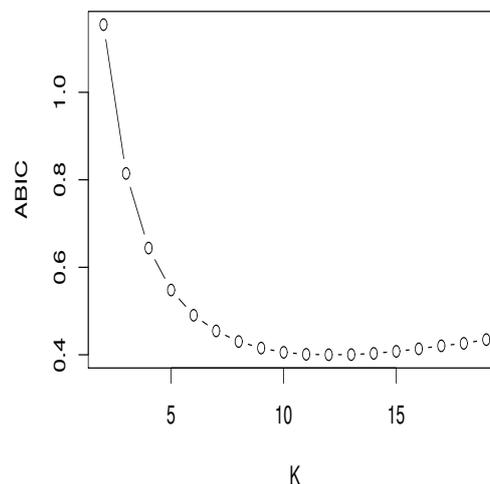
The number of components, K , is important for the ALDs mixture regression model and an optimal K will make a trade-off between the goodness of fit and model complexity. Here we adopt the heuristic information criterion called ABIC proposed in [12] to determine the number of components K . Specifically, the ABIC is defined as

$$ABIC(K) = -\frac{2l_{obs}(\hat{\theta}_K, \hat{b}_K; \mathcal{O}) + (2K + p) \ln n}{n}, \quad (5)$$

where $l_{obs}(\theta_K, b_K; \mathcal{O}) = \sum_{i=1}^n \log \left(\sum_{k=1}^K \frac{\pi_k \psi_k \tau_k (1 - \tau_k)}{\sigma} \right) \exp \left(-\rho_{\tau_k} \left(\frac{\psi_k (Y - X^T \beta - b_k)}{\sigma} \right) \right)$ is the log-likelihood function evaluated at the estimators $\hat{\theta}_K$ and \hat{b}_K ; $\hat{\theta}_K = (\hat{\pi}^T, \hat{\beta}^T, \hat{\sigma})^T$ and $\hat{b}_K = (\hat{b}_1, \dots, \hat{b}_K)^T$; $2K + p$ is the number of independent parameters; and $\mathcal{O} = \{X_i, Y_i\}_{i=1}^n$ is the observed data. An optimal K comes when the information value in (5) goes to minimum, i.e.,

$$\hat{K}_{opt} = \arg \min_K ABIC(K).$$

For the model (4) and the data described in Section 4.2, Figure 3 shows the ABIC values against different K , where each value is marked by circle and a trend by K is shown with line. It indicates that the optimal K is 12.

**Figure 3.** ABIC values against different K .

When K is selected, estimators can be acquired by the two-layer EM algorithm [12]. We use R software to do this work, then perform statistical inferences through bootstrap. As mentioned before, we have twenty parameters to be estimated, where the dataset has over 830 observations. The algorithm converges before 10,000th iteration comes. The empirical results are listed in Table 4, where estimates and their standard errors (in parentheses) are provided. The same presentation can be found in the other tables.

Table 4. Estimates of the ALDs mixture model.

Variable	β
<i>lngdpp</i>	0.9644 *** (1.2023)
<i>lnapop</i>	−0.2082 *** (0.0933)
<i>lnurb</i>	0.3314 *** (0.1493)
<i>lnenc</i>	0.2619 *** (0.1727)
<i>lnind</i>	2.3551 *** (0.3692)
$(lngdpp)^2$	−0.0325 *** (0.0588)
R^2	0.3683
$\hat{\sigma}$	0.7667

Notes: 1. $p < 0.01$ ‘***’, $p < 0.05$ ‘**’, $p < 0.1$ ‘*’; 2. Standard errors are provided in parentheses.

From Table 4, all estimators are significant. R squared shows that the STIRPAT model can explain 36.8% city carbon emission. The three percentiles, *lnurb*, *lnenc* and *lnind*, have positive coefficients, which means positive influence. However, the intensity of *lnind* is much greater than that of the other two. It means that each percent growth of production industry can increase more carbon emission than the other two factors. *lngdpp* shows a positive coefficient and its quadratic term has a negative one, proving an invert “U” shape Kuznets curve in Chinese cities. The quadratic form leads to a peak of carbon emission.

We also report the estimates for the mixing probability π and the corresponding location parameter *b* for each ALD component in Table 5. Further, we have the residual histogram in Figure 4. It indicates a non-normal distribution with two peaks. The red curve is the distribution of estimated ALDs mixture.

