Jackknife empirical likelihood for the correlation coefficient with additive distortion measurement errors

Article *in* Test · September 2024

DOI: 10.1007/s11749-024-00941-x

Da Chen
Georgia State University

Linlin Dai
Southwestern University of Finance and Economics

Yichuan Zhao
Georgia State University

TEST
https://doi.org/10.1007/s11749-024-00941-x

ORIGINAL PAPER

Jackknife empirical likelihood for the correlation coefficient with additive distortion measurement errors

Da Chen[1] · Linlin Dai[2] · Yichuan Zhao[1]

Received: 16 January 2022 / Accepted: 21 June 2024

Abstract
The correlation coefficient is fundamental in advanced statistical analysis. However, traditional methods of calculating correlation coefficients can be biased due to the existence of confounding variables. Such confounding variables could act in an additive or multiplicative fashion. To study the additive model, previous research has shown residual-based estimation of correlation coefficients. The powerful tool of empirical likelihood (EL) has been used to construct the confidence interval for the correlation coefficient. However, the methods so far only perform well when sample sizes are large. With small sample size situations, the coverage probability of EL, for instance, can be below 90% at confidence level 95%. On the basis of previous research, we propose new methods of interval estimation for the correlation coefficient using jackknife empirical likelihood, mean jackknife empirical likelihood and adjusted jackknife empirical likelihood. For better performance with small sample sizes, we also propose mean adjusted empirical likelihood. The simulation results show the best performance with mean adjusted jackknife empirical likelihood when the sample sizes are as small as 25. Real data analyses are used to illustrate the proposed approach.

Keywords Correlation coefficient · Distortion errors · Jackknife empirical likelihood · Adjusted jackknife empirical likelihood · Mean jackknife empirical likelihood · Mean adjusted jackknife empirical likelihood

Mathematics Subject Classification 62G10 · 62G20

1 Introduction

Confounding variables are common and inevitable during research, experiments, and real data analysis. The existence of such variables could lead to inflated bias and vari-

✉ Yichuan Zhao
yichuan@gsu.edu

1 Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303, USA

2 Center of Statistical Research, School of Statistics, Southwestern University of Finance and Economics, Chengdu 610000, Sichuan, China

Published online: 04 September 2024

ation, while affecting corresponding coefficients in linear regression. When studying the correlation coefficient between two variables, a confounding variable may exist in an additive way to the variables being studied, which could either underestimate or overestimate the true correlation coefficient between the variables of interest. The additive model is first introduced by Sentürk and Müller (2005) and can be modeled as:

$$\begin{cases} \tilde{X} = X + \psi(U), \\ \tilde{Y} = Y + \phi(U), \end{cases}$$

where $(\tilde{X}, \tilde{Y})$ are the observable variables, (X, Y) are the corresponding unobservable true values of interest that are independent of the observed confounding variable U. $\phi(U)$ and $\psi(U)$ are unknown functions of U with identifiability condition $E[\phi(U)] = E[\psi(U)] = 0$. Sentürk and Müller (2005) also proposed a model for multiplicative errors. The multiplicative error model suggests that $\tilde{X} = \psi(U)X$ and $\tilde{Y} = \phi(U)Y$ where $E[\phi(U)] = E[\psi(U)] = 1$. Different models of $\phi(\cdot)$ and $\psi(\cdot)$ have been studied under the scenario of multiplicative distortion measurement errors. Sentürk and Müller (2008) investigated linear and generalized linear models. Zhang et al. (2020) studied multiplicative regression models with distortion measurement errors and developed three kinds of estimators. By taking the logarithm of the response variable, a least squares estimator and a moment-based estimator can be calculated. Without the logarithmic transformation, a least product relative error estimator is proposed. As to the additive fashion of distortion measurement errors, Feng et al. (2020) established a residual based least squares estimator for linear regression models under restricted and unrestricted conditions. They also proposed a hypothesis testing method by introducing a test statistic according to the normalized difference of residual sums of squares under null and alternative hypotheses. In addition, an estimation and hypothesis testing method for the partial linear models are proposed by Zhang et al. (2017) under the additive fashion of distortion measurement errors.

To study the correlation coefficient between X and Y, Zhang et al. (2017) developed a direct plug-in estimator and a residual-based estimator. Under the method of direct plug-in estimator, estimators of $\phi(U)$ and $\psi(U)$, denoted as $\hat{\phi}(U)$ and $\hat{\psi}(U)$, are subtracted from the observed response and predictor to obtain calibrated X, Y used to be used in constructing the estimation for the correlation. On the other hand, based on the residuals

$$\begin{cases} e_{\tilde{X}U} = \tilde{X} - E[\tilde{X}|U], \\ e_{\tilde{Y}U} = \tilde{Y} - E[\tilde{Y}|U], \end{cases}$$

and the fact that $\rho(X, Y) = \rho(e_{\tilde{X}U}, e_{\tilde{Y}U})$, Zhang et al. (2017) proposed a residual estimation for $\rho(X, Y)$ using an empirical likelihood-based statistic and also constructed the confidence interval of $\rho(X, Y)$.

First introduced by Owen (1988, 1990, 2001), the empirical likelihood (EL) has shown its advantage as a non-parametric tool that does not need a distribution assumption. Huang and Zhao (2018) proposed to use empirical likelihood for bivariate survival function under univariate censoring. Cheng et al. (2012) introduced empirical likelihood inference for semiparametric additive isotonic regression. The method of adjusted empirical likelihood has been shown by Chen et al. (2008). The adjusted empirical likelihood has been applied to a variety of research fields since then. Zheng and Yu (2013) developed the adjusted empirical likelihood for the multivariate Cox model. Yu and Zhao (2019a) proposed empirical likelihood inference and adjusted empirical likelihood for semi-parametric transformation models with length-biased sampling. Wang et al. (2019) proposed the method of penalized empirical likelihood when dealing with the sparse Cox model. However, in situations with small sample sizes, the coverage probability and the confidence region could be inaccurate using empirical likelihood. Thus, Liang et al. (2019) proposed mean empirical likelihood (MEL) by constructing a pseudo dataset through the means of observed values. It has been shown that the method of MEL satisfies Wilks' theorem. Zhao et al. (2022) developed the sample empirical likelihood for general hypothesis tests on population parameters in the general estimating equation under complex survey data, and proposed a novel penalized sample empirical likelihood for variable selection. Bouzebda and Keziou (2024) investigated EL methods for the functional of nonparametric copula models. Yu and Bondell (2024) develop a fast and accurate approach to approximate posterior distributions in the Bayesian empirical likelihood framework. The idea is to combine the stochastic variational Bayes procedure with an adjusted empirical likelihood to resolve computational challenges. Fu et al. (2024) developed the robust penalized EL method for a sparse high-dimensional model in longitudinal data analysis. The proposed method can select important variables and estimating equations simultaneously, which enjoyed the robust and effective property.

When dealing with complicated statistics, the calculation of empirical likelihood can still be redundant. Thus, the jackknife empirical likelihood (JEL) approach was introduced by Jing et al. (2009) to simplify the application of empirical likelihood to complicated statistics. There exist various works on the applications of jackknife empirical likelihood. Sang et al. (2019) proposed JEL method for estimating Gini correlations. Lin et al. (2017) investigated the error variance using JEL in linear regression models. Liu and Liang (2017) developed JEL for the error variance in partially linear varying-coefficient errors-in-variables models. JEL is also introduced for the accelerated failure time model by Yu and Zhao (2019b). Then, Xu et al. (2022) extended the JEL to linear errors-in-variables models with missing data. The method of JEL was also applied to the Bayesian inference by Cheng and Zhao (2019), motivated by Bayesian empirical likelihood (Lazar 2003). When measuring the spread of data using mean absolute deviation, Zhao et al. (2015) have shown the JEL inference for mean absolute deviation. To compare two correlated Gini indices, Alemdjrodo and Zhao (2019) proposed a new method to reduce the computation in the jackknife empirical likelihood. According to Zhao et al. (2018), jackknife empirical likelihood can also be used for both the skewness and kurtosis. The method of adjusted jackknife empirical likelihood (AJEL) was proposed by Zhao et al. (2015) and Chen and Ning (2016). AJEL preserves the property of JEL while providing better coverage probability with slightly longer confidence intervals according to Zhao et al. (2015). Yang and Zhao (2017) applied both JEL and AJEL methods to obtain the quantile difference using smoothed non-parametric estimating equation. In addition, Huang et al. (2024) proposed jackknife empirical likelihood and its variants for the lower mean ratio. Hewage and Sang (2024) proposed jackknife empirical likelihood and weighted JEL confidence intervals for the categorical Gini correlation to improve the small sample performance.

Most recently, Pidgeon et al. (2024) develop jackknife empirical likelihood methods for the correlation coefficient with multiplicative distortion measurement errors, based on the estimator proposed in Zhang et al. (2014). For the new developments in EL and JEL methods, you may read Lazar (2021) and Liu and Zhao (2023).

In this paper, we propose new methods to estimate the correlation coefficient between the response and the predictor under the model of additive distortion. Based on the residual based estimator in Zhang et al. (2017), new interval estimations for the correlation coefficient are developed based on the JEL, mean jackknife empirical likelihood (MJEL) and adjusted jackknife empirical estimating procedures. Furthermore, to further improve the performance in terms of coverage probability for small sample sizes, we consider the confidence intervals using the mean adjusted jackknife empirical likelihood (MAJEL). The results of an extensive simulation study show that the coverages of the new confidence intervals are more accurate than its competitors, especially when the sample size is as small as 25.

The organization of the paper is as follows. In Sect. 2, the confidence interval based on the jackknife empirical likelihood, mean jackknife empirical likelihood and adjusted jackknife empirical likelihood are proposed. We also propose MAJEL to improve the performance for small sample size situations. In Sect. 3, an extensive simulation study is conducted. In Sect. 4, the real data analyses using the new methods are performed for illustrative purpose. A conclusion is made in Sect. 5. The proofs of theorems are provided in the Appendix.

2 Main results

2.1 Review of residual based estimator for correlation coefficients

In this section, we briefly review the residual based estimator for correlation coefficient, which is used by Zhang et al. (2017) to develop empirical likelihood method for correlation coefficients with confounding variables. We use similar notations, which are used in Zhang et al. (2017). The residuals of X and Y are denoted by

$$\begin{cases} \hat{e}_{i\tilde{X}U} = \tilde{X}_i - \hat{E}_h(\tilde{X}_i|U = U_i), \\ \hat{e}_{i\tilde{Y}U} = \tilde{Y}_i - \hat{E}_h(\tilde{Y}_i|U = U_i), \end{cases}$$

where

$$\hat{E}_h(\tilde{X}|U = u) = \frac{n^{-1}\sum_{j=1}^{n} K_h(U_j - u)\tilde{X}_j}{n^{-1}\sum_{j=1}^{n} K_h(U_j - u)},$$
$$\hat{E}_h(\tilde{Y}|U = u) = \frac{n^{-1}\sum_{j=1}^{n} K_h(U_j - u)\tilde{Y}_j}{n^{-1}\sum_{j=1}^{n} K_h(U_j - u)},$$

with kernel function $K_h(\cdot) = h^{-1}K(\cdot/h)$, where $h = \hat{\sigma}_U n^{-1/3}$ and $\hat{\sigma}_U$ is the sample standard deviation of U. Then, the residual based estimator in Zhang et al. (2017) is defined as

$$\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})} = \frac{\widehat{Cov}(e_{\tilde{Y}U}, e_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}} \hat{\sigma}^2_{e_{\tilde{X}U}}}},$$

where

$$\widehat{Cov}(e_{\tilde{Y}U}, e_{\tilde{X}U}) = n^{-1} \sum_{i=1}^{n} \hat{e}_{i\tilde{X}U} \hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{X}U} \bar{\hat{e}}_{\tilde{Y}U},$$

$$\hat{\sigma}^2_{e_{\tilde{X}U}} = n^{-1} \sum_{i=1}^{n} \hat{e}^2_{i\tilde{X}U} - [\bar{\hat{e}}_{\tilde{X}U}]^2,$$

$$\hat{\sigma}^2_{e_{\tilde{Y}U}} = n^{-1} \sum_{i=1}^{n} \hat{e}^2_{i\tilde{Y}U} - [\bar{\hat{e}}_{\tilde{Y}U}]^2$$

with $\bar{\hat{e}}_{\tilde{X}U} = n^{-1} \sum_{i=1}^{n} \hat{e}_{i\tilde{X}U}$ and $\bar{\hat{e}}_{\tilde{Y}U} = n^{-1} \sum_{i=1}^{n} \hat{e}_{i\tilde{Y}U}$.

