

# Variable Selection for Panel Data Linear Regression Models with Fixed Effects

Xinye Hui\*

College of Statistics and Mathematics, Anhui University of Finance and Economics, Bengbu, China

\*Corresponding Author: Xinye Hui

## ABSTRACT

This paper introduces a robust variable selection mechanism for fixed effect panel data models by integrating compound quantile regression with the adjusted MIXED penalty method. Initially, forward orthogonal deviation transformation is employed to eliminate the influence of fixed effects. Subsequently, the MIXED penalty is utilized to construct a penalized compound quantile regression objective function, facilitating simultaneous estimation of regression coefficients and variable selection. This method not only effectively eliminates the interference of fixed effects but also demonstrates outstanding robustness. Its performance with limited sample sizes was validated through simulation studies, and its practical value was illustrated through application in real data analysis.

## KEYWORDS

Compound Quantile Regression; MIXED Penalty; Fixed Effects; Variable Selection.

## 1. INTRODUCTION

Panel data is a special type of data observed over multiple time points for the same group of individuals, thereby blending the characteristics of time series data and cross-sectional data. Panel data offers richer information, higher data frequency, and larger sample sizes. These advantages make panel data analysis critically important and widespread in the fields of economics, finance, social sciences, and medical research. Variable selection in panel data is a crucial task in statistical analysis and econometrics, especially when dealing with high-dimensional data. Effective variable selection identifies the most significant variables impacting the dependent variable, simplifying the model to enhance its explanatory power and reduce overfitting. This is particularly important in panel data analysis, where researchers often face a plethora of potential explanatory variables. Additionally, understanding which factors significantly influence economic behavior and social phenomena is vital in economics and social science research. Variable selection helps researchers focus on the most explanatory variables to deeply understand the essence of economic and social phenomena. Kock [1] applied the bridge regression method to the variable selection problem in panel data models. Subsequently, Kock and Tang[2] explored high-dimensional dynamic panel data models based on the LASSO penalty, while Li and others[3] considered the form of varying coefficient models, utilizing spline methods and SCAD penalty for variable selection in varying coefficient models. Tang Lizhi[4] and others applied the Elastic Net penalty, suitable for highly correlated variables, to Bayesian quantile regression for variable selection and parameter estimation in panel data.

Most current research on high-dimensional panel data models focuses on average regression models. However, when dealing with real-world data, one often encounters data with outliers, peakedness, and heavy tails, among other heterogeneity features. In such cases, relying solely on average

regression analysis can no longer ensure the superiority and stability of the analysis. Faced with this data heterogeneity, compound quantile regression offers a more robust solution, demonstrating stronger robustness compared to traditional average regression methods. Since the introduction of compound quantile regression, it has been widely applied across various fields. Lv Yazhao and others[5] researched variable selection in partial linear single-index models using the penalized compound quantile regression method. Later, Wang Kangning and others[6], Weng Yuling and others[7], respectively, utilized compound quantile regression techniques to explore the application of functional linear models and longitudinal data models in statistical inference problems. Additionally, Wu Dongsheng and others[8] developed an improved two-stage estimation method for panel data models based on compound quantile regression, analyzing the effects of foreign trade on economic growth. Liu Huilan[9] conducted a comprehensive study on compound quantile regression.

In variable selection for panel data, scholars commonly use LASSO-type methods, with the most widely used being the adaptive LASSO. Adaptive LASSO was proposed by Zou in 2006, employing different weights for each coefficient, allowing the model to include more significant variables. This method possesses Oracle properties in fixed dimensions and was proven by Huang in 2009 to maintain these properties even as dimensions increase with sample size. These penalties essentially belong to the  $L_1$  penalty. Theoretically, the best penalty for variable selection is the  $L_0$  penalty, which directly penalizes the number of non-zero parameters in the model. Due to the non-convex nature of the  $L_0$  penalty, optimization problems with  $L_0$  penalties become very challenging. In practice, heuristic or approximate methods, such as forward selection or  $L_1$  penalties, are commonly used to approximate the objectives of  $L_0$  penalties. Zhu and Wang proposed the MIXED penalty to address the limitations of the  $L_0$  penalty. This method simulates the  $L_0$  penalty with a simpler function that is continuous and differentiable everywhere, and to balance the stability of variable selection, they incorporated the  $L_0$  penalty into the model, demonstrating its theoretical properties and performance in data.