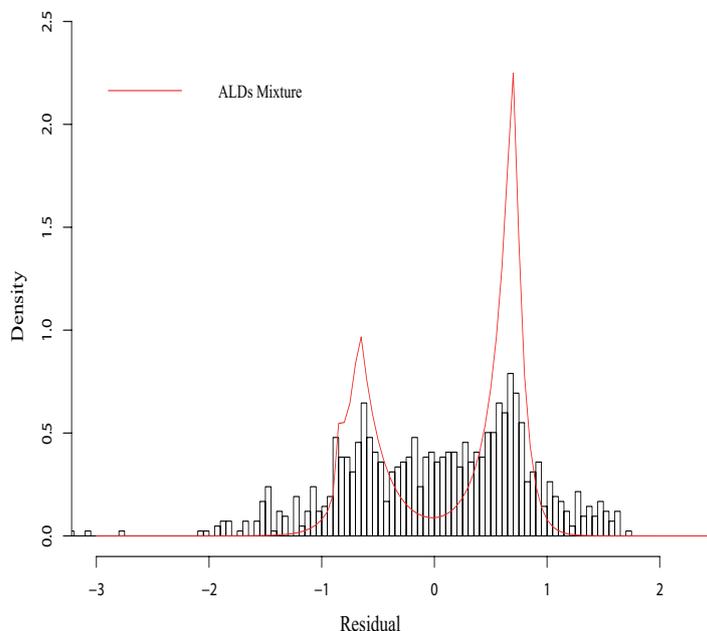


Figure 4. Residual distribution and ALDs mixture distribution.

Table 5. Estimators of ALD components.

τ	π	b
0.0769	1.94×10^{-8}	-1.0811
0.1538	4.15×10^{-2}	-0.852
0.2308	7.36×10^{-8}	-0.7009
0.3077	9.53×10^{-8}	-0.6997
0.3846	3.38×10^{-1}	-0.6654
0.4615	3.04×10^{-8}	-0.2435
0.5385	1.32×10^{-4}	-0.0939
0.6154	3.92×10^{-8}	0.6791
0.6923	4.96×10^{-1}	0.7286
0.7692	1.16×10^{-4}	0.6901
0.8462	1.24×10^{-1}	0.6954
0.9231	1.16×10^{-8}	1.0314

5. Model Comparison

To see whether the ALDs mixture model is more reliable than others, we compare it with 5 different estimations: OLS, LAD, Panel model with fixed individual effects, QR and CQR. Estimation accuracy of each is criticized or discussed through violations of error assumption and other reasons. Then we discuss improved properties of the ALDs mixture model. We depart the comparison into 2 stages.

5.1. ALDs Mixture Model vs. OLS, LAD and Panel Model with Fixed Individual Effects

On the first stage, the ALDs mixture model is compared with OLS, LAD and panel model with fixed individual effects which are common in literatures mentioned. We still use R software to estimate and list detailed estimators with inferences in Table 6. Standard errors (S.E.) are also provided beneath each coefficient.

Table 6. Comparison with OLS, LAD and panel model.

Variable	OLS Coefficients	Fixed Effect Coefficients	LAD Coefficients	ALD $K = 12$ Coefficients
<i>lngdpp</i>	-2.179 ** (1.1410)	1.622 *** (0.3613)	-1.234 (1.2551)	0.964 *** (1.2023)
<i>lnapop</i>	-0.058 (0.0545)	-0.046 (0.1172)	-0.196 *** (0.0595)	-0.208 *** (0.0933)
<i>lnurb</i>	0.480 *** (0.0799)	0.141 *** (0.0361)	0.655 *** (0.1180)	0.331 *** (0.1493)
<i>lnenc</i>	0.313 *** (0.0560)	0.081 ** (0.0339)	-0.082 (0.1155)	0.261 *** (0.1727)
<i>lnind</i>	2.889 *** (0.1716)	0.656 *** (0.0957)	1.946 *** (0.3354)	2.355 *** (0.3692)
$(\text{lngdpp})^2$	0.121 ** (0.0587)	-0.053 *** (0.0188)	0.073 (0.0634)	-0.032 *** (0.0588)

Notes: 1. $p < 0.01$ '***', $p < 0.05$ '**', $p < 0.1$ '*'; 2. Standard errors are provided in parentheses.