2.2 JEL for correlation coefficients with additive errors

In this section, we develop JEL methods for correlation coefficients with additive errors. Let $\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})}$ be the residual based estimator obtained as in the previous section and $\hat{\rho}_{i,(e_{\tilde{Y}U}, e_{\tilde{X}U})}$ denote the residual based estimator of $\rho(X, Y)$ calculated with the i^{th} observation deleted, where $i = 1, ..., n$. Let $\hat{V}_i$ denote the jackknife pseudo-value, which is obtained by

$$\hat{V}_i = n\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})} - (n-1)\hat{\rho}_{i,(e_{\tilde{Y}U}, e_{\tilde{X}U})};\ i = 1, ..., n.$$

The jackknife estimator $\hat{\rho}_J(X, Y)$ is defined as

$$\hat{\rho}_J(X, Y) = n^{-1} \sum_{i=1}^{n} \hat{V}_i.$$

The jackknife empirical likelihood ratio at $\rho(X, Y)$ can be then defined as

$$J(\rho(X, Y)) = \sup_{\mathbf{p}=(p_1,\ldots,p_n)} \left(\prod_{i=1}^{n} np_i;\ p_i \geq 0; \sum_{i=1}^{n} p_i = 1; \sum_{i=1}^{n} p_i(\hat{V}_i - \rho(X, Y)) = 0 \right).$$

We can calculate the $-2\log$ of the empirical likelihood ratio as

$$-2\log J(\rho(X, Y)) = 2\sum_{i=1}^{n} \log\{1 + \lambda((\hat{V}_i - \rho(X, Y))\},$$

where λ is the solution of the following equation

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{V}_i-\rho(X,Y)}{1+\lambda(\hat{V}_i-\rho(X,Y))}=0. \tag{2.1}$$

Then, we derive the Wilks' theorem as follows:

Theorem 2.1 *Assume that conditions (C1)–(C5) in the Appendix hold. Let $\rho_0(X,Y)$ be the true value of $\rho(X,Y)$. When $n\to\infty$, we have*

$$-2\log J(\rho_0(X,Y)) \xrightarrow{\mathcal{D}} \chi_1^2.$$

Following the theorem, the asymptotic $100(1-\alpha)\%$ JEL confidence interval for $\rho(X,Y)$ is obtained by

$$I^J_{\rho(X,Y)}=\{\rho_{(X,Y)}:-2\log R(\rho_{(X,Y)})\leq\chi^2_{1-\alpha}(1)\},$$

where $\chi^2_{1-\alpha}(1)$ is the $1-\alpha$ quantile of χ_1^2 and $0<\alpha<1$.

2.3 Adjusted and mean JEL for correlation coefficients with additive errors

Simulation studies have shown that under-coverage issues still exist when the sample size is smaller than 25. Thus, we use adjusted jackknife empirical likelihood to improve the performance of JEL. In order to construct an adjusted jackknife empirical likelihood ratio for $\rho(X,Y)$, first define $\widehat{W}_i$ as

$$\widehat{W}_i(\rho(X,Y))=\hat{V}_i-\rho(X,Y), i=1,...,n, \tag{2.2}$$

and then add one more pseudo value $\widehat{W}_{n+1}$ to $\widehat{W}_i$

$$\widehat{W}_{n+1}(\rho(X,Y))=-\frac{a_n}{n}\sum_{i=1}^{n}\widehat{W}_i(\rho(X,Y)),$$

where we let $a_n=\max(1,\log(n)/2)$ as recommended by Chen et al. (2008). The AJEL is an adjustment to the JEL. Thus, we can calculate the adjusted jackknife estimator as follows by implementing the adjustment to the jackknife estimator

$$\hat{\rho}_A(X,Y)=\hat{\rho}_J(X,Y)+\frac{1}{n+1}\sum_{i=1}^{n+1}\widehat{W}_i(\hat{\rho}_J(X,Y)).$$

The adjusted jackknife empirical likelihood ratio for $\rho(X,Y)$ is defined as

$$J_A(\rho(X,Y))=\sup_{\mathbf{p}=(p_1,\dots,p_{n+1})}\left(\prod_{i=1}^{n+1}(n+1)p_i:p_i\geq 0;\sum_{i=1}^{n+1}p_i=1;\sum_{i=1}^{n+1}p_i\widehat{W}_i(\rho(X,Y))=0\right).$$

Hence, the -2 log of adjusted jackknife empirical likelihood ratio at $\rho(X, Y)$ is

$$-2\log J_A(\rho(X,Y)) = 2\sum_{i=1}^{n+1}\log\{1+\lambda_a\widehat{W}_i\},$$

where λ_a is a solution to the following equation:

$$\frac{1}{n+1}\sum_{i=1}^{n+1}\frac{\widehat{W}_i(\rho(X,Y))}{1+\lambda_a\widehat{W}_i(\rho(X,Y))} = 0.$$

The Wilks' theorem also holds for the adjusted jackknife empirical likelihood and it states as follows:

Theorem 2.2 *Suppose that $\rho_0(X, Y)$ is the true value of $\rho(X, Y)$. Under the same assumptions in Theorem 2.1, as $n \to \infty$, we have*

$$-2\log J_A(\rho_0(X,Y)) \xrightarrow{\mathcal{D}} \chi_1^2.$$

Thus, following the theorem, the asymptotic $100(1-\alpha)\%$ AJEL confidence interval is defined as:

$$I^A_{\rho(X,Y)} = \{\rho_a(X,Y) : -2\log J_A(\rho_a(X,Y)) \le \chi^2_{1-\alpha}(1)\},$$

where $0 < \alpha < 1$.

By using AJEL, the length of AJEL confidence interval is usually longer than JEL and thus the coverage probability of AJEL is better in small sample cases.

To apply the method of mean jackknife empirical likelihood, we let M denote the pseudo vector calculated from $\hat{W}_i$ in Eq. (2.2) such that:

$$M = \left\{\frac{\hat{W}_i + \hat{W}_j}{2} : 1 \le i \le j \le n\right\}. \tag{2.3}$$

Through the above equation, the original $\hat{W}_i$ is expanded into a vector of size $N = n(n+1)/2$. Meanwhile, M maintains the same mean as $\hat{W}_i$. The expected value of the new M is close to 0. Similar to the adjusted jackknife estimator, the mean jackknife estimator can be defined as follows by adding an adjustment term to the jackknife estimator:

$$\hat{\rho}_M(X,Y) = \hat{\rho}_J(X,Y) + \frac{1}{N}\sum_{i=1}^{N} M_i(\hat{\rho}_J(X,Y)).$$

Then, we can construct the empirical likelihood based on the new vector M. The mean empirical likelihood ratio, denoted as $J_M(\rho(X, Y))$, is defined as:

$$J_M(\rho(X,Y)) = \max_{\mathbf{p}=(p_1,\ldots,p_n)} \left(\prod_{i=1}^{N} Np_i;\ p_i \geq 0; \sum_{i=1}^{N} p_i = 1; \sum_{i=1}^{N} p_i M_i(\rho(X,Y)) = 0 \right).$$

By the properties of empirical likelihood, the log-likelihood $l^M(\rho(X, Y))$ can then be calculated as:

$$\begin{aligned} l^M(\rho(X,Y)) &= \frac{-2\log J_M(\rho(X,Y))}{n+1} \\ &= \frac{2}{n+1}\sum_{i=1}^{N}\log(1+\lambda_m M_i(\rho(X,Y))), \end{aligned}$$

where λ_m is the solution of the following equation

$$N^{-1}\sum_{i=1}^{N}\frac{M_i(\rho(X,Y))}{1+\lambda_m M_i(\rho(X,Y))} = 0.$$

To construct the confidence interval of $\rho(X, Y)$, we obtain Wilks' theorem as Pidgeon et al. (2024) did:

Theorem 2.3 *Assume that the same conditions as we did in Theorem 2.1. We have that*

$$l^M(\rho_0(X,Y)) \xrightarrow{\mathcal{D}} \chi_1^2.$$

Then, the mean jackknife empirical likelihood confidence interval is defined as follows:

$$I^M_{\rho(X,Y)} = \{\rho(X,Y) : l^M\ (\rho(X,Y)) \leq \chi^2_{1-\alpha}(1)\},$$

where $0 < \alpha < 1$.

Like the jackknife empirical likelihood, there is no exact formula for the confidence interval of mean jackknife empirical likelihood. Thus, the calculation of CI for MJEL is the same as that for EL by fitting a vector of estimators into the non-parametric models and compare the $l^M(\rho(X, Y))$ to the desired quantile. In general cases, the length of confidence interval decreases with the larger sample size. Simulation studies have shown that the confidence interval of MJEL is longer than that of empirical likelihood.

2.4 Mean adjusted JEL for correlation coefficients with additive errors

To increase the performance in small sample situations, we combine the methods of MJEL and AJEL and propose the methods of mean adjusted jackknife empirical likelihood (MAJEL). For MAJEL, we first obtain the vector $\widehat{W}_i$ from Eq. (2.2) and then expand the vector using the equation similar to Eq. (2.3). The M^a for MAJEL is computed as follows:

$$M^a = \left\{ \frac{\widehat{W}_i + \widehat{W}_j}{2} : 1 \le i \le j \le n+1 \right\}.$$

The expectation of M^a remains close to 0 and M^a has $N^a = (n+1)(n+2)/2$ values. The mean adjusted jackknife estimator is then defined as

$$\hat{\rho}_{MA}(X, Y) = \hat{\rho}_A(X, Y) + \frac{1}{N^a} \sum_{i=1}^{N^a} M^a(\hat{\rho}_A(X, Y)).$$

The mean adjusted jackknife empirical likelihood ratio is then defined as follows:

$$J_{MA}(\rho(X, Y)) = \sup \left(\prod_{i=1}^{N^a} N^a p_i; \, p_i \ge 0; \sum_{i=1}^{N^a} p_i = 1; \sum_{i=1}^{N^a} p_i M_i^a = 0 \right).$$

The log-likelihood of MAJEL can be calculated by the following equation:

$$l^{MA}(\rho(X, Y)) = \frac{2}{n+2} \sum_{i=1}^{N^a} \log(1 + \lambda_{ma} M_i^a(\rho(X, Y))),$$

where λ_{ma} is the solution of the following equation:

$$\frac{1}{N^a} \sum_{i=1}^{N^a} \frac{M_i^a(\rho(X, Y))}{1 + \lambda_{ma} M_i^a(\rho(X, Y))} = 0.$$

We can also obtain the Wilks' theorem for MAJEL like Huang et al. (2024).

