



Improved Empirical Likelihood Inference for Multiplicative Regression with Independent and Longitudinal Data

Jiahui Xu¹, Jing Zhong¹, Yuying Xia¹, Minmin Zhang¹, Wei Chen^{2*}

¹Suzhou Institute of Technology, Jiangsu University of Science and Technology, Suzhou, China

²Campus of Zhangjiagang, Jiangsu University of Science and Technology, Suzhou, China

Email: *chenweixiyang@sina.com

How to cite this paper: Xu, J.H., Zhong, J., Xia, Y.Y., Zhang, M.M. and Chen, W. (2024) Improved Empirical Likelihood Inference for Multiplicative Regression with Independent and Longitudinal Data. *Open Access Library Journal*, **11**: e11919.

<https://doi.org/10.4236/oalib.1111919>

Received: July 8, 2024

Accepted: August 12, 2024

Published: August 15, 2024

Copyright © 2024 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Multiplicative regression model has been proven to be an excellent model for analyzing data with positive responses. When constructing the confidence regions of the regression parameters, one had to either directly estimate the asymptotic covariance matrix involving the estimation of the unknown density function of the model error, then the normal approximation can be conducted, or resort to the time-consuming resampling methods to avoid the difficulty of estimating the covariance matrix. Recently, an empirical likelihood (EL) approach has been proposed and has comparable performances when the sample size is moderate or large. However, all these methods become inefficient and unsatisfactory when the sample size is small. This paper proposed the adjusted empirical likelihood (AEL) approach to improve the performance of the EL combined with the least absolute relative error and the least product relative error criteria for independent data and longitudinal data with small sample sizes, respectively. It is shown that the adjusted empirical log-likelihood ratio is asymptotically Chi-squared distributed. Simulation studies indicate that the proposed AEL method performs better than EL method for small sample sizes and is as efficient as EL method when the sample size is moderate.

Subject Areas

Mathematical Statistics

Keywords

Adjusted Empirical Likelihood, Confidence Regions, Longitudinal Data, Multiplicative Regression, Relative Error

1. Introduction

In this paper, we consider the multiplicative regression models:

$$Y_i = \exp(\boldsymbol{\beta}^T \mathbf{X}_i) \varepsilon_i, \quad (1)$$

where Y is the response variable taking positive values, \mathbf{X} is a p -dimensional vector of covariates including the intercept as its first component, $\boldsymbol{\beta}$ is the corresponding p -dimensional vector of unknown regression coefficients, and ε is the positive error. When taking logarithm transformation on both sides of (1), the model becomes the known linear regression model relating $\log(Y)$ to \mathbf{X} , and the least square method or the least absolute deviation method may be applied. A common feature of both methods mentioned above is that the absolute error criterion is employed. However, as argued by [1], in many practical applications, the relative error is more of interest, and some earlier discussions and applications based on the relative error criterion can be found in the references of [1]. Based on a random sample (Y_i, X_i) from the model (1), [1] proposed the novel least absolute relative error criterion to estimate the regression parameters in model (1) by minimizing the following objective function

$$LARE(\boldsymbol{\beta}) = \sum_{i=1}^n \left\{ \left| \frac{Y_i - \exp(\mathbf{X}_i^T \boldsymbol{\beta})}{Y_i} \right| + \left| \frac{Y_i - \exp(\mathbf{X}_i^T \boldsymbol{\beta})}{\exp(\mathbf{X}_i^T \boldsymbol{\beta})} \right| \right\}, \quad (2)$$

which is conceptually meaningful and has some appealing merits such as the scale-free and unit-free characters. Under some regularity conditions, they also established the resulting estimators are consistent and asymptotical normality. However, due to the computational difficulty mainly caused by the non-smoothness of the loss function $LARE(\boldsymbol{\beta})$, [2] further investigated the least product relative error criterion, by which the loss function is defined as

$$LPRE(\boldsymbol{\beta}) = \sum_{i=1}^n \left\{ \left| \frac{Y_i - \exp(\mathbf{X}_i^T \boldsymbol{\beta})}{Y_i} \right| \times \left| \frac{Y_i - \exp(\mathbf{X}_i^T \boldsymbol{\beta})}{\exp(\mathbf{X}_i^T \boldsymbol{\beta})} \right| \right\}. \quad (3)$$

They also showed the estimators enjoy some asymptotical properties and that the simple plug-in variance estimation is reasonable. Following their works, some authors continued to extend the model (1) to some more flexible semi-parametric models, such as the partially linear MM ([3]-[6]), single-index MM ([7]-[9]), varying coefficient MM ([10]) and partially linear varying index MM ([11]).