This paper aims to explore how to effectively eliminate the influence of fixed effects and construct a robust variable selection process by combining the MIXED penalty mechanism. The research first employs a forward orthogonal deviation transformation matrix to eliminate fixed effect terms, then integrates compound quantile regression with MIXED penalty techniques to construct a penalized compound quantile regression objective function. This process not only estimates regression coefficients but also achieves the goal of variable selection.

## 2. THEORETICAL FOUNDATIONS AND METHODS

### 2.1. Fixed Effects Model for Panel Data

The fixed effects model for panel data is a commonly used method in panel data analysis, particularly suited for analyzing observational data collected over time from the same group of individuals (such as individuals, companies, countries, etc.). The core idea of the fixed effects model is to control for individual-specific effects that do not change over time, thereby accurately estimating the impact of explanatory variables on the dependent variable. The fixed effects model controls for individual heterogeneity that does not vary over time by introducing a unique constant term (the fixed effect) for each individual. This method assumes that these unobserved individual-specific effects can be fully captured by individual-specific intercepts. By controlling for individual fixed effects, the model can reduce biases due to omitted variables, enhancing the accuracy of the estimates. The fixed effects model uses only the information on changes over time within each individual for estimation, ignoring differences between individuals. Therefore, this model is particularly suitable for analyzing the impact of explanatory variables that change over time on the dependent variable. The fixed effects model for panel data can be represented by the following formula:

$$Y_{it} = \alpha_i + X_{it}\beta + \epsilon_{it}. \quad (1)$$

Where  $Y_{it}$  is the dependent variable, representing the observation of the  $i$ th individual at time  $t$ .  $\alpha_i$  is the individual fixed effect, a unique constant term for each individual to capture all individual characteristics that do not change over time but may influence the dependent variable.  $X_{it}$  is a  $k \times 1$  vector of explanatory variables, containing other variables that may affect  $y_{it}$  aside from the individual fixed effects.  $\beta$  is a  $k \times 1$  vector of coefficients, indicating the impact of explanatory variables on the dependent variable.  $\epsilon_{it}$  is the error term, representing the random disturbance for the  $i$ th individual at time  $t$ . In practical applications, a key task of the fixed effects model is to estimate the coefficients of the explanatory variables,  $\beta$ , while controlling for the individual fixed effects,  $\alpha_i$ . By using the fixed effects model, biases potentially introduced by omitting invariant or difficult-to-observe individual-specific variables can be eliminated, allowing for a more accurate estimation of the impact of explanatory variables on the dependent variable.

## 2.2. Composite Quantile Regression Estimation

Composite quantile regression estimation is a statistical method that combines multiple quantile regressions. Unlike traditional least squares regression, which focuses only on the conditional mean of the dependent variable, quantile regression allows researchers to explore the impact of independent variables on different quantiles of the dependent variable (such as the median, upper quartile, and lower quartile). By considering multiple quantiles simultaneously, composite quantile regression offers a more comprehensive method of data analysis. Its basic form is:

$$\min_{\beta} \sum_{q \in Q} \sum_{i=1}^n \rho_q(y_i - x_i' \beta).$$

where  $y_i$  is the dependent variable for the  $i$ th observation,  $x_i$  is the vector of independent variables for the  $i$ th observation, and  $\beta$  is the coefficient vector to be estimated.  $Q$  is the set of selected quantiles, for example,  $Q = \{0.25, 0.5, 0.75\}$  represents considering the lower quartile, median, and upper quartile, respectively.  $\rho_q$  is the quantile loss function, defined for the  $q$ th quantile as  $\rho_q(u) = u(q - I(u < 0))$ , where  $I$  is the indicator function. The advantages of composite quantile regression are numerous. It is more robust to outliers because it minimizes the quantile loss function instead of the squared loss function, thereby reducing the influence of outliers. By considering multiple quantiles of the dependent variable, it provides a more comprehensive understanding of the impact of independent variables. This is particularly useful for understanding the effects of variables on different parts of the distribution of the dependent variable. This method is applicable to various types of data distributions, especially when the distribution of the dependent variable deviates from normal distribution.