We discuss each estimates in Table 6. The OLS, as a benchmark, ranks first. There has been a short discussion about this estimation in empirical result section. We test the normality assumption there and prove an inefficiency of t tests and F test. Such an inefficiency leads to a poor reliability of the estimation result. Because we cannot make a right confidence in coefficients, there is no guarantee of estimation accuracy. If we step into the second method, the panel model with fixed individual effects, the same problem exists. They use effect terms, varying coefficients etc. to deal with heterogeneity. However, no matter what type a panel model has, all these ones share a common normality assumption of the final error without effects. Thus, a violation similar to what in OLS will lead to a similar inefficiency of inferences in spite of the exact type of a panel model.

To obtain a comparability with previous works, we use a panel model with fixed individual effects like what in [1] to do normality tests. Table 7 demonstrates the results of the same five tests used in

the empirical result section. We can see each of them rejects the null hypothesis that there is a normal distribution. Then in the same way as in the discussion about OLS, panel models are unreliable here.

Table 7. Normality tests on panel model.

Test	Statistics	<i>p</i> -Value
Anderson-Darling Test	7.1695	0.0000
Cramer-Von Mises Test	1.2134	0.0000
Lilliefors (Kolmogorov-Smirnov) Test	0.0669	0.0000
Pearson Chi-Square Test	73.4737	0.0000
Shapiro-Francia Test	0.9492	0.0000

The least absolute deviation (LAD) is known for its robustness especially in thick tail situations even with infinite variance. It is also listed here to see its effectiveness. Distinct from the previous two, LAD does not require a normality assumption but only one that the median is able to represent the whole error distribution. A visual inspection of whether such a representativeness met can be obtained from Figure 4. The figure clearly indicates a “valley” around the median at -0.098 . When considering the two peaks mentioned before, the reliability of LAD deserves a serious doubt. Besides, another criticism on LAD says that it is not a conditional mean estimate but describes the median. And what we do in this work is to find an improvement of the former. But the LAD is still able to work as a benchmark.

From discussion above, we can see the results of OLS, LAD, and panel models are unreliable. But things goes different in the ALDs mixture model. The ALDs mixture model does not require a normal error distribution but a mixture of ALDs. Besides, its result is a conditional mean estimate but not a median one just like that in LAD. As well not similar to ALD, such a mixture allows higher plasticity, where plural components are estimated and weighted to approximate the real error. Then the model is able to match various types of distributions even those without a regular form. When computed with the two-layer EM algorithm, the ALDs mixture model has efficient estimators and efficient inferences. Relative proof and Monte Carlo trails can be found in [12]. Therefore, when compared with the other four methods, the ALDs method is more reliable here and has a better estimation accuracy.

5.2. ALDs Mixture Model vs. QR and CQR

On the second stage, we compare the ALDs mixture model with quantile regressions (QR) and composite quantile regression (CQR). Similar to panel models, quantile regressions are widely used to inspect local behaviours at different quantile levels. A fact is that quantile regressions are indeed not conditional mean estimate but useful benchmarks. Because there is a natural logic that if a behaviour appears at each local position then it could be observed at the average level. The logic is applied here to prove the reliability of the ALDs mixture model.

We do quantile regressions at an often-used group of quantile levels from 0.1 to 0.9 by 0.2. Table 8 demonstrates detailed estimation results with inferences. On the results of $\ln g d p p$, we can see all coefficients positive and significant while each that of $(\ln g d p p)^2$ is negative and shows significance as well. A more intuitive inspection can be obtained in Figure 5. Combined with the first stage discussion, the results of QR prove the reliability of the ALDs mixture model in a different way.

The last estimation to compare with the ALDs mixture model is the CQR method. As mentioned before, the former is a generalization of the latter. We use *cqrReg* package in R to do CQR estimate and the coefficients with inferences by bootstrap are shown in Table 9. In the computation of the likelihood optimization, the alternating direction method of multipliers (ADMM) is used. In the table, we can see a highly similar result compared with that of the ALDs method in Table 4. Although we do not know the true value of the intensities, we are able to say that the ALDs mixture model works at least as good as the CQR method.

Table 8. Quantile Regression Result.