However, it should be noted that LARE estimators' asymptotic covariance matrixes involve the unknown density function of the error term, which makes the construction of the confidence region intervals of the regression parameters not straightforward. One way is to directly estimate the covariance via some nonparametric methods such as the kernel smooth method, but the efficiency may be unsatisfactory when the sample size is not large enough. As done in the papers mentioned above, one can resort to a resampling approach, also named

the random weighting method, which is computationally intensive and time-consuming, especially when the dimensionality of the covariates is more than one. What's worse, both alternatives perform poorly when the sample size is small, as presented in the results of [12]-[14].

To overcome those drawbacks mentioned above, in this paper, we follow the line of [15] to extend the AEL approach to the multiplicative model with independent data and longitudinal data, combining with the LARE and LPRE criteria, respectively. It is shown that the adjusted empirical log-likelihood ratio function at the true value of regression parameters converges weakly to the standard Chi-squared distribution. As will be seen in the simulation studies later, the improved coverage probability is achieved when the sample size is small, and becomes comparable with classical EL method when the sample size is moderate or large. This phenomenal turn-up for the analysis of both independent data and longitudinal data.

The paper is organized as follows. In Section 2, we introduce the AEL method for the multiplicative model with independent data based on the LARE criterion, and with longitudinal data based on the LPRE criterion, respectively. In addition, some asymptotic properties are provided, which are used to construct the confidence regions of regression parameters and make hypothesis tests. Extensive simulation studies are conducted to demonstrate the usefulness of the proposal and results are presented in Section 3. Finally, several conclusions and remarks are given in Section 4, and proofs are deferred to the Appendix.

2. Main Results

2.1. AEL for LARE Estimation with Independent Data

Suppose that (Y_i, X_i) is a random sample from the model (1). Then the LARE estimator can be rewritten as solution to the following estimating equation

$$0 = \sum_{i=1}^n [\varepsilon_i(\boldsymbol{\beta}) + \varepsilon_i^{-1}(\boldsymbol{\beta})] \text{sgn}(\varepsilon_i(\boldsymbol{\beta}) - 1) X_i, \tag{4}$$

where $\varepsilon_i(\boldsymbol{\beta}) = Y_i / \exp(X_i^T \boldsymbol{\beta})$ and $\text{sgn}(\cdot)$ denotes the sign function. Write $Z_i(\boldsymbol{\beta}) = [\varepsilon_i(\boldsymbol{\beta}) - \varepsilon_i^{-1}(\boldsymbol{\beta})] \text{sgn}(\varepsilon_i(\boldsymbol{\beta}) - 1) X_i$, $i = 1, \dots, n$. Under the later condition (A4), it holds that $E(Z_i(\boldsymbol{\beta}_0)) = 0$. Inspired by the EL approach, for a given $\boldsymbol{\beta}$, the EL ratio of $\boldsymbol{\beta}$ is defined as

$$R_n(\boldsymbol{\beta}) = \sup \left\{ \prod_{i=1}^n p_i : \sum_{i=1}^n p_i Z_i(\boldsymbol{\beta}) = 0, \sum_{i=1}^n p_i = 1, p_i \geq 0 \right\}.$$

According to the findings in [15], $R_n(\boldsymbol{\beta})$ is well defined only if 0 is inside of the convex hull of $\{Z_i(\boldsymbol{\beta}), i = 1, \dots, n\}$. To avoid the computational constraints, Chen (2006) proposed the adjusted EL approach. Explicitly, define $\bar{Z}_n(\boldsymbol{\beta}) = \sum_{i=1}^n Z_i(\boldsymbol{\beta}) / n$ and $Z_{n+1}(\boldsymbol{\beta}) = -a_n \bar{Z}_n(\boldsymbol{\beta})$, where $a_n = \max(1, \log(n)/2)$. Then the AEL ratio function is

$$\tilde{R}_n(\boldsymbol{\beta}; a_n) = \sup \left\{ \prod_{i=1}^{n+1} p_i : \sum_{i=1}^{n+1} p_i Z_i(\boldsymbol{\beta}) = 0, \sum_{i=1}^{n+1} p_i = 1, p_i \geq 0 \right\}. \tag{5}$$

For each given β , by the Lagrange multiplier method, (5) achieves its maximum at

$$p_i = \frac{1}{n+1} \frac{1}{1 + \lambda^T Z_i(\beta)}, i = 1, \dots, n+1,$$

where λ is the Lagrange multipliers and satisfies

$$\frac{1}{n+1} \sum_{i=1}^{n+1} \frac{Z_i(\beta)}{1 + \lambda^T Z_i(\beta)} = 0.$$

Furthermore, the proposed adjusted empirical log-likelihood ratio is given by

$$l^{ad}(\beta) = -2 \log(\tilde{R}_n(\beta; a_n)) = 2 \sum_{i=1}^{n+1} \log(1 + \lambda^T Z_i(\beta)).$$

To establish the asymptotic distribution of the proposed statistic, some regular conditions are required as follows.

(A1) The random error ε has a density function f , which is continuous in a neighborhood of 1.