## 2.3. MIXED Penalty

Suppose we construct a dataset that conforms to the Accelerated Failure Time (AFT) model as  $\log(T_i) = X_i^T \beta + \epsilon_i$ , where  $T_i = (t_1, t_2, t_3, \dots, t_n)^T$  is the response variable and also the survival time.  $X = (X_1, X_2, X_3, \dots, X_n) = (x_{ij})_{n \times p}$  is an  $n \times p$  matrix,  $\epsilon_i$  is independently and identically distributed random error with a mean of 0.  $\beta$  is a  $p$ -dimensional true parameter, we let some  $\beta_i = 0$  and some  $\beta_i \neq 0$ . As the sample size increases and some coefficients of the covariates are zero, we need to compress the components that were originally zero to zero while estimating the coefficients.

Generally, this is achieved by minimizing the loss function, while also considering penalizing the number and magnitude of the coefficients, so we have:

$$\hat{\beta}_p = \underset{\beta_p}{\operatorname{argmin}} Q_p(\beta_p), \quad Q_p(\beta_p) = \|Y - X\beta_p\|^2 + \sum_{j=1}^p p_{\lambda_n}(\beta_j).$$

If our penalty function is:

$$\sum_{j=1}^p p_{\lambda_n}(|\beta_j|) = \lambda \sum_{j=1}^p |\beta_j|$$

Therefore, for the survival time data we constructed, the loss function with a weighted LASSO regression is:

$$L(\beta) = \sum_{i=1}^n w_i (Y_{(i)} - X_{(i)}^T \beta)^2 + \lambda \sum_{j=1}^p |\beta_j|.$$

This is the loss function with weights for LASSO regression. Hui Zou proposed the Adaptive LASSO, which applies a different weight to each coefficient, allowing the penalty to adapt according to the magnitude of the coefficients. Compared to LASSO, this method prioritizes compressing those coefficients that contribute less to the model and imposes smaller penalties on larger or more important coefficients. It possesses Oracle properties, with the objective function being:

$$L(\beta) = \sum_{i=1}^n w_i (Y_{(i)} - X_{(i)}^T \beta)^2 + \lambda \sum_{j=1}^p w_j |\beta_j|.$$

where  $w_j$  is typically taken as the reciprocal of the absolute value of the least squares result.

If:

$$\sum_{j=1}^p p_{\lambda_n}(|\beta_j|) = \lambda \sum_{j=1}^p I\{\beta_j \neq 0\}.$$

This penalty mechanism is known as the  $l_0$  penalty, which theoretically possesses very valuable properties, mainly by counting the number of non-zero elements in a vector. Applying the  $l_0$  penalty in the model optimization process can effectively reduce the proportion of non-zero elements in the model parameters, thereby constructing a sparse model. However, the  $l_0$  penalty is a non-convex optimization problem, and seeking its global optimum is an NP-hard problem. Moreover, due to its discontinuity at zero, slight changes in parameters can cause drastic fluctuations in solutions,

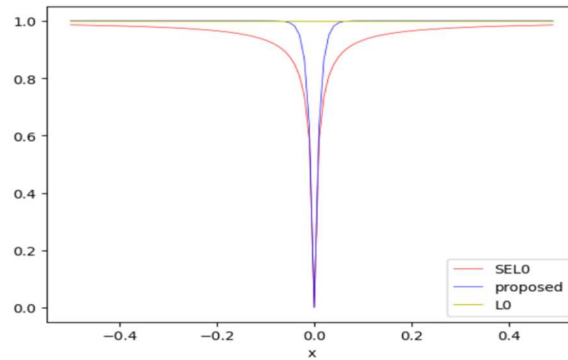
negatively affecting the stability of parameter estimation. In this context, Dicker, Huang, and Lin proposed the SELO method, with the base form of its penalty function as:

$$p_{SELO}(\beta_j) = \frac{\lambda}{(\log(2))\log\left(\frac{|\beta_j|}{(|\beta_j| + \tau)} + 1\right)}.$$

where  $\lambda$  is a tuning parameter, and when  $\tau$  is set to 0.01, the SELO penalty achieves results very close to the  $l_0$  penalty and ensures the function is smooth and differentiable everywhere. Zhu and Wang used a simpler, continuous, and everywhere differentiable function to simulate the  $l_0$  penalty, avoiding the drawbacks of the  $l_0$  penalty while achieving parameter selection as much as possible, constructing the following penalty function:

$$p_a(|\beta_j|) = 1 - e^{-\frac{|\beta_j|}{a}}.$$

Taking  $a$  as a very small positive number, usually set to 0.01. When  $a$  is small, we can deduce that  $\lim_{a \rightarrow 0^+} (1 - e^{-|x|/a}) = I\{x \neq 0\}$ . Figure 3-1 illustrates the simulation of the penalty function,  $l_0$  penalty, and SELO penalty function within the interval  $[-0.5, 0.5]$ :



**Figure 1.** Penalty Function Graphs

From Figure 1, it can be observed that the simulated penalty function is closer to the  $l_0$  penalty than the SELO penalty. Due to the continuous and differentiable characteristics of the simulated penalty function, it will be more stable in variable selection compared to the  $l_0$  penalty, and nearly unbiased as the sample size approaches infinity. To combine the convexity and sparsity of the  $l_0$  penalty, Zhu and Wang proposed the MIXED penalty:

$$\lambda_1 \sum_{j=1}^p p_a(|\beta_j|) + \lambda_2 \sum_{j=1}^p |\beta_j|.$$

where  $\lambda_1, \lambda_2$ , and  $a$  are tuning parameters, and  $p_a(|\beta_j|) = 1 - e^{-|\beta_j|/a}$ . Building on the foundation of the MIXED penalty, we consider the form of the adaptive LASSO, incorporating the adaptive idea into the penalty term, thus our penalty term is set as:

$$\lambda_1 \sum_{j=1}^p p_a(|\beta_j|) + \lambda_2 \sum_{j=1}^p \frac{|\beta_j|}{|\hat{\beta}_j|}$$

where  $\hat{\beta}_j$  can be obtained from the results of the composite quantile regression estimation without penalties.

## 2.4. Estimation of Composite Quantile Regression with Penalty Terms

Let  $Y_i = (Y_{i1}, \dots, Y_{iT})^T$ ,  $X_i = (X_{i1}, \dots, X_{iT})^T$ ,  $\epsilon_i = (\epsilon_{i1}, \dots, \epsilon_{iT})^T$ , and  $C_i = 1_T c_i$ , where  $1_T$  is a  $T$ -dimensional vector of ones. Rewriting equation (1) as:

$$Y_i = X_i \beta + C_i + \epsilon_i, \quad i = 1, 2, \dots, n. \quad (2)$$

To eliminate fixed effect terms, we adopt the concept of the forward orthogonal deviation transformation proposed by Arellano. Define the  $(T-1) \times T$  filter matrix  $D$  as:

$$D = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix}$$

By multiplying both sides of equation (2) by the matrix  $(DD^T)^{-1/2}D$ , we obtain:

$$Y_i^* = X_i^* \beta + \epsilon_i^*, \quad i = 1, 2, \dots, n. \quad (3)$$

where  $Y_i^* = (DD^T)^{-1/2}DY_i$ ,  $X_i^* = (DD^T)^{-1/2}DX_i$ , and  $\epsilon_i^* = (DD^T)^{-1/2}D\epsilon_i$ . Equation (3) eliminates fixed effects through the forward orthogonal deviation transformation. Arellano noted that a key feature of this method is the avoidance of serial correlation in the transformed error terms, i.e.,  $Var(\epsilon_i) = \sigma^2 I_T \Rightarrow Var(\epsilon_i^*) = \sigma^2 I_{(T-1)}$ . After the transformation, the sample size changes from  $nT$  to  $n(T-1)$ . For simplicity, we denote  $N = n(T-1)$ . Let  $Y_i^* = (Y_{i1}^*, \dots, Y_{i(T-1)}^*)^T$ ,  $X_i^* = (X_{i1}^*, \dots, X_{i(T-1)}^*)^T$ ,  $\epsilon_i^* = (\epsilon_{i1}^*, \dots, \epsilon_{i(T-1)}^*)^T$ . To select non-zero variables, similar to the ideas of Zou and Yuan, we construct the MIXED penalty composite quantile regression estimation  $\hat{\beta}^{MCQR}$  for equation (3) defined as:

$$\arg \min_{b_1, \dots, b_K, \beta} \left\{ \sum_{k=1}^K \left( \sum_{i=1}^n \sum_{t=1}^{T-1} \rho_{\tau_k}(Y_{it}^* - X_{it}^{*T} \beta - b_k) \right) + \lambda_1 \sum_{j=1}^p \left( 1 - e^{-|\beta_j|/a} \right) + \lambda_2 \sum_{j=1}^p \frac{|\beta_j|}{|\hat{\beta}_j^{CQ} |^2} \right\}$$

where  $\lambda_1$  and  $\lambda_2$  are penalty parameters,  $\hat{\beta}^{CQR}$  is the composite quantile regression estimation obtained without penalty,  $\lambda_N = 0$ ,  $\rho_{\tau_k}(r)$  is the quantile loss function  $\rho_{\tau_k}(r) = \tau_k r - rI(r < 0)$  for  $k = 1, 2, \dots, K$  with  $0 < \tau_1 < \tau_2 < \dots < \tau_K < 1$  and  $b_k$  is the  $\tau_k$  quantile of  $\epsilon_i^*$ . Equidistant quantiles are typically chosen,  $\tau_k = k/(K + 1)$ , for  $k = 1, \dots, K$ . In this paper,  $K = 5$  is chosen for both simulation studies and real data analysis. To determine  $\lambda_1$  and  $\lambda_2$ , we employ a grid search method to find the optimal combination.

### 3. NUMERICAL SIMULATION

To verify the finite sample properties of the method proposed in this article, this section conducts a model study with simulated data:

$$Y_{it} = X_{it}^T \beta + c_i + \epsilon_{it}, \quad i = 1, 2, \dots, n, \quad t = 1, 2, \dots, 5,$$

where  $\beta = (3, 1.5, 0, 0, 0, 0, 0, 0)^T$ . The covariates  $X$  are generated from a multivariate normal distribution with mean 0 and covariance matrix  $\text{Cov}(X_{it,l}, X_{it,s}) = 0.2^{|l-s|}$  ( $l, s = 1, 2, \dots, 9$ ).  $c_i = X_{i,1} + V_i$  for  $i = 2, \dots, n$ , where  $V_i \sim N(0, 1)$ ,  $X_{i,1} = \frac{1}{5} \sum_{t=1}^5 X_{it,1}$ , and  $c_1 = -\sum_{i=2}^n c_i$ .

The model error considers three different scenarios: normal distribution, t-distribution, and Cauchy distribution.,  $\epsilon_{it} \sim N(0, 1)$ ,  $\epsilon_{it} \sim t(3)$ , and  $\epsilon_{it} \sim 0.5C(0, 1)$ . The sample sizes are set to  $n = 200, 300$ , and  $400$ , and each scenario is run 100 times.