Theorem 2.4 *Under the same assumptions in Theorem 2.1, as $n \to \infty$, we have*

$$l^{MA}(\rho_0(X, Y)) \xrightarrow{\mathcal{D}} \chi_1^2.$$

From the theorem, we can construct the asymptotic $100(1-\alpha)\%$ MAJEL confidence interval as follows:

$$I_{\rho(X,Y)}^{MA} = \{\rho(X, Y) : l^{MA}(\rho(X, Y)) \le \chi_{1-\alpha}^2(1)\},$$

where $0 < \alpha < 1$.

MAJEL performs better than JEL, MJEL and AJEL when the sample size is small. The average length of the MAJEL confidence interval is longer than AJEL and MJEL.

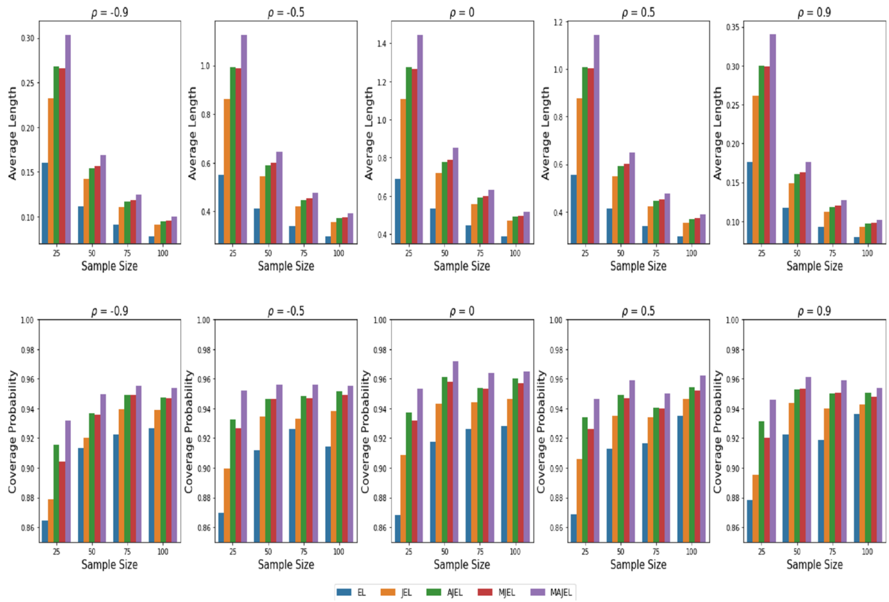

Fig. 1 Average length and coverage probability trend plot

3 Simulation study

For the simulation study, (X, Y) is generated by multivariate normal distribution with $\mu = (2, 4)$ and $\rho(X, Y) = -0.9, -0.5, 0, 0.5, 0.9$. To ensure X and Y are generated with predefined correlation coefficient, we let the $(1, 1)$ and $(2, 2)$ of the covariance matrix to be 1 and the $(1, 2)$ and $(2, 1)$ elements equal to the predefined correlation coefficients. U is simulated with uniform, normal, Beta and Weibull distributions. The uniform distribution of U, $U(a, b)$, is generated with $a = 1$ and $b = 7$ while $\psi(U) = U - 4$ and $\phi(U) = 4 - U$. The normal distribution of $\mu = 2$ and $\sigma = 1$ is used to generate U. We set $\psi(U) = U - 2$ and $\phi(U) = 2 - U$. In the Beta distribution, $Beta(\alpha, \beta)$, we let $\alpha = 2$ and $\beta = 8$ such that $\psi(U) = U - 0.2$ and $\phi(U) = 0.2 - U$. For the Weibull distribution, $W(\lambda, \kappa)$, we have $\lambda = 1$ and $\kappa = 1.2$. We also let $\psi(U) = U - 0.9407$ and $\phi(U) = 0.9407 - U$ to ensure $E[\psi(U)] = 0$ and $E[\phi(U)] = 0$. The observed values $(\tilde{X}, \tilde{Y})$ are set up as:

$$\begin{cases} \tilde{X} = X + \psi(U), \\ \tilde{Y} = Y + \phi(U). \end{cases}$$

We have also simulated the non-linear scenario under uniform distribution, $U(0, 1)$, such that $\psi(U) = 1/3 - U^2$ and $\phi(U) = U^3 - 1/4$.

Each simulation was repeated 2,000 times with the sample size $n = 25, 50, 75, 100$. For the kernel function, we choose to use the Epanechnikov kernel functions, $K(t) = 0.75(1 - t^2)^+$, as suggested by Zhang et al. (2017). The bandwidth is chosen as suggested by Silverman (1986) such that $h = \hat{\sigma}_U n^{-1/3}$, where $\hat{\sigma}_U$ is the sample standard deviation of U. Five methods, EL, JEL, AJEL, MJEL, and MAJEL are compared in terms of estimators, coverage probability (CP) and average lengths (AL) of 95% confidence intervals. The results are shown in the following Tables 1, 2, 3, 4, 5.

Figure 1 shows trends of average length and coverage probability from the simulation under Weibull distribution. The trends are similar across all scenarios with different $\rho(X, Y)$, but actual values vary. If $\rho(X, Y)$ is closer to 0, average length tends to be longer.

Conclusions from the simulations are as follows:

(1) The average lengths of confidence intervals from all five jackknife methods are longer than those from EL method. The average length of the AJEL confidence interval is longer than that of JEL but the lengths of MJEL and AJEL confidence intervals are close. The average length of the MJEL confidence interval is shorter than that of AJEL when the sample size is as small as 25. When the sample size is greater than 50, MJEL has longer confidence interval length than AJEL.
(2) All new methods give better performance with an increase in the sample size.
(3) Both the MJEL and AJEL have similar performances. The MAJEL show the improvement from MJEL and AJEL.
(4) The performance of estimators varies under the same sample size. However, with larger sample sizes, JEL gives better estimators than the EL method.
(5) The estimator, coverage probability and average length are consistent under the normal, Beta and Weibull distributions. The methods are consistent regardless of symmetric or asymmetric distributions.
(6) The simulation study from nonlinear scenario shows consistent performance for all proposed methods.
(7) The MAJEL outperforms all the other methods in the small sample size ($n = 25$) situations. When the sample sizes are larger than 50, the MAJEL has over-coverage issues, which is similar to what we have observed in the MJEL method.

4 Real data applications

To compare the new methods to the original EL method proposed by Zhang et al. (2017), we use the 1993 new car data and Boston house price data to conduct the real data analysis. The 1993 new car data is collected by Lock (1993). The data has 93 observations with 27 variables. We choose the horsepower as $\tilde{X}$ and the highway MPG as $\tilde{Y}$. Weights of cars are considered as the confounding distortion errors U. The Boston house data is retrieved from Harrison and Rubinfeld (1978) and is also used to demonstrate EL method by Zhang et al. (2017). The data contains 506 observations with 14 variables. We study the correlation coefficient between the house prices ($medv$) and distance to employment centers (dis), where dis is considered as $\tilde{X}$ and $medv$ is considered as $\tilde{Y}$. The lower status of the population ($lstat$) is considered as the confounding variable U. In the real data analysis, to calculate the bandwidth h for the kernel function $K_h(\cdot) = h^{-1}K(\cdot/h)$, we let $h = \hat{\sigma}_U n^{-1/3}$ where $\hat{\sigma}_U$ is the sample standard deviation of confounding variable. For the first part of the real data analysis, we compare the EL, JEL, AJEL, MJEL, and MAJEL methods using the whole dataset.

Table 1 Comparison of all methods under the uniform distribution

$\rho(X,Y)$	n	EL			JEL			AJEL			MJEL			MAJEL		
		$\hat{\rho}(X,Y)$	AL	CP	$\hat{\rho}_J(X,Y)$	AL	CP	$\hat{\rho}_A(X,Y)$	AL	CP	$\hat{\rho}_M(X,Y)$	AL	CP	$\hat{\rho}_{MA}(X,Y)$	AL	CP
– 0.9	25	– 0.895	0.154	0.845	– 0.900	0.240	0.885	– 0.900	0.276	0.901	– 0.900	0.274	0.907	– 0.900	0.294	**0.920**
	50	– 0.898	0.109	0.907	– 0.900	0.145	0.935	– 0.900	0.156	0.941	– 0.900	0.159	0.944	– 0.900	0.166	**0.955**
	75	– 0.899	0.088	0.918	– 0.900	0.111	**0.947**	– 0.900	0.117	0.953	– 0.900	0.119	0.956	– 0.900	0.122	0.960
	100	– 0.900	0.076	0.928	– 0.900	0.091	0.944	– 0.900	0.095	0.946	– 0.900	0.096	**0.949**	– 0.900	0.098	0.951
– 0.5	25	– 0.492	0.533	0.850	– 0.499	0.889	0.893	– 0.499	1.023	0.913	– 0.499	1.016	0.916	– 0.499	1.090	**0.933**
	50	– 0.499	0.398	0.905	– 0.501	0.553	0.935	– 0.501	0.598	0.943	– 0.501	0.607	0.945	– 0.501	0.633	**0.953**
	75	– 0.499	0.331	0.917	– 0.501	0.425	0.938	– 0.501	0.449	0.946	– 0.501	0.455	**0.950**	– 0.501	0.469	0.956
	100	– 0.498	0.290	0.918	– 0.499	0.355	0.947	– 0.499	0.371	**0.950**	– 0.499	0.374	0.954	– 0.499	0.383	0.959
0	25	– 0.016	0.675	0.863	0.001	1.151	0.907	0.001	1.325	0.925	0.001	1.314	0.930	0.001	1.410	**0.947**
	50	– 0.003	0.518	0.903	0.006	0.731	0.945	0.006	0.789	**0.951**	0.006	0.802	0.954	0.006	0.836	0.964
	75	– 0.001	0.434	0.912	0.003	0.564	0.944	0.003	0.596	**0.952**	0.003	0.604	0.955	0.003	0.623	0.961
	100	– 0.003	0.381	0.919	– 0.001	0.470	0.946	– 0.001	0.491	**0.952**	– 0.001	0.495	0.954	– 0.001	0.507	0.959
0.5	25	0.474	0.545	0.856	0.511	0.914	0.902	0.511	1.052	0.914	0.511	1.045	0.923	0.511	1.121	**0.933**
	50	0.484	0.404	0.907	0.500	0.561	0.937	0.500	0.606	**0.948**	0.500	0.615	0.954	0.500	0.641	0.959
	75	0.494	0.332	0.910	0.503	0.426	0.937	0.503	0.450	0.942	0.503	0.456	0.947	0.503	0.470	**0.953**
	100	0.496	0.290	0.933	0.502	0.356	**0.953**	0.502	0.371	0.956	0.502	0.375	0.958	0.502	0.384	0.963
0.9	25	0.868	0.192	0.842	0.909	0.300	0.917	0.909	0.345	0.929	0.909	0.341	0.928	0.909	0.367	**0.943**
	50	0.886	0.119	0.892	0.903	0.157	0.929	0.903	0.170	0.935	0.903	0.172	0.939	0.903	0.179	**0.946**
	75	0.892	0.093	0.905	0.902	0.116	0.935	0.902	0.122	0.938	0.902	0.124	0.942	0.902	0.128	**0.946**
	100	0.895	0.079	0.926	0.902	0.094	0.935	0.902	0.099	0.939	0.902	0.099	0.942	0.902	0.102	**0.946**

The bold values demonstrate the coverage probability which is the closest to the confidence level 0.95 among all the methods. From the number of values in bold font, we find that the proposed four JEL methods outperform the existing EL method in terms of coverage probability