(A2) $P(\varepsilon > 0) = 1$.

(A3) X is bounded and does not concentrate on any hyperplane of $p-1$ dimension.

(A4) $E[(\varepsilon + \varepsilon^{-1}) \text{sgn}(\varepsilon - 1)] = 0$.

(A5) $E[(\varepsilon + \varepsilon^{-1})^2] < \infty$.

Conditions (A1)-(A3) are common requirements in the study of multiplicative regression models. (A4) is an identification condition for the LARE estimation, which is similar to the zero mean condition in the classical linear mean regression. (A5) is required for proof, which is also used in [12].

Theorem 1 Suppose conditions (A1)-(A5) hold and $a_n = o(n)$. As $n \rightarrow \infty$, $l^{ad}(\beta_0) = -2 \log(\tilde{R}_n(\beta_0; a_n))$ converges in distribution to a standard chi-squared random variables with p degrees of freedom.

For fixed $c \in (0, 1)$, a natural asymptotic $1-c$ confidence region for β_0 based on the empirical likelihood ratio is given by

$$\mathcal{C}_1 = \{ \beta : l^{ad}(\beta) \leq \chi_{p,c}^2 \},$$

where $\chi_{p,c}^2$ is the upper c -quantile of the chi-squared distribution in Theorem 1.

2.2. AEL for LPRE Estimation with Longitudinal Data

Consider a longitudinal study consisting of n subjects with the j -th subject having n_j observations. For each $i = 1, \dots, n$ and $j = 1, \dots, n_j$, as notations presented in model (1), (Y_{ij}, X_{ij}) is a random sample from the model

$$Y_{ij} = \exp(\beta^T X_{ij}) \varepsilon_{ij}, \tag{6}$$

where X_{ij} are known bounded design vectors, errors ε_{ij} across subjects are independent but may be dependent within the same subject. As proposed in [13], the LPRE estimator can be defined as solution to the following estimating equation

$$0 = \sum_{i=1}^n \sum_{j=1}^{n_i} [-\varepsilon_{ij}(\beta) + \varepsilon_{ij}^{-1}(\beta)] X_{ij}, \tag{7}$$

where $\varepsilon_{ij}(\boldsymbol{\beta}) = Y_{ij} / \exp(\mathbf{X}_{ij}^T \boldsymbol{\beta})$. For notation simplicity, we write $W_i(\boldsymbol{\beta}) = \sum_{j=1}^{n_i} [\varepsilon_{ij}(\boldsymbol{\beta}) - \varepsilon_{ij}^{-1}(\boldsymbol{\beta})] \mathbf{X}_{ij}$, $i = 1, \dots, n$. $\bar{W}_n(\boldsymbol{\beta}) = \sum_{i=1}^n W_i(\boldsymbol{\beta}) / n$ and $W_{n+1}(\boldsymbol{\beta}) = -a_n \bar{W}_n(\boldsymbol{\beta})$, where $a_n = \max(1, \log(n)/2)$. Under the later condition (B4), it holds that $E(W_i(\boldsymbol{\beta}_0)) = 0$. Following the idea of the block EL approach in [16], for a given $\boldsymbol{\beta}$, the AEL ratio function of $\boldsymbol{\beta}$ is defined as

$$\tilde{R}_n(\boldsymbol{\beta}) = \sup \left\{ \prod_{i=1}^{n+1} p_i : \sum_{i=1}^n p_i W_i(\boldsymbol{\beta}) = 0, \sum_{i=1}^{n+1} p_i = 1, p_i \geq 0 \right\}. \tag{8}$$

For each given $\boldsymbol{\beta}$, by the Lagrange multiplier method, (8) achieves its maximum at

$$p_i = \frac{1}{n+1} \frac{1}{1 + \lambda^T W_i(\boldsymbol{\beta})}, i = 1, \dots, n+1,$$

where λ is the Lagrange multipliers and satisfies

$$\frac{1}{n+1} \sum_{i=1}^{n+1} \frac{W_i(\boldsymbol{\beta})}{1 + \lambda^T W_i(\boldsymbol{\beta})} = 0.$$

Furthermore, the proposed adjusted empirical log-likelihood ratio is given by

$$\tilde{l}^{ad}(\boldsymbol{\beta}) = -2 \log(\tilde{R}_n(\boldsymbol{\beta}; a_n)) = 2 \sum_{i=1}^{n+1} \log(1 + \lambda^T W_i(\boldsymbol{\beta})).$$

To establish the asymptotic distribution of the proposed statistic, some regular conditions are required as follows.

- (B1) The random errors ε_{ij} have a common density function f and $P(\varepsilon_{ij} > 0) = 1$.
- (B2) $E[\varepsilon - \varepsilon^{-1}] = 0$.
- (B3) There exists a constant $\delta > 0$ such that $E[(\varepsilon + \varepsilon^{-1})^{2+\delta}] < \infty$.
- (B4) $\max_{i \geq 1} n_i < \infty$, and $D_n = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \mathbf{X}_{ij} \mathbf{X}_{ij}^T$ converges to a positive definite matrix D .