Variables	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
<i>lngdpp</i>	1.1615 *** (0.2041)	0.8314 *** (0.1508)	0.9248 *** (0.1533)	0.9336 *** (0.0723)	1.0673 *** (0.1147)
<i>lnapop</i>	-0.1848* (0.0963)	-0.0772 (0.1017)	-0.2801 *** (0.0780)	-0.3285 *** (0.0328)	-0.3063 *** (0.1153)
<i>lnurb</i>	0.3521* (0.1850)	0.5344 *** (0.1203)	0.5800 *** (0.1351)	0.4168 *** (0.0977)	0.3338 ** (0.1419)
<i>lnenc</i>	0.8111 *** (0.1123)	0.2063 (0.2340)	-0.1872 *** (0.1020)	-0.0342 *** (0.0533)	0.2154 ** (0.0905)
<i>lnind</i>	3.5585 *** (0.3893)	2.4306 *** (0.3322)	1.8876* (0.3349)	2.0771 (0.1498)	1.4560 *** (0.4942)
$(\text{lngdpp})^2$	-0.0622 *** (0.0170)	-0.0385 *** (0.0114)	-0.0357 *** (0.0127)	-0.0287 *** (0.0061)	-0.0418 *** (0.0098)

Notes: 1. $p < 0.01$ '***', $p < 0.05$ '**', $p < 0.1$ '*'; 2. Standard errors are provided in parentheses.

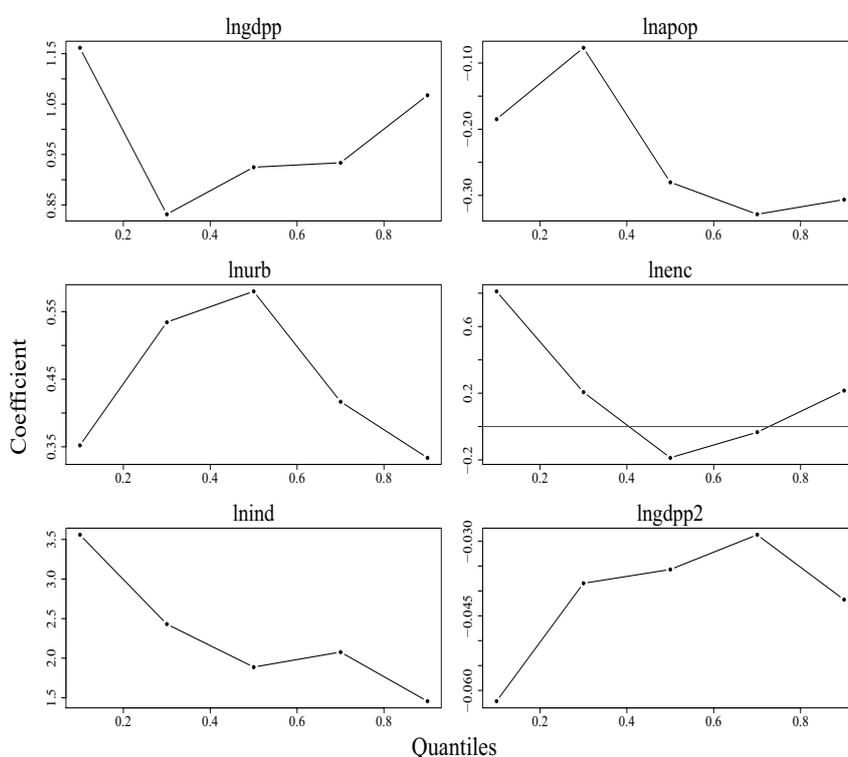


Figure 5. Estimators varying with quantile level in quantile regressions.

Table 9. Estimation results of Composite Quantile Regression.

Variable	β
<i>lngdpp</i>	1.251 *** (0.6808)
<i>lnapop</i>	-0.2054 *** (0.0688)
<i>lnurb</i>	0.4375 *** (0.1121)
<i>lnenc</i>	0.0691 *** (0.0976)
<i>lnind</i>	2.1622 *** (0.2969)
$(\text{lngdpp})^2$	-0.0531 *** (0.0356)

Notes: 1. $p < 0.01$ '***', $p < 0.05$ '**', $p < 0.1$ '*'; 2. Standard errors are provided in parentheses.