Table 2 Comparison of all methods under the normal distribution

$\rho(X,Y)$	n	EL			JEL			AJEL			MJEL			MAJEL		
		$\hat{\rho}(X,Y)$	AL	CP	$\hat{\rho}_J(X,Y)$	AL	CP	$\hat{\rho}_A(X,Y)$	AL	CP	$\hat{\rho}_M(X,Y)$	AL	CP	$\hat{\rho}_{MA}(X,Y)$	AL	CP
− 0.9	25	− 0.895	0.162	0.854	− 0.903	0.255	0.893	− 0.903	0.294	0.909	− 0.903	0.292	0.913	− 0.903	0.314	**0.928**
	50	− 0.898	0.112	0.910	− 0.900	0.151	0.935	− 0.900	0.163	0.943	− 0.900	0.166	0.948	− 0.900	0.173	**0.951**
	75	− 0.899	0.090	0.917	− 0.900	0.113	0.938	− 0.900	0.120	0.941	− 0.900	0.122	0.944	− 0.900	0.125	**0.948**
	100	− 0.898	0.078	0.916	− 0.899	0.095	0.942	− 0.899	0.099	0.947	− 0.899	0.100	**0.951**	− 0.899	0.102	0.956
− 0.5	25	− 0.485	0.551	0.835	− 0.496	0.935	0.895	− 0.496	1.076	0.911	− 0.496	1.071	0.917	− 0.496	1.150	**0.929**
	50	− 0.494	0.408	0.894	− 0.499	0.576	0.926	− 0.499	0.622	0.934	− 0.499	0.634	0.938	− 0.499	0.661	**0.944**
	75	− 0.5	0.337	0.909	− 0.504	0.439	0.942	− 0.504	0.464	**0.950**	− 0.504	0.471	0.955	− 0.504	0.485	0.965
	100	− 0.496	0.295	0.918	− 0.498	0.368	0.947	− 0.498	0.384	**0.951**	− 0.498	0.388	0.954	− 0.498	0.398	0.960
0	25	− 0.005	0.695	0.841	0.002	1.215	0.904	0.002	1.398	0.926	0.002	1.392	0.930	0.002	1.493	**0.945**
	50	− 0.002	0.529	0.886	0.001	0.758	0.928	0.001	0.819	0.937	0.001	0.834	0.942	0.001	0.870	**0.951**
	75	0.003	0.444	0.918	0.004	0.586	**0.953**	0.004	0.619	0.960	0.004	0.630	0.965	0.004	0.649	0.970
	100	0.003	0.388	0.930	0.004	0.488	**0.957**	0.004	0.510	0.958	0.004	0.515	0.959	0.004	0.528	0.963
0.5	25	0.481	0.551	0.864	0.500	0.928	0.906	0.500	1.068	0.927	0.500	1.061	0.928	0.500	1.139	**0.944**
	50	0.490	0.410	0.905	0.499	0.578	0.936	0.499	0.624	0.944	0.499	0.635	**0.950**	0.499	0.661	0.959
	75	0.490	0.341	0.915	0.497	0.443	0.941	0.497	0.468	0.947	0.497	0.475	**0.952**	0.497	0.490	0.960
	100	0.496	0.295	0.924	0.500	0.367	0.946	0.500	0.383	**0.952**	0.500	0.387	0.953	0.500	0.396	0.957
0.9	25	0.884	0.177	0.856	0.904	0.280	0.894	0.904	0.322	0.912	0.904	0.320	0.911	0.904	0.344	**0.927**
	50	0.894	0.114	0.886	0.901	0.154	0.915	0.901	0.167	0.926	0.901	0.169	0.930	0.901	0.177	**0.939**
	75	0.895	0.093	0.906	0.899	0.118	0.936	0.899	0.124	0.941	0.899	0.126	0.942	0.899	0.130	**0.950**
	100	0.899	0.078	0.929	0.901	0.094	0.945	0.901	0.098	**0.950**	0.901	0.099	0.951	0.901	0.102	0.954

The bold values demonstrate the coverage probability which is the closest to the confidence level 0.95 among all the methods. From the number of values in bold font, we find that the proposed four JEL methods outperform the existing EL method in terms of coverage probability

Table 3 Comparison of all methods under the Beta distribution

$\rho(X,Y)$	n	EL			JEL			AJEL			MJEL			MAJEL		
		$\hat{\rho}(X,Y)$	AL	CP	$\hat{\rho}_J(X,Y)$	AL	CP	$\hat{\rho}_A(X,Y)$	AL	CP	$\hat{\rho}_M(X,Y)$	AL	CP	$\hat{\rho}_{MA}(X,Y)$	AL	CP
− 0.9	25	− 0.894	0.161	0.857	− 0.902	0.248	0.891	− 0.902	0.285	0.905	− 0.902	0.284	0.910	− 0.902	0.304	**0.924**
	50	− 0.898	0.110	0.912	− 0.902	0.148	0.924	− 0.902	0.160	0.936	− 0.902	0.163	0.941	− 0.902	0.169	**0.948**
	75	− 0.899	0.089	0.925	− 0.901	0.112	0.945	− 0.901	0.118	0.947	− 0.901	0.120	**0.950**	− 0.901	0.124	0.956
	100	− 0.899	0.077	0.916	− 0.900	0.093	0.940	− 0.900	0.097	0.946	− 0.900	0.098	0.949	− 0.900	0.101	**0.951**
− 0.5	25	− 0.494	0.544	0.863	− 0.511	0.908	0.907	− 0.511	1.044	0.923	− 0.511	1.039	0.926	− 0.511	1.114	**0.939**
	50	− 0.499	0.407	0.904	− 0.505	0.566	0.943	− 0.505	0.612	**0.950**	− 0.505	0.623	0.954	− 0.505	0.649	0.960
	75	− 0.498	0.338	0.913	− 0.503	0.435	0.937	− 0.503	0.459	0.945	− 0.503	0.466	**0.950**	− 0.503	0.481	0.955
	100	− 0.499	0.293	0.909	− 0.501	0.360	0.940	− 0.501	0.376	0.943	− 0.501	0.380	0.946	− 0.501	0.389	**0.948**
0	25	− 0.001	0.685	0.834	0.003	1.172	0.891	0.003	1.349	0.910	0.003	1.341	0.915	0.003	1.439	**0.931**
	50	0.002	0.528	0.888	0.000	0.745	0.929	0.000	0.805	0.938	0.000	0.819	0.946	0.000	0.854	**0.952**
	75	0.002	0.442	0.915	0.002	0.577	0.940	0.002	0.609	**0.948**	0.002	0.619	0.952	0.002	0.638	0.957
	100	− 0.003	0.386	0.918	− 0.003	0.480	0.946	− 0.003	0.502	**0.950**	− 0.003	0.507	0.953	− 0.003	0.519	0.956
0.5	25	0.493	0.544	0.857	0.507	0.908	0.900	0.507	1.044	0.919	0.507	1.039	0.925	0.507	1.115	**0.941**
	50	0.497	0.404	0.898	0.503	0.563	0.924	0.503	0.608	0.930	0.503	0.619	0.936	0.503	0.645	**0.944**
	75	0.497	0.337	0.912	0.501	0.434	0.945	0.501	0.459	**0.951**	0.501	0.466	0.955	0.501	0.480	0.961
	100	0.497	0.294	0.915	0.500	0.363	0.940	0.500	0.379	0.943	0.500	0.383	0.945	0.500	0.392	**0.949**
0.9	25	0.895	0.159	0.848	0.901	0.246	0.877	0.901	0.283	0.896	0.901	0.282	0.902	0.901	0.302	**0.923**
	50	0.899	0.110	0.885	0.902	0.146	0.915	0.902	0.158	0.922	0.902	0.161	0.928	0.902	0.167	**0.932**
	75	0.900	0.089	0.929	0.901	0.111	0.941	0.901	0.118	0.946	0.901	0.119	**0.947**	0.901	0.123	0.954
	100	0.899	0.077	0.916	0.900	0.093	0.944	0.900	0.098	0.946	0.900	0.099	**0.948**	0.900	0.101	0.953

The bold values demonstrate the coverage probability which is the closest to the confidence level 0.95 among all the methods. From the number of values in bold font, we find that the proposed four JEL methods outperform the existing EL method in terms of coverage probability

Table 4 Comparison of all methods under the Weibull distribution

$\rho(X,Y)$	n	EL			JEL			AJEL			MJEL			MAJEL		
		$\hat{\rho}(X,Y)$	AL	CP	$\hat{\rho}_J(X,Y)$	AL	CP	$\hat{\rho}_A(X,Y)$	AL	CP	$\hat{\rho}_M(X,Y)$	AL	CP	$\hat{\rho}_{MA}(X,Y)$	AL	CP
− 0.9	25	− 0.895	0.160	0.865	− 0.902	0.233	0.882	− 0.902	0.268	0.903	− 0.902	0.266	0.909	− 0.902	0.286	**0.923**
	50	− 0.898	0.112	0.914	− 0.900	0.142	0.923	− 0.900	0.154	0.930	− 0.900	0.157	0.936	− 0.900	0.163	**0.943**
	75	− 0.899	0.091	0.923	− 0.900	0.110	0.937	− 0.900	0.116	0.943	− 0.900	0.118	0.946	− 0.900	0.122	**0.953**
	100	− 0.900	0.077	0.927	− 0.901	0.091	0.938	− 0.901	0.095	0.941	− 0.901	0.096	0.944	− 0.901	0.098	**0.951**
− 0.5	25	− 0.493	0.550	0.870	− 0.503	0.862	0.906	− 0.503	0.991	0.923	− 0.503	0.987	0.927	− 0.503	1.059	**0.943**
	50	− 0.497	0.410	0.912	− 0.500	0.545	0.933	− 0.500	0.588	0.942	− 0.500	0.599	0.947	− 0.500	0.625	**0.953**
	75	− 0.499	0.339	0.926	− 0.501	0.421	0.937	− 0.501	0.445	0.944	− 0.501	0.451	**0.949**	− 0.501	0.465	0.956
	100	− 0.497	0.297	0.915	− 0.500	0.356	0.944	− 0.500	0.371	0.946	− 0.500	0.375	**0.949**	− 0.500	0.384	0.953
0	25	− 0.006	0.691	0.868	0.000	1.110	0.908	0.000	1.277	0.925	0.000	1.270	0.931	0.000	1.362	**0.943**
	50	− 0.001	0.532	0.918	0.003	0.718	0.943	0.003	0.775	**0.956**	0.003	0.789	0.963	0.003	0.823	0.969
	75	0.001	0.444	0.926	0.004	0.559	0.945	0.004	0.590	**0.951**	0.004	0.599	0.955	0.004	0.618	0.962
	100	− 0.001	0.389	0.928	0.001	0.469	**0.949**	0.001	0.490	0.953	0.001	0.495	0.955	0.001	0.507	0.959
0.5	25	0.483	0.556	0.869	0.510	0.878	0.903	0.510	1.011	0.919	0.510	1.007	0.923	0.510	1.080	**0.939**
	50	0.490	0.412	0.913	0.500	0.547	0.941	0.500	0.591	0.946	0.500	0.601	**0.949**	0.500	0.626	0.956
	75	0.496	0.340	0.917	0.502	0.422	0.933	0.502	0.446	0.938	0.502	0.452	0.944	0.502	0.466	**0.949**
	100	0.497	0.296	0.935	0.502	0.354	0.944	0.502	0.370	**0.950**	0.502	0.373	0.952	0.502	0.382	0.958
0.9	25	0.884	0.176	0.878	0.906	0.259	0.891	0.906	0.298	0.908	0.906	0.296	0.912	0.906	0.318	**0.925**
	50	0.892	0.117	0.923	0.900	0.148	0.939	0.900	0.160	0.943	0.900	0.163	**0.948**	0.900	0.170	0.954
	75	0.895	0.093	0.919	0.900	0.112	0.947	0.900	0.119	**0.950**	0.900	0.120	0.954	0.900	0.124	0.960
	100	0.897	0.079	0.936	0.901	0.093	0.945	0.901	0.097	0.947	0.901	0.098	**0.951**	0.901	0.100	0.953