Conditions (B1)-(B4) are also used in [13].

Theorem 2 Suppose conditions (B1)-(B4) hold and $a_n = o(n)$. As $n \rightarrow \infty$, $\tilde{l}^{ad}(\boldsymbol{\beta}_0) = -2 \log(\tilde{R}_n(\boldsymbol{\beta}_0; a_n))$ converges in distribution to a standard chi-squared random variables with p degrees of freedom.

For fixed $c \in (0, 1)$, a natural asymptotic $1-c$ confidence region for $\boldsymbol{\beta}_0$ based on the empirical likelihood ratio is given by

$$\mathfrak{C}_2 = \{ \boldsymbol{\beta} : \tilde{l}^{ad}(\boldsymbol{\beta}) \leq \chi_{p,c}^2 \},$$

where $\chi_{p,c}^2$ is the upper c -quantile of the chi-squared distribution in Theorem 1.

3. Simulation Studies

In this section, we investigate the finite sample performance of the proposed adjusted empirical likelihood method (AEL) with the ordinary empirical likelihood method (EL) for analyzing independent and longitudinal data, respectively. For comparison, we compute the coverage probabilities of the simultaneous confi-

dence regions for all regression coefficients under various settings. Besides, we also consider the power behavior of the methods mentioned above.

3.1. Comparisons for Independent Data

Similar to the model used in Section 3 of [12], the studies in this part are based on the following model,

$$Y_i = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{12i}) \varepsilon_i, \quad (9)$$

where X_{1i} and X_{2i} are independent standard normal random variables and $\beta_0^T = (\beta_0, \beta_1, \beta_2)$ take a value (1, 1, 1). The random error ε_i is assumed to be independent of (X_{1i}, X_{2i}) , and its logarithm $\log(\varepsilon_i)$ comes from two different distributions: one is the standard normal distribution $N(0,1)$, the other is the uniform distribution $U(-2,2)$, which are denoted by $LN(0,1)$ and $LU(-2,2)$, respectively. As is proposed in Section 3, the confidence region at nominal level α for AEL and EL are constructed by $I_{1,\alpha}$ and $I_{2,\alpha}$, which are defined as

$$I_{1,\alpha} = \{\beta : l^{ad}(\beta) \leq \chi_\alpha^2(3)\},$$

$$I_{2,\alpha} = \{\beta : -2\log(R_n(\beta)) \leq \chi_\alpha^2(3)\},$$

respectively. Here and after, α is chosen to 0.05 or 0.1, and $\chi_\alpha^2(3)$ is the 100 α % -level quantile of a Chi-squared distribution $\chi^2(3)$.

Furthermore, based on the results in Theorem 1, we evaluate the power of the proposed methods by testing the hypothesis $H_0 : \beta_0^T = (b_0, b_1, b_2)$ against $H_1 : \beta_0^T \neq (b_0, b_1, b_2)$, where we consider four combinations of $(b_0, b_1, b_2)^T$, namely, (0.5, 0.5, 0.5), (0.8, 0.8, 0.8), (0.9, 0.9, 0.9) and (0.8, 0.5, 0.8). For each specific setting, 1000 independent replications are simulated based on the sample sizes 25, 50, 100, 200, 300, 500, 1000, respectively. In contrast to the simulation design in [12] where only the moderate sample sizes are studied, we also consider the small sample size cases.

Simulation results are summarized in **Table 1**, **Table 2**. It can be seen from **Table 1** that both AEL and EL methods provide accurate coverage probabilities of confidence regions when the sample size is moderate (≥ 200). Especially, the coverage probabilities for the two methods are close to the nominal level and usually comparable as the sample size increases. But when the sample size is small (≤ 100), the coverage probabilities via AEL method are much better than that of the EL method under all settings, although there still exists a certain gap with the nominal levels. As presented in **Table 2**, when the null hypothesis H_0 is far away from the true value of the regression coefficients, such as the cases $H_0 : \beta_0^T = (0.5, 0.5, 0.5)$ or (0.8, 0.8, 0.8), both tests are powerful and comparable, even when the sample size is small, or there exists only one component of $(b_0, b_1, b_2)^T$ is far away from the true value of the regression parameters, such as the case $H_0 : \beta_0^T = (0.8, 0.5, 0.8)$. Meanwhile, when H_0 lies in the nearby of the true value of the regression coefficients, the AEL test is slightly less powerful than the EL test for small sample sizes and the difference between them becomes

less and negligible with the increase in the sample size. This result is not strange and similar performance still occurs in **Table 3** of [12], which shows that the confidence region based on the AEL method contains the confidence region based on the EL method and the cost is not large and deserved compared to the improvement of the coverage probability for samples with small sample sizes.