6. Discussion

The ALDs mixture model indicates existence of a Chinese Kuznets curve in cities. It helps to learn a peak of carbon emission in China. With our Chinese cities' dataset from 2002 to 2012, we launch at a 2012 Chinese GDP per capita level at 31,171 yuan (2002 Beijing constant price) [1] and the ALD method's result shows a symmetry axis at $\ln 14.837$. It means an 8.7% GDP per capita growth if the country wants to reach the peak in 2030.

Meanwhile, we compare the results of other research to those of this paper. In 2017, Mi et al. [30] use IMEC (Integrated Model of Economy and Climate) model to do forecasting. They say China will reach an emission peak at 11.20 Gt ($1Gt = 1 \times 10^9$ tons) in 2026. In another work by Li et al. [24], three scenarios are applied. These scenarios based on STIRPAT model lead to a similar result that China is able to achieve the goal in a range from 2024 to 2030. Similar to this, Ren and Xia [31] as well use scenario method to insight into this topic. They employ Markov chains distinct from STIRPAT model. However, their work tells us that China will stand on a razor's edge in 2030. The country needs to do more to avoid a default. When compared, our result performs coherent.

Corresponding to the peak estimation, discussion about policy making matters. Fan et al. [32] talk about heterogeneity of carbon emission in policy-making. Huang and Meng [25] use spatial method to look into it further. In this paper, we use ALDs mixture model to obtain a more accurate estimation. The result characterizes carbon emission at an entirety level. It is helpful when environmental policies work as part of the unified structural reformation in China. Xu et al. [33] analyze this Chinese structural reformation in carbon-creating industries. They warn a policy's inefficiency at city level when there is no carbon trading market (CTM) or other designed new policies. A factor-flowing market at city level can improve the situation. Figure 6 shows that different cities stand distinctly on the distribution. Each color differentiates one city from another. This corresponds with Xu, Chen and Chen's result. Then we can say that there should be more policy autonomy left to cities in the future.

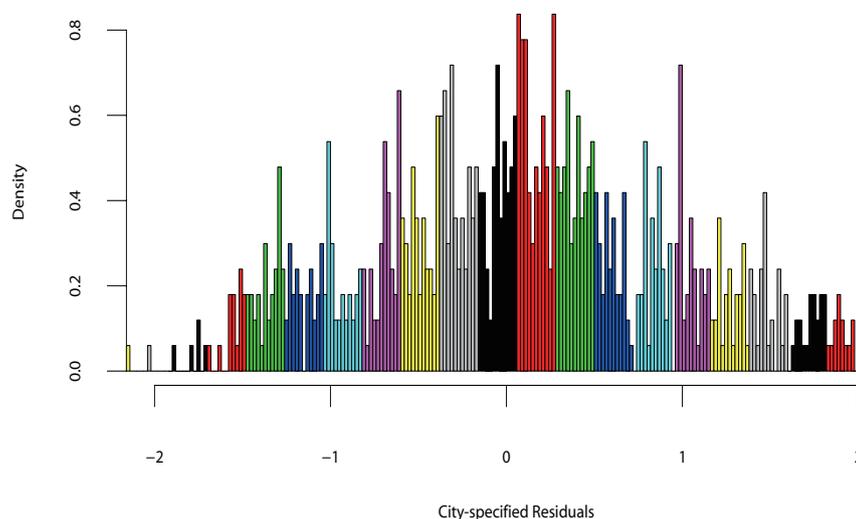


Figure 6. Residual distribution dyed by city tags.

7. Conclusions

This paper aims to obtain a more accurate estimator of a STIRPAT model. Further, a quadratic term of logarithm GDP per capita is added to identify the Kuznets curve in Chinese cities. We employ an ALDs mixture model to do estimate then solve the model with a two-layer EM algorithm. The empirical result supports the existence of EKC in Chinese cities. Further, we do model comparison to validate our result, which shows the efficiency of the ALD method. The method works as good as other estimations

and greater than some of them. At the end, a discussion about policy-making suggests more focuses on heterogeneity in cities. Because of the heterogeneity, different cities have different economy-carbon types. When making decisions, cities should have more autonomy to adapt their specific characters for a better development. On the other hand, GDP is still the main factor of carbon deduction in China. A more green growth driven by technology innovation, economic structural transformation etc. can help cities reach their peaks more rapidly.