The bold values demonstrate the coverage probability which is the closest to the confidence level 0.95 among all the methods. From the number of values in bold font, we find that the proposed four JEL methods outperform the existing EL method in terms of coverage probability

Table 5 Comparison of all methods in nonlinear scenario under the uniform distribution

$\rho(X,Y)$	n	EL			JEL			AJEL			MJEL			MAJEL		
		$\hat{\rho}(X,Y)$	AL	CP	$\hat{\rho}_J(X,Y)$	AL	CP	$\hat{\rho}_A(X,Y)$	AL	CP	$\hat{\rho}_M(X,Y)$	AL	CP	$\hat{\rho}_{MA}(X,Y)$	AL	CP
− 0.9	25	− 0.894	0.156	0.848	− 0.901	0.244	0.891	− 0.901	0.281	0.904	− 0.901	0.278	0.910	− 0.901	0.299	**0.924**
	50	− 0.897	0.109	0.906	− 0.900	0.146	0.934	− 0.900	0.157	0.943	− 0.900	0.160	**0.946**	− 0.900	0.167	0.955
	75	− 0.898	0.089	0.916	− 0.900	0.111	**0.948**	− 0.900	0.118	0.953	− 0.900	0.119	0.956	− 0.900	0.123	0.961
	100	− 0.899	0.076	0.927	− 0.900	0.091	0.944	− 0.900	0.095	0.946	− 0.900	0.096	0.948	− 0.900	0.098	**0.952**
− 0.5	25	− 0.486	0.537	0.848	− 0.500	0.900	0.901	− 0.500	1.035	0.916	− 0.500	1.029	0.920	− 0.500	1.104	**0.933**
	50	− 0.496	0.400	0.903	− 0.501	0.557	0.935	− 0.501	0.601	0.943	− 0.501	0.611	**0.947**	− 0.501	0.637	0.954
	75	− 0.497	0.331	0.915	− 0.501	0.426	0.938	− 0.501	0.450	0.946	− 0.501	0.457	**0.950**	− 0.501	0.470	0.956
	100	− 0.497	0.290	0.918	− 0.499	0.356	0.947	− 0.499	0.372	**0.951**	− 0.499	0.375	0.953	− 0.499	0.384	0.957
0	25	− 0.002	0.675	0.856	− 0.001	1.158	0.911	− 0.001	1.333	0.930	− 0.001	1.323	0.931	− 0.001	1.419	**0.946**
	50	0.002	0.518	0.907	0.004	0.732	0.943	0.004	0.791	**0.952**	0.004	0.803	0.957	0.004	0.837	0.966
	75	0.002	0.434	0.910	0.002	0.564	0.947	0.002	0.596	**0.953**	0.002	0.605	0.957	0.002	0.623	0.961
	100	− 0.001	0.381	0.918	− 0.002	0.470	**0.949**	− 0.002	0.491	0.952	− 0.002	0.496	0.956	− 0.002	0.508	0.959
0.5	25	0.492	0.534	0.852	0.506	0.900	0.901	0.506	1.035	0.918	0.506	1.029	0.923	0.506	1.104	**0.936**
	50	0.492	0.400	0.903	0.498	0.557	0.941	0.498	0.602	0.946	0.498	0.612	**0.953**	0.498	0.637	0.959
	75	0.499	0.330	0.908	0.502	0.424	0.936	0.502	0.448	0.946	0.502	0.454	**0.948**	0.502	0.468	0.955
	100	0.499	0.289	0.934	0.501	0.355	**0.952**	0.501	0.370	0.959	0.501	0.374	0.961	0.501	0.383	0.965
0.9	25	0.894	0.158	0.860	0.903	0.250	0.889	0.903	0.287	0.906	0.903	0.285	0.912	0.903	0.306	**0.926**
	50	0.897	0.109	0.905	0.900	0.145	0.932	0.900	0.156	0.940	0.900	0.159	0.942	0.900	0.166	**0.950**
	75	0.898	0.088	0.907	0.900	0.110	0.935	0.900	0.116	0.939	0.900	0.118	0.945	0.900	0.122	**0.952**
	100	0.899	0.076	0.918	0.901	0.091	0.934	0.901	0.095	0.941	0.901	0.096	0.944	0.901	0.098	**0.949**

The bold values demonstrate the coverage probability which is the closest to the confidence level 0.95 among all the methods. From the number of values in bold font, we find that the proposed four JEL methods outperform the existing EL method in terms of coverage probability

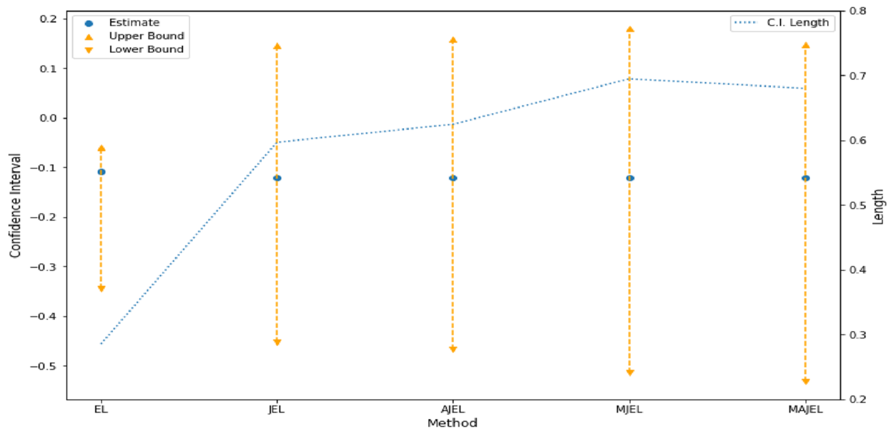

Fig. 2 Confidence intervals for the new car dataset

Table 6 1993 new car data analysis

Method	Estimator	Lower bound (LB)	Upper bound (UB)	Length
EL	− 0.1083	− 0.3451	− 0.0599	0.2852
JEL	− 0.1220	− 0.4518	0.1450	0.5968
AJEL	− 0.1220	− 0.4671	0.1574	0.6245
MJEL	− 0.1220	− 0.5145	0.1803	0.6948
MAJEL	− 0.1220	− 0.5320	0.1481	0.6801

Figure 2 concludes the analysis result for the new car dataset. The vertical lines show the confidence intervals for all methods including EL. The blue line shows the change of confidence interval length. The results shown in Table 6 and Fig. 2 are consistent with simulation studies that the 95% confidence intervals of new methods are longer than those of EL. The MAJEL confidence interval is longer than AJEL confidence interval. The naive correlation coefficient between horsepower and highway MPG from the 1993 cars data, $\rho(\tilde{X}, \tilde{Y})$, is -0.8107, which indicates that there exists a strong negative correlation between MPG and horsepower. After taking the confounding variable of weight into consideration, the 95% confidence intervals of all proposed methods contain zero, meaning the horsepower and MPG are uncorrelated (Table 6). However, the EL confidence interval proposed by Zhang et al. (2017) does not include zero.

For the Boston housing price analysis, the naive correlation coefficient, $\rho(\tilde{X}, \tilde{Y})$, between the distance and median price is 0.2499. However, the new methods indicate a negative correlation between these two variables. Also, all confidence intervals are exclusively less than zero, meaning the distance to employment centers and house prices are negatively correlated (Table 7). The next part of the real data analysis focuses on the fact that new methods outperform EL with the small sample size. Thus, the Boston data set is partitioned into five sets depending on the lower status of the population. The partition and the results are as shown in Table 8. A plot is drawn in

Table 7 Comparison of Boston house price analysis

Method	Estimator	Lower bound (LB)	Upper bound (UB)	Length
EL	− 0.2522	− 0.3420	− 0.1549	0.1871
JEL	− 0.2478	− 0.3524	− 0.1495	0.2029
AJEL	− 0.2478	− 0.3536	− 0.1484	0.2052
MJEL	− 0.2478	− 0.3531	− 0.1492	0.2039
MAJEL	− 0.2478	− 0.3543	− 0.1481	0.2062

Fig. 3 to demonstrate how all the methods perform in the partitioned Boston house price analysis.

The results show that the correlation coefficient between the distance and house price increases with the increase of the lower status of population (lstat). House price is positively correlated with the distance to employment centers when lstat is greater than 15 and the correlation is moderate. Also, when the lower status of population (lstat) is less than 5, house prices can be moderately correlated to the distance in a negative fashion.

5 Conclusions

In this paper, we proposed four JEL methods for a correlation analysis when the response variable is influenced by a confounding variable and the error terms are assumed to be additions to the unobserved true values of interest. By the nature of JEL, MJEL and AJEL, all confidence intervals are larger than those of empirical likelihood by Zhang et al. (2017), thus providing better coverage than the EL. Both the AJEL and MJEL have longer confidence intervals than JEL. Performance of these methods vary depending on $\rho(X, Y)$ and n. The underlying theoretical cause for such variety is challenging to prove and still under investigation. The AJEL provides longer confidence intervals compared to the MJEL, when sample sizes are as small as 25. When the sample size, n, is between 50 and 100, the length of MJEL is larger than that of AJEL. Both the MJEL and AJEL have larger length than the JEL. When the true value of $\rho(X, Y) = 0$, all proposed methods generate longer confidence intervals compared with other situations. In the cases with $25 \leq n \leq 100$, all new methods provide better coverage probability compared to the conventional empirical likelihood. But the performance varies case by case. Both the MJEL and AJEL could have over-coverage when the sample size is greater than or equal to 75. Generally, the MAJEL shows better performances and is recommended when $n = 50$ and $n = 25$. In cases of $n = 75$ and $n = 100$, the coverage probabilities among AJEL, MJEL, and MAJEL vary depending on the underlying distribution of U and functions of $\psi(U)$, $\phi(U)$ yet the general performances of all three are similar. When the sample size is between 50 and 100, the MAJEL provides marginally better performance while consumes more computation resources. Thus, the recommendation in such scenarios depends on the actual case to be applied and available computation resources. When applied to real

Table 8 Partitioned Boston house price analysis

lstat n	(0,5] 62	(5,10] 157	(10,15] 125	(15,20] 88	(20,100] 74
$\hat{\rho}$(LB, UB)	− 0.465 (− 0.630, − 0.242)	− 0.387 (− 0.506, − 0.236)	− 0.180 (− 0.301, − 0.033)	0.360 (0.174, 0.524)	0.437 (0.290, 0.565)
$\hat{\rho}_J$(LB, UB)	− 0.493 (− 0.717, − 0.227)	− 0.391 (− 0.542, − 0.228)	− 0.197 (− 0.339, − 0.033)	0.374 (0.082, 0.636)	0.460 (0.292, 0.647)
$\hat{\rho}_A$(LB, UB)	− 0.493 (− 0.732, − 0.210)	− 0.391 (− 0.553, − 0.223)	− 0.197 (− 0.344, − 0.027)	0.374 (0.068, 0.649)	0.460 (0.283, 0.658)
$\hat{\rho}_M$(LB, UB)	− 0.493 (− 0.723, − 0.200)	− 0.391 (− 0.557, − 0.223)	− 0.197 (− 0.341, − 0.019)	0.374 (0.029, 0.675)	0.460 (0.284, 0.657)
$\hat{\rho}_{MA}$(LB, UB)	− 0.493 (− 0.738, − 0.181)	− 0.391 (− 0.562, − 0.218)	− 0.197 (− 0.346, − 0.013)	0.374 (0.013, 0.689)	0.460 (0.274, 0.668)