Table 1. Coverage probabilities of the confidence region for β_0 with independent data.

n	$\varepsilon \sim LN(0,1)$				$\varepsilon \sim LU(-2,2)$			
	0.9		0.95		0.9		0.95	
	AEL	EL	AEL	EL	AEL	EL	AEL	EL
25	0.850	0.773	0.908	0.849	0.853	0.798	0.915	0.856
50	0.860	0.826	0.904	0.881	0.895	0.864	0.924	0.894
100	0.870	0.855	0.931	0.922	0.871	0.852	0.951	0.944
200	0.874	0.868	0.942	0.933	0.908	0.898	0.945	0.939
300	0.883	0.879	0.940	0.933	0.897	0.893	0.961	0.959
500	0.885	0.879	0.948	0.947	0.918	0.913	0.953	0.951
1000	0.901	0.899	0.944	0.943	0.912	0.911	0.941	0.939

Table 2. Powers of tests at nominal levels 0.1 and 0.05 for $H_0 : \beta_0 = (b_0, b_1, b_2)$ with independent data.

(b_0, b_1, b_2)	$\varepsilon \sim LN(0,1)$				$\varepsilon \sim LU(-2,2)$			
	0.1		0.05		0.1		0.05	
	AEL	EL	AEL	EL	AEL	EL	AEL	EL
(0.5, 0.5, 0.5)								
$n = 50$	0.998	0.999	0.999	0.999	1	1	0.998	0.998
$n = 100$	1	1	1	1	1	1	1	1
$n = 200$	1	1	1	1	1	1	1	1
(0.8, 0.8, 0.8)								
$n = 50$	0.638	0.682	0.493	0.544	0.579	0.623	0.425	0.485
$n = 100$	0.885	0.895	0.823	0.841	0.849	0.865	0.773	0.799
$n = 200$	0.986	0.987	0.988	0.988	0.988	0.990	0.983	0.986
(0.9, 0.9, 0.9)								
$n = 50$	0.285	0.326	0.201	0.240	0.240	0.265	0.142	0.173
$n = 100$	0.396	0.426	0.302	0.322	0.349	0.374	0.250	0.275
$n = 200$	0.637	0.651	0.493	0.508	0.551	0.567	0.459	0.480
(0.8, 0.5, 0.8)								
$n = 50$	0.994	0.995	0.987	0.990	0.986	0.993	0.969	0.977
$n = 100$	1	1	1	1	1	1	1	1
$n = 200$	1	1	1	1	1	1	1	1

Table 3. Coverage probabilities of the confidence region for β_0 with longitudinal data.

n	S1				S2			
	0.9		0.95		0.9		0.95	
	AEL	EL	AEL	EL	AEL	EL	AEL	EL
$\varepsilon \sim LN(0,1)$								
25	0.850	0.811	0.931	0.886	0.842	0.797	0.920	0.872
50	0.882	0.850	0.920	0.898	0.856	0.823	0.916	0.898
100	0.877	0.867	0.934	0.925	0.888	0.875	0.925	0.917
200	0.902	0.893	0.945	0.938	0.901	0.897	0.938	0.938
300	0.889	0.884	0.946	0.944	0.896	0.890	0.948	0.944
500	0.887	0.885	0.950	0.948	0.896	0.892	0.950	0.946
$\varepsilon \sim LU(-2,2)$								
25	0.895	0.848	0.949	0.922	0.878	0.833	0.929	0.882
50	0.898	0.879	0.953	0.943	0.872	0.854	0.939	0.920
100	0.889	0.880	0.958	0.949	0.900	0.888	0.956	0.952
200	0.906	0.901	0.947	0.940	0.895	0.889	0.940	0.935
300	0.908	0.905	0.943	0.943	0.888	0.883	0.963	0.960
500	0.917	0.914	0.949	0.948	0.913	0.910	0.954	0.952

3.2. Comparisons for Longitudinal Data

Similar to the simulation design in [13], we consider the following model,

$$Y_{ij} = \exp(\beta_1 X_{1ij} + \beta_2 X_{2ij}) \varepsilon_{ij}, i = 1, \dots, n; j = 1, \dots, 5, \tag{10}$$

where X_{1ij} and X_{2ij} are independent standard normal random variables and $(\beta_1, \beta_2) = (1, 1)$. For each i , the random error ε_{ij} is assumed to be independent of (X_{1ij}, X_{2ij}) , and its logarithm $\log(\varepsilon_{ij})$ comes from two different distributions, which are denoted by $LN(0, 1)$ and $LU(-2, 2)$, respectively. Furthermore, the covariance matrix of the logarithm vector $(\log(\varepsilon_{i1}), \dots, \log(\varepsilon_{i5}))$ has the form (σ_{jk}) , where $\sigma_{jk} = 0.5$ or $\sigma_{jk} = 0.5^{|j-k|}$. Denote these two matrices by S1 and S2, respectively.