Acknowledgments: This work was financially supported in part by the National Natural Science Foundation of China (Grant No. 71371021, 71333014, 11701023, 71420107025), Beijing Planning Program of Philosophy and Social Science (Grant No. 11JGC102) and Humanities and Social Sciences Planning Fund of Ministry of Education (Grant No. 17YJA790097).

Author Contributions: Shanshan Wang and Haitao Zheng conceived and designed the empirical study; Tianhao Zhao and Jie Hu performed the empirical study; Tianhao Zhao analyzed the data; Shanshan Wang contributed analysis tools; Tianhao Zhao, Haitao Zheng, Jie hu and Shanshan Wang wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ALD	Asymmetric Laplace Distribution
CQR	Composite Quantile Regression
EKC	Environmental Kuznets Curve
LAD	Least Absolute Derivation
OLS	Ordinary Least Square
QR	Quantile Regression
STIRPAT	Stochastic Impacts by Regression on Population, Affluence, and Technology
IPAT	Human Impact, Population, Affluence and Technology
EM	Expectation-Maximum
PLS	Partial Least Square
LMDI	Log-Mean Divisia Index

References

1. Zheng, H.; Hu, J.; Guan, R.; Wang, S. Examining Determinants of CO₂ Emissions in 73 Cities in China. *Sustainability* **2016**, *8*, 1296.
2. Shuai, S.; Yang, L.L.; Cao, J.H. Study on Influencing Factors of CO₂ Emissions from Industrial Energy Consumption: An Empirical Analysis Based on STIRPAT Model and Industrial Sectors' Dynamic Panel Data in Shanghai. *J. Financ. Econ.* **2010**, *36*, 17–28.
3. Li, B.; Liu, X.; Li, Z. Using the STIRPAT model to explore the factors driving regional CO₂ emissions: A case of Tianjin, China. *Nat. Hazards* **2015**, *76*, 1667–1685.
4. Wang, Z.; Yin, F.; Zhang, Y.; Zhang, X. An empirical research on the influencing factors of regional CO₂ emissions: Evidence from Beijing city, China. *Appl. Energy* **2012**, *100*, 277–284.
5. Xie, R.; Fang, J.; Liu, C. The effects of transportation infrastructure on urban carbon emissions. *Appl. Energy* **2017**, *196*, 199–207.
6. Shao, S.; Yang, L.; Gan, C.; Cao, J.; Geng, Y.; Guan, D. Using an extended LMDI model to explore techno-economic drivers of energy-related industrial CO₂ emission changes: A case study for Shanghai (China). *Renew. Sustain. Energy Rev.* **2016**, *55*, 516–536.
7. Wang, M.; Che, Y.; Yang, K.; Wang, M.; Xiong, L.; Huang, Y. A local-scale low-carbon plan based on the STIRPAT model and the scenario method: The case of Minhang District, Shanghai, China. *Energy Policy* **2011**, *39*, 6981–6990.
8. Guan, D.; Barker, T. Low-carbon development in the least developed region: A case study of Guangyuan, Sichuan province, southwest China. *Nat. Hazards* **2012**, *62*, 243–254.
9. Mi, Z.; Meng, J.; Guan, D.; Shan, Y.; Liu, Z.; Wang, Y.; Feng, K.; Wei, Y.M. Pattern changes in determinants of Chinese emissions. *Environ. Res. Lett.* **2017**, 074003.