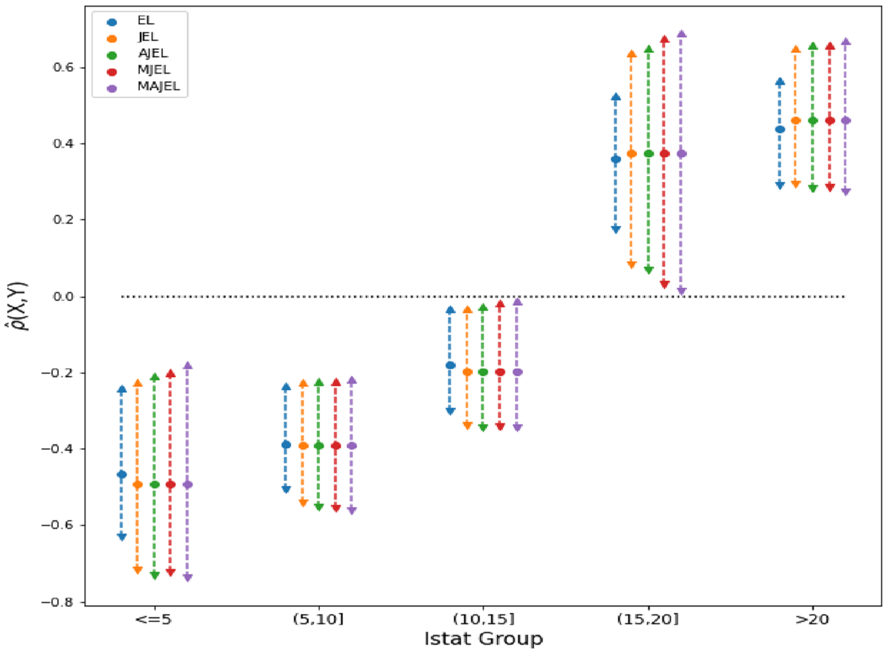

Fig. 3 Confidence intervals for the partitioned house price dataset

data sets, the new methods make it more convenient to partition a data set into smaller subgroups without losing efficacy such that the analysis of trend is possible even when we are dealing with small data sets. At the same time, it is understood that the independence between (X, Y) and the confounding variable is less common in real situations. The relationship between the performance of proposed methods and the levels of dependence remains to be uncovered in future researches. Additionally, the tool of jackknife empirical likelihood can be further applied to scenarios where the measurement error acts as a factor to the unobserved variables of interest.

A Appendix

A.1 Assumptions

Five conditions are needed to obtain asymptotic results:

- (C1) The density function $f_U(u)$ of the random variable U is bounded away from zero and its derivative is bounded on $\mathcal{U}$, which is a compact support set of U.
- (C2) $\phi(\cdot), \psi(\cdot)$ have third-order bounded and continuous derivatives with $E[\phi(U)] = 0$ and $E[\psi(U)] = 0$.
- (C3) The kernel function $K(\cdot)$ is a univariate bounded, continuous and symmetric density function about zero with $\int t^2 K(t)dt < \infty$. The second derivative of $K(\cdot)$ is bounded on $\mathcal{R}$.

(C4) $E[|X|^4] < \infty$, $E[|Y|^4] < \infty$.
(C5) As $n \to \infty$, $nh^4 \to 0$, $\log^2 n/(nh^2) \to 0$.

A.2 Proofs of theorems

The proofs of theorems follow the similar arguments as Huang et al. (2024) and Pidgeon et al. (2024) did. Recall that

$$\hat{\rho}_{i,(e_{\tilde{Y}U},e_{\tilde{X}U})} = \frac{\widehat{Cov}_{-i}(e_{\tilde{Y}U}, e_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{-i,e_{\tilde{Y}U}}\hat{\sigma}^2_{-i,e_{\tilde{X}U}}}},$$

where

$$\begin{aligned}
\widehat{Cov}_{-i}(e_{\tilde{Y}U}, e_{\tilde{X}U}) &= \frac{1}{n-1}\sum_{j\neq i}^{n}\hat{e}_{j\tilde{X}U}\hat{e}_{j\tilde{Y}U} - \left(\frac{1}{n-1}\sum_{j\neq i}^{n}\hat{e}_{j\tilde{X}U}\right)\left(\frac{1}{n-1}\sum_{j\neq i}^{n}\hat{e}_{j\tilde{Y}U}\right),\\
\hat{\sigma}^2_{-i,e_{\tilde{X}U}} &= \frac{1}{n-1}\sum_{j\neq i}^{n}\hat{e}^2_{j\tilde{X}U} - \left(\frac{1}{n-1}\sum_{j\neq i}^{n}\hat{e}_{j\tilde{X}U}\right)^2 \quad \text{and}\\
\hat{\sigma}^2_{-i,e_{\tilde{Y}U}} &= \frac{1}{n-1}\sum_{j\neq i}^{n}\hat{e}^2_{j\tilde{Y}U} - \left(\frac{1}{n-1}\sum_{j\neq i}^{n}\hat{e}_{j\tilde{Y}U}\right)^2.
\end{aligned}$$

By some basic derivation, it is straightforward to have

$$\begin{aligned}
\widehat{Cov}_{-i}(e_{\tilde{Y}U}, e_{\tilde{X}U}) &= \frac{n}{n-1}\widehat{Cov}(e_{\tilde{Y}U}, e_{\tilde{X}U}) - \frac{n}{(n-1)^2}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U}),\\
\hat{\sigma}^2_{-i,e_{\tilde{X}U}} &= \frac{n}{n-1}\hat{\sigma}^2_{e_{\tilde{X}U}} - \frac{n}{(n-1)^2}(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})^2,\\
\hat{\sigma}^2_{-i,e_{\tilde{Y}U}} &= \frac{n}{n-1}\hat{\sigma}^2_{e_{\tilde{Y}U}} - \frac{n}{(n-1)^2}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^2.
\end{aligned}$$

Then, it follows that

$$\begin{aligned}
\hat{\rho}_{i,(e_{\tilde{Y}U},e_{\tilde{X}U})} &= \frac{\widehat{Cov}(e_{\tilde{Y}U}, e_{\tilde{X}U}) - \frac{1}{(n-1)}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}} - \frac{1}{(n-1)}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^2}\sqrt{\hat{\sigma}^2_{e_{\tilde{X}U}} - \frac{1}{(n-1)}(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})^2}}\\
&= \frac{\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} - \frac{1}{(n-1)}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})/\sqrt{\hat{\sigma}^2_{e_{\tilde{X}U}}\hat{\sigma}^2_{e_{\tilde{Y}U}}}}{\sqrt{1 - \frac{1}{(n-1)}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^2/\hat{\sigma}^2_{e_{\tilde{Y}U}}}\sqrt{1 - \frac{1}{(n-1)}(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})^2/\hat{\sigma}^2_{e_{\tilde{X}U}}}}.
\end{aligned} \tag{A.1}$$

Recall that

$$e_{i\tilde{Y}U} = \tilde{Y}_i - E[\tilde{Y}_i|U_i] = Y_i - E(Y_i), \tag{A.2}$$

$$e_{i\tilde{X}U} = \tilde{X}_i - E[\tilde{X}_i|U_i] = X_i - E(X_i), \tag{A.3}$$

by the independence between U and (Y, X). According to Lemma A.1 in Zhang et al. (2017) and the definition of $\hat{e}_{i\tilde{Y}U}$, $\hat{e}_{i\tilde{X}U}$, we have

$$\hat{e}_{i\tilde{Y}U} = e_{i\tilde{Y}U} + S_\phi(U_i)h^2 + O_p(\tau_{n,h}), \tag{A.4}$$

$$\hat{e}_{i\tilde{X}U} = e_{i\tilde{X}U} + S_\psi(U_i)h^2 + O_p(\tau_{n,h}), \tag{A.5}$$

where $S_\phi(u) = -\mu_2 f_U'(u)\phi'(u)/f_U(u) - (\mu_2/2)\phi''(u)$, $S_\psi(u) = -\mu_2 f_U'(u)\psi'(u)/f_U(u) - (\mu_2/2)\psi''(u)$ with $\mu_2 = \int K(u)u^2 du$ and $\tau_{n,h} = h^3 + \sqrt{\log n/(nh)}$. In addition, by conditions (C1)–(C3), $|S_\phi(u)|$ and $|S_\psi(u)|$ is bounded. Therefore, we have the following expansion:

$$\begin{aligned}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^2 &= \{(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U}) + (S_\phi(U_i) - \bar{S}_\phi(U))h^2 + O_p(\tau_{n,h})\}^2 \\ &= (e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})^2 + 2(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})(S_\phi(U_i) - \bar{S}_\phi(U))h^2 + O_p(\tau_{n,h}),\end{aligned} \tag{A.6}$$

where $\bar{e}_{\tilde{Y}U} = n^{-1}\sum_{j=1}^n e_{j\tilde{Y}U}$ and $\bar{S}_\phi(U) = n^{-1}\sum_{j=1}^n S_\phi(U_i)$. By the assumption (C4) with finite fourth moments and the modification of Lemma 3 in Owen (1990), we have $\max_{1\le i\le n}(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})^2 = o_p(n^{1/2})$. Moreover, since $U_i, i = 1, \ldots, n$ are *i.i.d.* random variables and $S_\phi(x)$ is bounded, one has $\max_{1\le i\le n}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^2 = o_p(n^{1/2})$. Using (A.11) and (A.12) in Zhang et al. (2017), we have

$$\max_{1\le i\le n} \frac{(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^2}{(n-1)\hat{\sigma}^2_{e_{\tilde{Y}U}}} = o_p(n^{-1/2}), \tag{A.7}$$

and similarly,

$$\max_{1\le i\le n} \frac{(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})^2}{(n-1)\hat{\sigma}^2_{e_{\tilde{X}U}}} = o_p(n^{-1/2}). \tag{A.8}$$

Let σ_Y^2 and σ_X^2 denote the variances of Y and X, respectively. To prove Theorem 2.1, we need the following two lemmas.