Furthermore, based on the results in Theorem 2, we evaluate the power of the proposed methods by testing the hypothesis $H_0: \beta_0^T = (b_1, b_2)$ against $H_1: \beta_0^T \neq (b_1, b_2)$. For each specific setting, 1000 independent replications are simulated based on the sample sizes 25, 50, 100, 200, 300, 500, respectively. In contrast to the simulation design in [13], where only the moderate sample sizes are studied, we also consider the small sample size cases.

Table 3 presents the coverage probabilities of the confidence regions at normal levels of 0.9 and 0.95, respectively. Again, we find the fact that both AEL and EL methods provide accurate coverage probabilities of confidence regions when the sample size is moderate (≥ 200). Especially, the coverage probabilities for the two methods are close to the nominal level and usually comparable as the sample

size increases. But when the sample size is small (≤ 300), the coverage probabilities via AEL method are much better than that of the EL method under all settings, although there still exists a certain gap with the nominal levels.

4. Conclusion

This paper introduces an efficient and easy-to-implemented adjusted empirical likelihood-based method for constructing the confidence region of the regression parameters in the multiplicative regression model with independent and longitudinal data, respectively. The proposed procedures can avoid estimating the nuisance parameters and perform well even if when the size of sample is relatively small. In the future, extensions can be made to handle an increasing number of covariates as done in [14], and the cases where the response variables are missing or censored.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Chen, K., Guo, S., Lin, Y. and Ying, Z. (2010) Least Absolute Relative Error Estimation. *Journal of the American Statistical Association*, **105**, 1104-1112. <https://doi.org/10.1198/jasa.2010.tm09307>
- [2] Chen, K., Lin, Y., Wang, Z. and Ying, Z. (2016) Least Product Relative Error Estimation. *Journal of Multivariate Analysis*, **144**, 91-98. <https://doi.org/10.1016/j.jmva.2015.10.017>
- [3] Zhang, Q. and Wang, Q. (2013) Local Least Absolute Relative Error Estimating Approach for Partially Linear Multiplicative Model. *Statistica Sinica*, **23**, 1091-1116. <https://doi.org/10.5705/ss.2012.133>
- [4] Zhang, J., Feng, Z. and Peng, H. (2018) Estimation and Hypothesis Test for Partial Linear Multiplicative Models. *Computational Statistics & Data Analysis*, **128**, 87-103. <https://doi.org/10.1016/j.csda.2018.06.017>
- [5] Chen, Y. and Liu, H. (2021) A New Relative Error Estimation for Partially Linear Multiplicative Model. *Communications in Statistics-Simulation and Computation*, **52**, 4962-4980.
- [6] Chen, W. and Wan, M. (2023) SIMEX Estimation of Partially Linear Multiplicative Regression Model with Mismeasured Covariates. *Symmetry*, **15**, Article 1833. <https://doi.org/10.3390/sym15101833>
- [7] Liu, H. and Xia, X. (2018) Estimation and Empirical Likelihood for Single-Index Multiplicative Models. *Journal of Statistical Planning and Inference*, **193**, 70-88. <https://doi.org/10.1016/j.jspi.2017.08.003>
- [8] Zhang, J., Zhu, J. and Feng, Z. (2019) Estimation and Hypothesis Test for Single-Index Multiplicative Models. *Test*, **28**, 242-268. <https://doi.org/10.1007/s11749-018-0586-2>
- [9] Zhang, J., Cui, X. and Peng, H. (2020) Estimation and Hypothesis Test for Partial Linear Single-Index Multiplicative Models. *Annals of the Institute of Statistical Mathematics*, **72**, 699-740. <https://doi.org/10.1007/s10463-019-00706-6>
- [10] Hu, D.H. (2019) Local Least Product Relative Error Estimation for Varying Coeffi-