10. Chambers, A.; Nakicenovic, N. *World Energy Outlook 2008*; International Energy Agency: Paris, France, 2008.
11. Zou, H.; Yuan, M. Composite quantile regression and the oracle model selection theory. *Ann. Stat.* **2008**, *36*, 1108–1126.
12. Wang, S.; Xiang, L. Two-layer EM algorithm for ALD mixture regression models: A new solution to composite quantile regression. *Comput. Stat. Data Anal.* **2017**, *115*, 136–154.
13. Grossman, G.M.; Krueger, A.B. Environmental Impacts of a North American Free Trade Agreement. *Soc. Sci. Electron. Publ.* **2000**, *8*, 223–250.
14. Meadows, D.H.; Meadows, D.L.; Randers, J.; Behrens, W.W. *The Limits to Growth*; London Earth Island Limited: London, UK, 1972; Volume 14, pp. 213–244.
15. Panayotou, T. *Empirical Tests and Policy Analysis of Environmental Degradation at Different Stages of Economic Development*; ILO Working Papers; Technology and Employment Programme: Geneva, Switzerland, 1993; Volume 4.
16. Zheng, H.; Huai, W.; Huang, L. Relationship Between Pollution and Economic Growth in China: Empirical Evidence from 111 Cities. *J. Urban Environ. Eng.* **1982**, *9*, 22–31.
17. Dietz, T.; Rosa, E.A. Effects of Population and Affluence on CO₂ Emissions. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 175.
18. Wang, P.; Wu, W.; Zhu, B.; Wei, Y. Examining the impact factors of energy-related CO₂ emissions using the STIRPAT model in Guangdong Province, China. *Appl. Energy* **2013**, *106*, 65–71.
19. Zhu, Q.; Peng, X. The impacts of population change on carbon emissions in China during 1978–2008. *Environ. Impact Assess. Rev.* **2012**, *36*, 1–8.
20. Hao, Y.; Chen, H.; Wei, Y.M.; Li, Y.M. The influence of climate change on CO₂ (carbon dioxide) emissions: An empirical estimation based on Chinese provincial panel data. *J. Clean. Prod.* **2016**, *131*, 667–677.
21. Zhang, C.; Lin, Y. Panel estimation for urbanization, energy consumption and CO₂ emissions: A regional analysis in China. *Energy Policy* **2012**, *49*, 488–498.
22. Su, Y.; Chen, X.; Li, Y.; Liao, J.; Ye, Y.; Zhang, H.; Huang, N.; Kuang, Y. China's 19-year city-level carbon emissions of energy consumptions, driving forces and regionalized mitigation guidelines. *Renew. Sustain. Energy Rev.* **2014**, *35*, 231–243.
23. Dong, L.; Fujita, T.; Zhang, H.; Dai, M.; Fujii, M.; Ohnishi, S.; Geng, Y.; Liu, Z. Promoting low-carbon city through industrial symbiosis: A case in China by applying HPIMO model. *Energy Policy* **2013**, *61*, 864–873.
24. Li, L.; Lei, Y.; He, C.; Wu, S.; Chen, J. Prediction on the Peak of the CO₂ Emissions in China Using the STIRPAT Model. *Adv. Meteorol.* **2016**, 2016.
25. Huang, B.; Meng, L. Convergence of per capita carbon dioxide emissions in urban China: A spatio-temporal perspective. *Appl. Geogr.* **2013**, *40*, 21–29.
26. Koenker, R.; Machado, J.A. Goodness of fit and related inference processes for quantile regression. *J. Am. Stat. Assoc.* **1999**, *94*, 1296–1310.
27. Yu, K.; Zhang, J. A Three-Parameter Asymmetric Laplace Distribution and Its Extension. *Commun. Stat. Theory Methods* **2005**, *34*, 1867–1879.
28. Mi, Z.; Zhang, Y.; Guan, D.; Shan, Y.; Liu, Z.; Cong, R.; Yuan, X.C.; Wei, Y.M. Consumption-based emission accounting for Chinese cities. *Appl. Energy* **2016**, 184.
29. Asghar, G.; Saleh, Z. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *Int. J. Endocrinol. Metab.* **2012**, *10*, 486.
30. Mi, Z.; Wei, Y.M.; Wang, B.; Meng, J.; Liu, Z.; Shan, Y.; Liu, J.; Guan, D. Socioeconomic impact assessment of China's CO₂ emissions peak prior to 2030. *J. Clean. Prod.* **2017**, *142*, 2227–2236.
31. Ren, F.; Xia, L. Analysis of China's Primary Energy Structure and Emissions Reduction Targets by 2030 Based on Multiobjective Programming. *Math. Probl. Eng.* **2017**, 2017, doi:10.1155/2017/1532539.
32. Fan, Y.; Liu, L.C.; Wu, G.; Wei, Y.M. Analyzing impact factors of CO₂ emissions using the STIRPAT model. *Environ. Impact Assess. Rev.* **2006**, *26*, 377–395.
33. Xu, L.; Chen, N.; Chen, Z. Will China make a difference in its carbon intensity reduction targets by 2020 and 2030? *Appl. Energy* **2017**, *203*, 874–882.