Lemma A.1 *Under the conditions (C1)–(C5), one has that as $n \to \infty$,*

$$\sqrt{n}\left\{\frac{1}{n}\sum_{i=1}^n \hat{V}_i - \rho_0(X, Y)\right\} \xrightarrow{\mathcal{D}} N(0, \sigma^2_{\rho_0}),$$

where

$$\sigma^2_{\rho_0} = \frac{1}{4}\rho_0(X, Y)^2\left\{\frac{E[\{X - E(X)\}]^4}{\sigma_X^4} + 2\frac{E[\{X - E(X)\}^2\{Y - E(Y)\}^2]}{\sigma_X^2\sigma_Y^2}\right.$$

$$+\frac{E[\{Y-E(Y)\}]^4}{\sigma_Y^4}\Bigg\} - \rho_0(X,Y)\Bigg\{\frac{E[\{X-E(X)\}^3\{Y-E(Y)\}]}{\sigma_X^3\sigma_Y}$$
$$+\frac{E[\{X-E(X)\}\{Y-E(Y)\}^3]}{\sigma_X\sigma_Y^3}\Bigg\} + \frac{E[\{X-E(X)\}^2\{Y-E(Y)\}^2]}{\sigma_X^2\sigma_Y^2}.$$

Proof We notice that $\sigma^2_{\rho_0}$ is the same as the $\sigma^2_{\rho(X,Y)}$ given on pages 667–668 of Zhang et al. (2017). We first consider to get the expansion of $\hat{\rho}_{i,(e_{\tilde{Y}U},e_{\tilde{X}U})}$. By (A.7), (A.8) and the Taylor's expansion, (A.1) can be written as

$$\begin{aligned}
\hat{\rho}_{i,(e_{\tilde{Y}U},e_{\tilde{X}U})} &= \left\{\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} - \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{(n-1)\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}}\right\} \\
&\quad\times\left\{1+\frac{1}{2}\Delta_i+\frac{3}{8}\Delta_i^2+o_p(n^{-3/2})\right\} \\
&= \hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} - \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{(n-1)\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} + \frac{\Delta_i\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{2} \\
&\quad - \frac{1}{2}\Delta_i\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{(n-1)\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} + \frac{3}{8}\hat{\rho}\Delta_i^2 \\
&\quad - \frac{3}{8}\Delta_i^2\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{(n-1)\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} + o_p(n^{-3/2}),
\end{aligned} \tag{A.9}$$

where

$$\Delta_i = \frac{(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2}{(n-1)\hat{\sigma}^2_{e_{\tilde{X}U}}} + \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2}{(n-1)\hat{\sigma}^2_{e_{\tilde{Y}U}}} - \frac{(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2}{(n-1)^2\hat{\sigma}^2_{e_{\tilde{X}U}}\hat{\sigma}^2_{e_{\tilde{Y}U}}}.$$

Since $\max_{1\le i\le n}(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^4 = o_p(n)$ and $\max_{1\le i\le n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^4 = o_p(n)$, one has

$$\Delta_i^2 = \frac{(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^4}{(n-1)^2(\hat{\sigma}^2_{e_{\tilde{X}U}})^2} + \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^4}{(n-1)^2(\hat{\sigma}^2_{e_{\tilde{Y}U}})^2} + \frac{2(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2}{(n-1)^2\hat{\sigma}^2_{e_{\tilde{X}U}}\hat{\sigma}^2_{e_{\tilde{Y}U}}} + o_p(n^{-3/2}). \tag{A.10}$$

Then, according to $\max_{1\le i\le n}|\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U}| = o_p(n^{1/4})$, $\max_{1\le i\le n}|\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U}| = o_p(n^{1/4})$ and the expansion (A.9), we have

$$\begin{aligned}
\hat{V}_i &= n\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} - (n-1)\hat{\rho}_{i,(e_{\tilde{Y}U},e_{\tilde{X}U})} \\
&= \hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} + \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} - \frac{(n-1)\Delta_i\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{2}
\end{aligned}$$

$$+\frac{1}{2}\Delta_i \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} - \frac{3}{8}(n-1)\hat{\rho}\Delta_i^2$$
$$+\frac{3}{8}\Delta_i^2 \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} + o_p(n^{-1/2}),$$
$$= \hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} + \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} - \frac{(n-1)\Delta_i\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{2}$$
$$+\frac{1}{2}\Delta_i \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} - \frac{3}{8}(n-1)\hat{\rho}\Delta_i^2 + o_p(n^{-1/2}). \quad \text{(A.11)}$$

Then, it follows from (A.11) that

$$\frac{1}{n}\sum_{i=1}^{n}\hat{V}_i = \hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} + \frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}}$$
$$-\frac{1}{2}\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2}{\hat{\sigma}^2_{e_{\tilde{Y}U}}} + \frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2}{\hat{\sigma}^2_{e_{\tilde{X}U}}}\right\}$$
$$+\frac{\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{2(n-1)}\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2}{\hat{\sigma}^2_{e_{\tilde{X}U}}\hat{\sigma}^2_{e_{\tilde{Y}U}}}$$
$$+\frac{1}{2(n-1)}\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^3}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^{1/2}(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{3/2}}$$
$$+\frac{1}{2(n-1)}\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^3(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^{3/2}(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{1/2}}$$
$$-\frac{3\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{8(n-1)}\left\{\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^4}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^2} + \frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^4}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^2}\right\}$$
$$-\frac{3\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{4(n-1)}\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2}{\hat{\sigma}^2_{e_{\tilde{X}U}}\hat{\sigma}^2_{e_{\tilde{Y}U}}} + o_p(n^{-1/2})$$
$$= \hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} - \frac{\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{4(n-1)}\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2}{\hat{\sigma}^2_{e_{\tilde{X}U}}\hat{\sigma}^2_{e_{\tilde{Y}U}}}$$
$$-\frac{3\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{8(n-1)}\left\{\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^4}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^2} + \frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^4}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^2}\right\}$$
$$+\frac{1}{2(n-1)}\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^3}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^{1/2}(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{3/2}}$$

$$
\begin{aligned}
&+\frac{1}{2(n-1)}\frac{n^{-1}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^3(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^{3/2}(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{1/2}}+o_p(n^{-1/2})\\
:=\,&\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}-\frac{\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{4(n-1)}\frac{\Pi_1}{\hat{\sigma}^2_{e_{\tilde{X}U}}\hat{\sigma}^2_{e_{\tilde{Y}U}}}-\frac{3\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{8(n-1)}\left\{\frac{\Pi_2}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^2}+\frac{\Pi_3}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^2}\right\}\\
&+\frac{1}{2(n-1)}\left\{\frac{\Pi_4}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^{1/2}(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{3/2}}+\frac{\Pi_5}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^{3/2}(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{1/2}}\right\}+o_p(n^{-1/2}).
\end{aligned}
$$

We now show that Π_1, Π_2, Π_3, Π_4 and Π_5 are at the order of $O_p(1)$. By (A.4), (A.5) and (A.6), it is straightforward to derive

$$
\begin{aligned}
\Pi_1&=\frac{1}{n}\sum_{i=1}^{n}(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2\\
&=\frac{1}{n}\sum_{i=1}^{n}\left\{(e_{i\tilde{Y}U}-\bar{e}_{\tilde{Y}U})^2+2(e_{i\tilde{Y}U}-\bar{e}_{\tilde{Y}U})(S_\phi(U_i)-\bar{S}_\phi(U))h^2+O_p(\tau_{n,h})\right\}\\
&\quad\times\{(e_{i\tilde{X}U}-\bar{e}_{\tilde{X}U})^2+2(e_{i\tilde{X}U}-\bar{e}_{\tilde{X}U})(S_\psi(U_i)-\bar{S}_\psi(U))h^2+O_p(\tau_{n,h})\}\\
&=\frac{1}{n}\sum_{i=1}^{n}(e_{i\tilde{Y}U}-\bar{e}_{\tilde{Y}U})^2(e_{i\tilde{X}U}-\bar{e}_{\tilde{X}U})^2\\
&\quad+\frac{2h^2}{n}\sum_{i=1}^{n}(e_{i\tilde{X}U}-\bar{e}_{\tilde{X}U})^2(e_{i\tilde{Y}U}-\bar{e}_{\tilde{Y}U})(S_\phi(U_i)-\bar{S}_\phi(U))\\
&\quad+\frac{2h^2}{n}\sum_{i=1}^{n}(e_{i\tilde{Y}U}-\bar{e}_{\tilde{Y}U})^2(e_{i\tilde{X}U}-\bar{e}_{\tilde{X}U})(S_\psi(U_i)-\bar{S}_\psi(U))+O_p(\tau_{n,h})\\
&=\frac{1}{n}\sum_{i=1}^{n}(e_{i\tilde{Y}U}-\bar{e}_{\tilde{Y}U})^2(e_{i\tilde{X}U}-\bar{e}_{\tilde{X}U})^2+o_p(1),\\
&=\frac{1}{n}\sum_{i=1}^{n}e^2_{i\tilde{Y}U}e^2_{i\tilde{X}U}-2\bar{e}_{\tilde{Y}U}\frac{1}{n}\sum_{i=1}^{n}e_{i\tilde{Y}U}e^2_{i\tilde{X}U}+\bar{e}^2_{\tilde{Y}U}\frac{1}{n}\sum_{i=1}^{n}e^2_{i\tilde{X}U}-2\bar{e}_{\tilde{X}U}\frac{1}{n}\sum_{i=1}^{n}e^2_{i\tilde{Y}U}e_{i\tilde{X}U}\\
&\quad+4\bar{e}_{\tilde{Y}U}\bar{e}_{\tilde{X}U}\frac{1}{n}\sum_{i=1}^{n}e_{i\tilde{Y}U}e_{i\tilde{X}U}+\bar{e}^2_{\tilde{X}U}\frac{1}{n}\sum_{i=1}^{n}e^2_{i\tilde{Y}U}-3\bar{e}^2_{\tilde{Y}U}\bar{e}^2_{\tilde{X}U}+o_p(1),
\end{aligned}\tag{A.12}
$$

where the second to last equality is obtained by the conditions (C2) and (C5) and the law of large number theory for the i.i.d. random variable $S_\phi(U_i)$, $S_\psi(U_i)$, $e_{i\tilde{Y}U}$, $e_{i\tilde{X}U}$. Since $\bar{e}_{\tilde{Y}U}=o_p(1)$ and $\bar{e}_{\tilde{X}U}=o_p(1)$ from the central limit theory and the moment condition (C4) holds, it follows from (A.12) that

$$
\begin{aligned}
\Pi_1&=\frac{1}{n}\sum_{i=1}^{n}e^2_{i\tilde{Y}U}e^2_{i\tilde{X}U}-2\bar{e}_{\tilde{Y}U}\frac{1}{n}\sum_{i=1}^{n}e_{i\tilde{Y}U}e^2_{i\tilde{X}U}+\bar{e}^2_{\tilde{Y}U}\frac{1}{n}\sum_{i=1}^{n}e^2_{i\tilde{X}U}-2\bar{e}_{\tilde{X}U}\frac{1}{n}\sum_{i=1}^{n}e^2_{i\tilde{Y}U}e_{i\tilde{X}U}\\
&\quad+4\bar{e}_{\tilde{Y}U}\bar{e}_{\tilde{X}U}\frac{1}{n}\sum_{i=1}^{n}e_{i\tilde{Y}U}e_{i\tilde{X}U}+\bar{e}^2_{\tilde{X}U}\frac{1}{n}\sum_{i=1}^{n}e^2_{i\tilde{Y}U}-3\bar{e}^2_{\tilde{Y}U}\bar{e}^2_{\tilde{X}U}+o_p(1)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n} e_{i\tilde{Y}U}^2 e_{i\tilde{X}U}^2 - o_p(1)\{E(e_{\tilde{Y}U}e_{\tilde{X}U}^2) + o_p(1)\} + o_p(1)\{E(e_{\tilde{X}U}^2) + o_p(1)\} \\
&\quad - o_p(1)\{E(e_{\tilde{Y}U}^2 e_{\tilde{X}U}) + o_p(1)\} + o_p(1)\{E(e_{\tilde{Y}U}e_{\tilde{X}U}) + o_p(1)\} \\
&\quad + o_p(1)\{E(e_{i\tilde{Y}U}^2) + o_p(1)\} + o_p(1) \\
&= E(e_{\tilde{Y}U}^2 e_{\tilde{X}U}^2) + o_p(1).
\end{aligned} \tag{A.13}
$$

Therefore, by (A.2) and (A.3), the law of large number yields that as $n \to \infty$,

$$
\Pi_1 \xrightarrow{p} E(e_{\tilde{Y}U}^2 e_{\tilde{X}U}^2) = E[\{Y - E(Y)\}^2\{X - E(X)\}^2]. \tag{A.14}
$$