- cient Multiplicative Regression Model. *Acta Mathematicae Applicatae Sinica, English Series*, **35**, 274-286. <https://doi.org/10.1007/s10255-018-0794-2>
- [11] Chen, Y., Liu, H. and Ma, J. (2022) Local Least Product Relative Error Estimation For Single-Index Varying-Coefficient Multiplicative Model With Positive Responses. *Journal of Computational and Applied Mathematics*, **415**, Article 114478. <https://doi.org/10.1016/j.cam.2022.114478>
- [12] Li, Z., Lin, Y., Zhou, G. and Zhou, W. (2014) Empirical Likelihood for Least Absolute Relative Error Regression. *Test*, **23**, 86-99. <https://doi.org/10.1007/s11749-013-0343-5>
- [13] Xiao, L. and Wang, W. (2017) Asymptotics for Least Product Relative Error Estimation and Empirical Likelihood with Longitudinal Data. *Journal of the Korean Statistical Society*, **46**, 375-389. <https://doi.org/10.1016/j.jkss.2016.12.001>
- [14] Li, Z., Liu, Y. and Liu, Z. (2017) Empirical Likelihood and General Relative Error Criterion with Divergent Dimension. *Statistics*, **51**, 1006-1022. <https://doi.org/10.1080/02331888.2017.1296443>
- [15] Chen, J., Variyath, A.M. and Abraham, B. (2008) Adjusted Empirical Likelihood and Its Properties. *Journal of Computational and Graphical Statistics*, **17**, 426-443. <https://doi.org/10.1198/106186008X321068>
- [16] You, J., Chen, G. and Zhou, Y. (2006) Block Empirical Likelihood for Longitudinal Partially Linear Regression Models. *Canadian Journal of Statistics*, **34**, 79-96. <https://doi.org/10.1002/cjs.5550340107>

Appendix A. Proofs of Theorems 1-2

Proof of Theorem 1. Under conditions (A1)-(A5), according to the lemma 1 in [12], we have $\bar{Z}_n(\beta_0) = O_p(1/\sqrt{n})$, $Z^*(\beta_0) = \max_{1 \leq i \leq n} \|Z_i(\beta_0)\| = o_p(\sqrt{n})$ and $\bar{W}(\beta_0) = \frac{1}{n} \sum_{i=1}^n Z_i(\beta_0) Z_i(\beta_0)^T = E[(\varepsilon + \varepsilon^{-1})^2] E[XX^T] + o_p(1) \doteq S + o_p(1)$, where $\|\cdot\|$ denotes the Euclidean norm.

Let $\rho = \|\lambda\|$ and $u = \lambda/\rho$, which implies $\|u\| = 1$. Note that

$$0 = \sum_{i=1}^{n+1} \frac{Z_i(\beta)}{1 + \lambda^T Z_i(\beta)} = \sum_{i=1}^{n+1} Z_i(\beta) - \sum_{i=1}^{n+1} \frac{Z_i(\beta) Z_i(\beta)^T \lambda}{1 + \lambda^T Z_i(\beta)}.$$

Multiplying u^T/n to both sides of (3), we get

$$0 = \frac{u^T}{n} \sum_{i=1}^{n+1} \frac{Z_i(\beta)}{1 + \lambda^T Z_i(\beta)} = \frac{u^T}{n} \sum_{i=1}^{n+1} Z_i(\beta) - \frac{\rho}{n} \sum_{i=1}^{n+1} \frac{(u^T Z_i(\beta))^2}{1 + \lambda^T Z_i(\beta)} \doteq I_1 - I_2.$$

Some algebraic calculations yield that $I_1 = u^T \bar{Z}_n(\beta)(1 - a_n/n)$. From the inequality $u^T Z_i(\beta) \leq \|u\| \|Z_i(\beta)\| = \|Z_i(\beta)\| \leq Z^*(\beta_0)$, it hold that

$$0 < 1 + \lambda^T Z_i(\beta) \leq 1 + |\lambda^T Z_i(\beta)| \leq 1 + \rho Z^*(\beta_0),$$

$$\frac{1}{1 + \lambda^T Z_i(\beta)} \geq \frac{1}{1 + \rho Z^*(\beta_0)}.$$

Thus

$$I_2 \geq \frac{\rho}{n} \sum_{i=1}^n \frac{(u^T Z_i(\beta))^2}{1 + \lambda^T Z_i(\beta)} = \frac{\rho}{n(1 + \rho Z^*(\beta_0))} \sum_{i=1}^n (u^T Z_i(\beta))^2.$$

Combining the results above, we have

$$\left| \frac{a_n}{n} u^T \bar{Z}_n(\beta) \right| \leq \frac{a_n}{n} \|\bar{Z}_n(\beta)\| = O_p(n^{-3/2} a_n),$$

and

$$0 \leq u^T \bar{Z}_n(\beta) - \frac{\rho}{n(1 + \rho Z^*(\beta_0))} \sum_{i=1}^n (u^T Z_i(\beta))^2 + O_p(n^{-3/2} a_n).$$

Conditions (A3) and (A5) imply that S is positive definite. Using the fact that

$$\sum_{i=1}^n (u^T Z_i(\beta))^2 / n = u^T \left(\sum_{i=1}^n Z_i(\beta) Z_i(\beta)^T / n \right) u = u^T S u + o_p(1) \geq \lambda_1 + o_p(1),$$

where $\lambda_1 > 0$ denotes the smallest eigenvalue of S , it holds that

$$\begin{aligned} \frac{\rho}{n(1 + \rho Z^*(\beta_0))} &\leq \left[u^T \bar{Z}_n(\beta) + O_p(n^{-3/2} a_n) \right] / \frac{1}{n} \sum_{i=1}^n (u^T Z_i(\beta))^2 \\ &\leq \frac{u^T \bar{Z}_n(\beta) + O_p(n^{-3/2} a_n)}{\lambda_1 + o_p(1)}. \end{aligned}$$