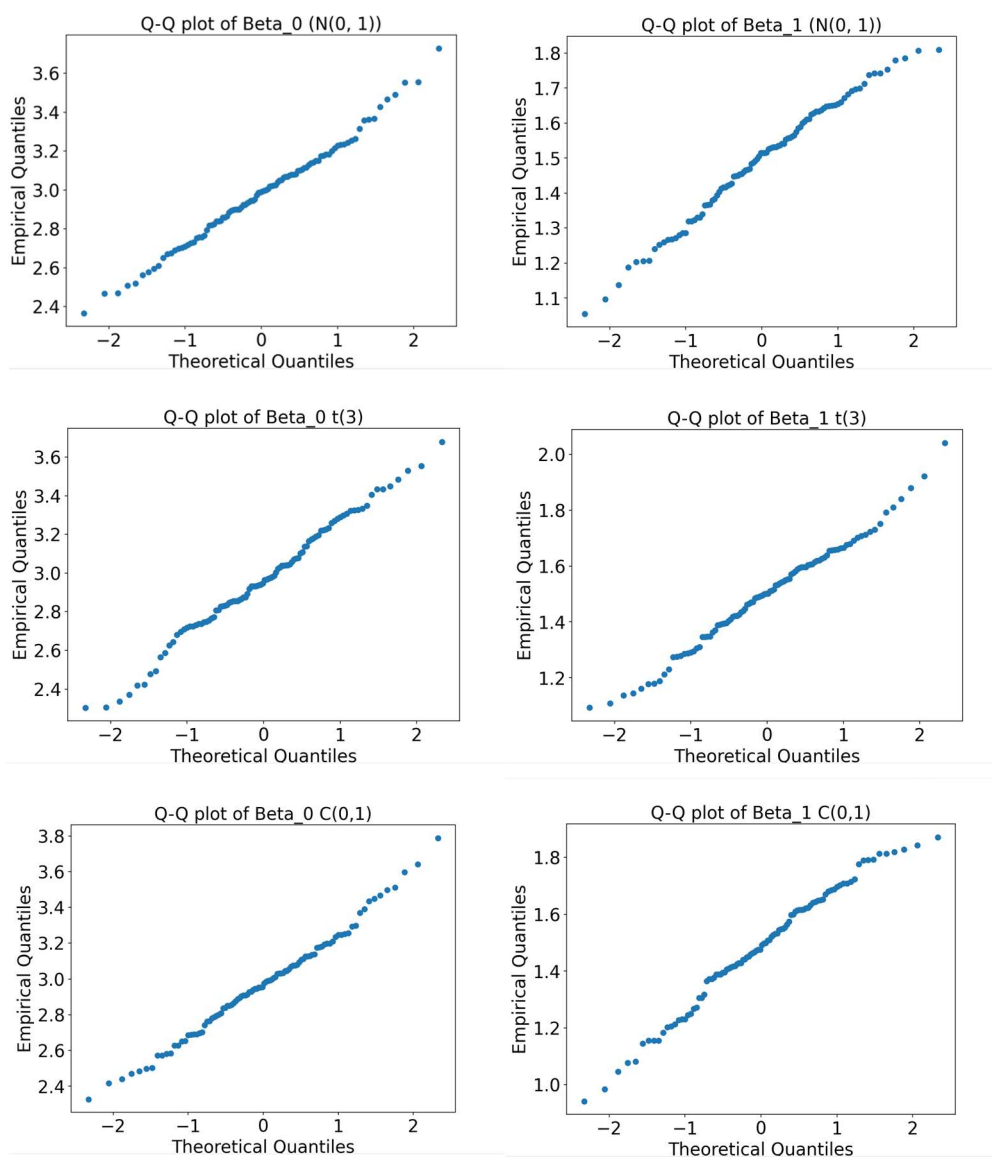
**Table 1.** Results of the Two Methods Under Three Error Distributions

$\epsilon$	$n$	LS-ALASSO			MCQR-ALASSO		
		GMSE	I	C	GMSE	I	C
$N(0,1)$	200	0.0028	0	6.88	0.0019	0	8
	300	0.0015	0	6.94	0.0009	0	8
	400	0.0009	0	6.96	0.0005	0	8
$t(3)$	200	0.0087	0	6.85	0.0035	0	8
	300	0.0038	0	6.93	0.0015	0	8
	400	0.0034	0	6.94	0.0014	0	8
$C(0,1)$	200	38.4521	0.318	6.80	0.0035	0	8
	300	2.5896	0.344	6.86	0.0014	0	8
	400	1.4567	0.350	6.92	0.0009	0	8

To compare the simulation results, we contrasted the outcomes of the newly proposed model with those obtained using the adaptive LASSO penalty based on least squares. The accuracy of variable selection and parameter estimation was measured using the Generalized Mean Square Error (GMSE) defined as:

$$GMSE(\hat{\beta}) = (\hat{\beta} - \beta)^T E(\tilde{X}\tilde{X}^T)(\hat{\beta} - \beta)$$