Together with $u^T \bar{Z}_n(\beta) = O_p(n^{-1/2})$ and $a_n = o(n)$, it follows that $\rho = \|\lambda\| = O_p(n^{-1/2})$. Recall that

$$\begin{aligned}
 0 &= \sum_{i=1}^{n+1} \frac{1}{n} \frac{Z_i(\boldsymbol{\beta})}{1 + \lambda^T Z_i(\boldsymbol{\beta})} \\
 &= \sum_{i=1}^{n+1} \frac{1}{n} \frac{1 - (Z_i(\boldsymbol{\beta})^T \lambda)^2 + (Z_i(\boldsymbol{\beta})^T \lambda)^2}{1 + \lambda^T Z_i(\boldsymbol{\beta})} \\
 &= \frac{1}{n} \sum_{i=1}^{n+1} Z_i(\boldsymbol{\beta}) - \left(\frac{1}{n} \sum_{i=1}^{n+1} Z_i(\boldsymbol{\beta}) Z_i(\boldsymbol{\beta})^T \lambda \right) + \frac{1}{n} \sum_{i=1}^{n+1} \frac{Z_i(\boldsymbol{\beta}) (Z_i(\boldsymbol{\beta})^T \lambda)^2}{1 + \lambda^T Z_i(\boldsymbol{\beta})} \\
 &= \bar{Z}_n(\boldsymbol{\beta}) (1 - a_n/n) - \left(\bar{W}(\boldsymbol{\beta}) + a_n^2 \|\bar{Z}_n(\boldsymbol{\beta})\|^2 / n \right) \lambda \\
 &\quad + \frac{1}{n} \sum_{i=1}^n \frac{Z_i(\boldsymbol{\beta}) (Z_i(\boldsymbol{\beta})^T \lambda)^2}{1 + \lambda^T Z_i(\boldsymbol{\beta})} + \frac{1}{n} \frac{Z_{n+1}(\boldsymbol{\beta}) (Z_{n+1}(\boldsymbol{\beta})^T \lambda)^2}{1 + \lambda^T Z_{n+1}(\boldsymbol{\beta})},
 \end{aligned}$$

we have $0 = \bar{Z}_n(\boldsymbol{\beta}) - \bar{W}(\boldsymbol{\beta}) \lambda + o_p(n^{-1/2})$. Then it follows that $\lambda = \bar{W}^{-1}(\boldsymbol{\beta}) \bar{Z}_n(\boldsymbol{\beta}) + o_p(n^{-1/2})$. Furthermore,

$$\begin{aligned}
 l^{ad}(\boldsymbol{\beta}) &= 2 \sum_{i=1}^{n+1} \log(1 + \lambda^T Z_i(\boldsymbol{\beta})) \\
 &= 2 \sum_{i=1}^{n+1} \left[\lambda^T Z_i(\boldsymbol{\beta}) - (\lambda^T Z_i(\boldsymbol{\beta}))^2 / 2 \right] + o_p(1) \\
 &= 2 \lambda^T \sum_{i=1}^{n+1} Z_i(\boldsymbol{\beta}) - \lambda^T \sum_{i=1}^{n+1} Z_i(\boldsymbol{\beta}) Z_i(\boldsymbol{\beta})^T \lambda + o_p(1) \\
 &= 2(n - a_n) \left[\bar{W}^{-1}(\boldsymbol{\beta}) \bar{Z}_n(\boldsymbol{\beta}) + o_p(n^{-1/2}) \right]^T \bar{Z}_n(\boldsymbol{\beta}) \\
 &\quad - \bar{Z}_n(\boldsymbol{\beta})^T \bar{W}^{-1}(\boldsymbol{\beta}) \left[n \bar{W}(\boldsymbol{\beta}) + a_n^2 \bar{Z}_n(\boldsymbol{\beta}) \bar{Z}_n(\boldsymbol{\beta})^T \right] \bar{W}^{-1}(\boldsymbol{\beta}) \bar{Z}_n(\boldsymbol{\beta}) + o_p(1) \\
 &= n \bar{Z}_n(\boldsymbol{\beta})^T \bar{W}^{-1}(\boldsymbol{\beta}) \bar{Z}_n(\boldsymbol{\beta}) + o_p(1) \\
 &= n \bar{Z}_n(\boldsymbol{\beta})^T S^{-1} \bar{Z}_n(\boldsymbol{\beta}) + o_p(1) \rightarrow_d \chi_p^2.
 \end{aligned}$$

As a result, The Theorem 1 is proved.

Proof of Theorem 2. The proof is similar to the line of Theorem 1 except some modifications as done in [13]. Hence it is omitted here.