To assess the accuracy of variable selection, we also calculated the average number of zero variables correctly estimated as zero, denoted as "C", and the average number of non-zero variables incorrectly set to zero, denoted as "I". The simulation results are shown in Table 1. From Table 1, it can be observed that the MCQR-ALASSO technique excels in reducing model complexity and enhancing estimation accuracy, evident under the influence of different error distribution types. Compared to the LS-ALASSO method, MCQR-ALASSO performs better in terms of complexity and accuracy, especially when facing model errors that follow a Cauchy distribution, where the LS-ALASSO method's Generalized Least Squares Error (GMSE) is excessively high, and its accuracy in model variable selection is poorer, often incorrectly zeroing non-zero variables. Therefore, compared to LS-ALASSO, the MCQR-ALASSO technique proposed in this study not only significantly simplifies the model structure but also demonstrates greater robustness. For the same model error, the MCQR-ALASSO method reduces the GMSE value as the sample size increases and correctly identifies the true model.



**Figure 2.** QQ Plots of Coefficient Estimates Under Different Error Distributions at  $n=200$

As the sample size increases, the GMSE value of the CQR-ALASSO method shows a decreasing trend, and it can accurately identify the true model. Furthermore, to test whether the non-zero

coefficients estimated by MCQR-ALASSO tend towards asymptotic normality, we conducted 100 repeated experiments and drew Q-Q plots for the non-zero coefficients  $\beta_1$  and  $\beta_2$ . At a sample size of  $n=200$ , we displayed Q-Q plots under three different model error distributions (Figure 2), and similar phenomena were observed for  $n=300$  and  $n=400$ . By analyzing the Q-Q plots shown in Figure 1, we find that the data points are generally distributed around a straight line, indicating that the non-zero coefficients estimated by MCQR-ALASSO indeed exhibit the characteristics of an asymptotic normal distribution.

#### 4. CASE STUDY

In this section, we examine the "Produce" dataset available in the "plm" R package, which includes economic growth indicators for 48 states in the United States from 1970 to 1986. These indicators encompass highways and streets ("hwy"), water supply and sewage systems ("water"), personal capital stock ("pc"), state gross product ("gsp"), labor input ("emp"), and state unemployment rate ("unemp"). Munnell (1990) utilized this dataset to analyze the impact of public capital, especially highways, streets, water supply, and sewage systems on economic growth. In our modeling analysis, we consider both "water" and "hwy" as components of public capital and explore the potential impact of these indicators and their quadratic terms on economic growth. By constructing the corresponding model, we aim to deeply understand how public capital promotes state-level economic growth and to examine the contribution of other economic indicators to growth, including their nonlinear effects. This approach helps to comprehensively assess the role of public capital investment in enhancing regional economic development and provides data support for policy-making.

$$gsp_{it} = \beta_1 water_{it} + \beta_2 water_{it}^2 + \beta_3 hwy_{it} + \beta_4 hwy_{it}^2 + \beta_5 pc_{it} + \beta_6 pc_{it}^2 + \beta_7 emp_{it} + \beta_8 emp_{it} + \beta_9 unemp_{it} + \beta_{10} unemp_{it}^2 + c_i + \varepsilon_{it}, i = 1, \dots, 48, t = 1, \dots, 17.$$

**Table 2.** Simulation Results of the Two Methods

Parameter	LS-A	MCQR	Parameter	LS-A	MCQR
<i>water</i>	0	0	<i>pc</i> <sup>2</sup>	-0.0489	-0.0589
<i>water</i> <sup>2</sup>	0.0016	0	<i>emp</i>	0	0
<i>hwy</i>	1.1056	0.7056	<i>emp</i> <sup>2</sup>	0.0564	0.0726
<i>hwy</i> <sup>2</sup>	-0.0646	0	<i>unemp</i>	-0.0049	0
<i>pc</i>	1.1441	1.2347	<i>unemp</i> <sup>2</sup>	0	0

In conducting a similar simulation study, we analyzed the "Produce" dataset using both LS-ALASSO and MCQR-ALASSO methods to explore the impact of different variables on economic growth. According to the results shown in Table 2, the MCQR-ALASSO method identified four key variables: highways and streets ("hwy") and personal capital stock ("pc") demonstrated an inverted "U" shaped relationship with economic growth, while labor input ("emp") exhibited a "U" shaped relationship. Meanwhile, the impact of water supply and sewage systems ("water") and unemployment rate ("unemp") on economic growth was considered negligible. In contrast, the LS-ALASSO method identified seven significant variables, showing clear differences in the variables selected by the two methods. This discrepancy mainly stems from assumptions about the distribution of model errors. The MCQR-ALASSO method, possibly due to its better handling of errors from non-normal distributions, identified variables more closely related to economic growth. This indicates that

choosing the appropriate statistical method is particularly crucial when dealing with economic data that may deviate from a normal error distribution. By improving upon composite quantile regression, MCQR-ALASSO offers an effective mechanism for variable selection, capable of identifying factors that significantly impact the target variable within complex data structures. Therefore, it is essential to employ advanced methods adaptable to data characteristics when analyzing issues like economic growth, to ensure the accuracy and reliability of research finding.

## REFERENCES

- [1] Kock , B. (2013).Oracle efficient variable selection in random and fixed effects panel data models. *Econometric Theory*, 29(1), 115-152. <https://doi.org/10.1017/S0266466612000230>.
- [2] Kock ,B. Tang , H. (2019).Uniform inference in high-dimensional dynamic panel data models with approximately sparse fixed effects. *Econometric Theory*,35(2), 295-359. <https://doi.org/10.48550/arXiv.1501.00478>.
- [3] Li ,Y., Zeng, X .B. (2014).Research on the penalized likelihood variable selection method for panel data models. *Statistical Research*, 31(03), 83-89. <https://doi.org/10.19343/j.cnki.11-1302/c.2014.03.012>.
- [4] Tang, L. Z., Li Y. J., Zhao L J. (2020).Research on the Bayesian Elastic Net quantile regression method for panel data and its application. *Statistical Research* , 37(03),94-113. <https://doi.org/10.19343/j.cnki.11-1302/c.2020.03.008>.
- [5] Lv, Y. Z., Zhang, R. Q., Zhao, W. H., Liu, J. C. (2014). Composite Quantile Regression and Variable Selection for Partial Linear Single-Index Models. *Scientia Sinica Mathematica*, 44(12), 1299-1322.
- [6] Wang, K. N., Li, S. M., Lin, L. (2020). Composite Quantile Regression and Variable Selection for Longitudinal Data Based on the Copula Function. *Scientia Sinica Mathematica*, 50(8), 1097-1116.
- [7] Weng, Y. L., Yu, P., Zhang, Z. Z. (2019). Composite Quantile Estimation for Functional Linear Models with Dependent Errors. *Journal of Applied Probability and Statistics*, 35(4), 360-372.
- [8] Wu, D. S., Yu, L., Yang, Y. P. (2019). The Impact of China's Foreign Trade on Economic Growth - Based on an Improved Two-Stage Panel Data Quantile Regression. *Journal of Mathematical Statistics and Management*, 38(4), 571-579.
- [9] Liu, H. L. (2016). *Studies on Statistical Models Based on Composite Quantile Regression Methods*. Chongqing: Chongqing University.
- [10] Li, G. R., Lian, H., Lai, P., Peng, H. (2015). Variable Selection for Fixed Effects Varying Coefficient Models. *Acta Mathematica Sinica (English Series)*, 31(1), 91-110. <https://doi.org/10.1007/s10114-015-3159-2>.
- [11] Wei, J. (2019). *High-Dimensional Testing Research in Panel Data Models*. China Economics Publishing House.
- [12] Gu, Z. T., Song, Z. F., Li, Y. (2020). Research on the Construction of Enhanced Index Funds Based on LASSO Variable Selection and Multifactor Models. *Journal of Mathematical Statistics and Management*, 39(3), 417-428. <https://doi.org/10.13860/j.cnki.sltj.20200122-006>.