As for Π_2, similar to expansion (A.12) and (A.13), one has that as $n \to \infty$,

$$
\begin{aligned}
\Pi_2 &= \frac{1}{n}\sum_{i=1}^{n}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^4 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})^4 + 4(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})^3(S_\phi(U_i) - \bar{S}_\phi(U))h^2 + O_p(\tau_{n,h})\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(e_{i\tilde{Y}U}^4 - 4e_{i\tilde{Y}U}^3\bar{e}_{\tilde{Y}U} + 6e_{i\tilde{Y}U}^2\bar{e}_{\tilde{Y}U}^2 - 4e_{i\tilde{Y}U}\bar{e}_{\tilde{Y}U}^3 + \bar{e}_{\tilde{Y}U}^4\right) \\
&\quad + \frac{4h^2}{n}\sum_{i=1}^{n}\left(e_{i\tilde{Y}U}^3 - 3e_{i\tilde{Y}U}^2\bar{e}_{\tilde{Y}U} + 3e_{i\tilde{Y}U}\bar{e}_{\tilde{Y}U}^2\right)(S_\phi(U_i) - \bar{S}_\phi(U)) + o_p(1) \\
&\xrightarrow{p} E(e_{\tilde{Y}U}^4) \\
&= E\{Y - E(Y)\}^4.
\end{aligned} \tag{A.15}
$$

In the similar way, as $n \to \infty$,

$$
\Pi_3 = \frac{1}{n}\sum_{i=1}^{n}(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})^4 \xrightarrow{p} E(e_{\tilde{X}U}^4) = E\{X - E(X)\}^4. \tag{A.16}
$$

On the other hand, one can derive that as $n \to \infty$,

$$
\begin{aligned}
\Pi_4 &= \frac{1}{n}\sum_{i=1}^{n}(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^3(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})^3(e_{i\tilde{X}U} - \bar{e}_{\tilde{X}U}) + h^2(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})^3(S_\psi(U_i) - \bar{S}_\psi(U))\right\} \\
&\quad + \frac{3h^2}{n}\sum_{i=1}^{n}(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})^2(e_{i\tilde{X}U} - \bar{e}_{\tilde{X}U})(S_\phi(U_i) - \bar{S}_\phi(U)) + o_p(1)
\end{aligned}
$$

$$\begin{aligned}&\xrightarrow{p} E(e^3_{\tilde{Y}U} e_{\tilde{X}U})\\&= E[\{Y - E(Y)\}^3\{X - E(X)\}].\end{aligned} \tag{A.17}$$

And similarly, as $n \to \infty$,

$$\Pi_5 \xrightarrow{p} E(e_{\tilde{Y}U} e^3_{\tilde{X}U}) = E[\{Y - E(Y)\}\{X - E(X)\}^3]. \tag{A.18}$$

Therefore, by (A.14)–(A.18) and Theorem 2.2 in Zhang et al. (2017), we have $n^{-1}\sum_{i=1}^n \hat{V}_i = \hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})} + o_p(n^{-1/2})$. As a result, as $n \to \infty$,

$$\begin{aligned}\sqrt{n}\left\{\frac{1}{n}\sum_{i=1}^n \hat{V}_i - \rho_0(X, Y)\right\} &= \sqrt{n}\{\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})} - \rho_0(X, Y)\} + o_p(1)\\&\xrightarrow{\mathcal{D}} N(0, \sigma^2_{\rho_0}).\end{aligned}$$

□

Lemma A.2 *Under the conditions (C1)–(C5), we have that as $n \to \infty$,*

$$S_n = \frac{1}{n}\sum_{i=1}^n \{\hat{V}_i - \rho_0(X, Y)\}^2 \xrightarrow{p} \sigma^2_{\rho_0}.$$

Proof Since $\max_{1\le i\le n} |\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U}| = o_p(n^{1/4})$, $\max_{1\le i\le n} |\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U}| = o_p(n^{1/4})$, (A.7) and (A.8) hold, we have $\max_{1\le i\le n} \Delta_i^2 = o_p(n^{-1})$ and

$$\max_{1\le i\le n}\left|\Delta_i \frac{(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}} \hat{\sigma}^2_{e_{\tilde{X}U}}}}\right| = o_p(1).$$

It follows from (A.11) that

$$\begin{aligned}\hat{V}_i = \hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})} &+ \frac{(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}} \hat{\sigma}^2_{e_{\tilde{X}U}}}} - \frac{1}{2}\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})} \frac{(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^2}{\hat{\sigma}^2_{e_{\tilde{Y}U}}}\\&- \frac{1}{2}\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})} \frac{(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})^2}{\hat{\sigma}^2_{e_{\tilde{X}U}}} + o_p(1).\end{aligned} \tag{A.19}$$

Then, we have

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\hat{V}_i^2 &= \frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})} + \frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} - \frac{1}{2}\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2}{\hat{\sigma}^2_{e_{\tilde{Y}U}}}\right. \\
&\qquad \left. -\frac{1}{2}\hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}\frac{(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2}{\hat{\sigma}^2_{e_{\tilde{X}U}}} + o_p(1)\right\}^2 \\
&= \hat{\rho}^2_{(e_{\tilde{Y}U},e_{\tilde{X}U})} + \frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2}{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}} \\
&\quad + \frac{\hat{\rho}^2_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{4}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^4}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^2} + \frac{2}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^2(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^2}{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}\right. \\
&\qquad \left. +\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^4}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^2}\right\} \\
&\quad - \hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})^3(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{3/2}(\hat{\sigma}^2_{e_{\tilde{X}U}})^{1/2}}\right. \\
&\qquad \left. +\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{e}_{i\tilde{Y}U}-\bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U}-\bar{\hat{e}}_{\tilde{X}U})^3}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{1/2}(\hat{\sigma}^2_{e_{\tilde{X}U}})^{3/2}}\right\} + o_p(1) \\
&= \hat{\rho}^2_{(e_{\tilde{Y}U},e_{\tilde{X}U})} + \frac{\Pi_1}{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}} + \frac{\hat{\rho}^2_{(e_{\tilde{Y}U},e_{\tilde{X}U})}}{4}\left\{\frac{\Pi_2}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^2} + \frac{2\Pi_1}{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}} + \frac{\Pi_3}{(\hat{\sigma}^2_{e_{\tilde{X}U}})^2}\right\} \\
&\quad - \hat{\rho}_{(e_{\tilde{Y}U},e_{\tilde{X}U})}\left\{\frac{\Pi_4}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{3/2}(\hat{\sigma}^2_{e_{\tilde{X}U}})^{1/2}} + \frac{\Pi_5}{(\hat{\sigma}^2_{e_{\tilde{Y}U}})^{1/2}(\hat{\sigma}^2_{e_{\tilde{X}U}})^{3/2}}\right\} + o_p(1).
\end{aligned}
$$

From Lemma A.1, one has $n^{-1}\sum_{i=1}^{n}\hat{V}_i = \rho_0(X,Y) + O_p(n^{-1/2})$. According to (A.14)–(A.18) and Theorem 2.2 in Zhang et al. (2017), we have that as $n \to \infty$,

$$
\begin{aligned}
S_n &= \frac{1}{n}\sum_{i=1}^{n}\{\hat{V}_i - \rho_0(X,Y)\}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\hat{V}_i^2 - 2\rho_0(X,Y)\frac{1}{n}\sum_{i=1}^{n}\hat{V}_i + \rho_0(X,Y)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\hat{V}_i^2 - \rho_0(X,Y)^2 + O_p(n^{-1/2}) \\
&\xrightarrow{p} \frac{E[\{Y-E(Y)\}^2\{X-E(X)\}^2]}{\sigma_Y^2\sigma_X^2} \\
&\quad + \frac{\rho_0(X,Y)^2}{4}\left\{\frac{E\{Y-E(Y)\}^4}{\sigma_Y^4} + \frac{2E[\{Y-E(Y)\}^2\{X-E(X)\}^2]}{\sigma_Y^2\sigma_X^2} + \frac{E\{X-E(X)\}^4}{\sigma_X^4}\right\} \\
&\quad - \rho_0(X,Y)\left\{\frac{E[\{Y-E(Y)\}^3\{X-E(X)\}]}{\sigma_Y^3\sigma_X} + \frac{E[\{Y-E(Y)\}\{X-E(X)\}^3]}{\sigma_Y\sigma_X^3}\right\} \\
&= \sigma^2_{\rho_0}.
\end{aligned}
$$

Proof of Theorem 2.1 Let $W_n = \max_{1\le i\le n} |\hat{V}_i - \rho_0(X,Y)|$. We first show that $W_n = o_p(n^{1/2})$. In fact, it follows from the expansion (A.19) and Lemma A.1 that for any $1 \le i \le n$,

$$
\begin{aligned}
|\hat{V}_i - \rho_0(X,Y)| &\le |\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})} - \rho_0(X,Y)| + \frac{|(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})|}{\sqrt{\hat{\sigma}^2_{e_{\tilde{Y}U}}\hat{\sigma}^2_{e_{\tilde{X}U}}}} \\
&\quad + \frac{1}{2}|\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})}| \frac{(\hat{e}_{i\tilde{Y}U} - \bar{\hat{e}}_{\tilde{Y}U})^2}{\hat{\sigma}^2_{e_{\tilde{Y}U}}} + \frac{1}{2}|\hat{\rho}_{(e_{\tilde{Y}U}, e_{\tilde{X}U})}| \frac{(\hat{e}_{i\tilde{X}U} - \bar{\hat{e}}_{\tilde{X}U})^2}{\hat{\sigma}^2_{e_{\tilde{X}U}}} + o_p(1) \\
&= \frac{|(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})(e_{i\tilde{X}U} - \bar{e}_{\tilde{X}U})\{1 + o_p(1)\}|}{\sqrt{\sigma^2_Y \sigma^2_X + o_p(1)}} \\
&\quad + \frac{1}{2}\frac{(e_{i\tilde{Y}U} - \bar{e}_{\tilde{Y}U})^2\{|\rho_0(X,Y)| + o_p(1)\}\{1 + o_p(1)\}}{\sigma^2_Y + o_p(1)} \\
&\quad + \frac{1}{2}\frac{(e_{i\tilde{X}U} - \bar{e}_{\tilde{X}U})^2\{|\rho_0(X,Y)| + o_p(1)\}\{1 + o_p(1)\}}{\sigma^2_X + o_p(1)} + o_p(1),
\end{aligned}
$$

where the last equality can be similarly obtained by (A.11), (A.12), (A.14) and (A.15) in Zhang et al. (2017) and (A.6). Since $\{e_{i\tilde{Y}U}, e_{i\tilde{X}U}\}_{i=1}^n$ are i.i.d. samples with finite fourth moments, we have $W_n = o_p(n^{1/2})$ by the modification of Lemma 3 in Owen (1990). Note that from equation (2.1), we have

$$
\begin{aligned}
0 &= \frac{1}{n}\left|\sum_{i=1}^n \{\hat{V}_i - \rho_0(X,Y)\} - \lambda \sum_{i=1}^n \frac{\{\hat{V}_i - \rho_0(X,Y)\}^2}{1 + \lambda\{\hat{V}_i - \rho_0(X,Y)\}}\right| \\
&\ge \frac{|\lambda| S_n}{1 + |\lambda| W_n} - \frac{1}{n}\left|\sum_{i=1}^n \{\hat{V}_i - \rho_0(X,Y)\}\right|. \qquad \text{(A.20)}
\end{aligned}
$$

Combining Lemma A.1, Lemma A.2, and the inequality (A.20), one can obtain

$$
\lambda = S_n^{-1}\left\{\frac{1}{n}\sum_{i=1}^n \hat{V}_i - \rho_0(X,Y)\right\} + o_p(n^{-1/2}).
$$

Let $\gamma_i = \lambda\{\hat{V}_i - \rho_0(X,Y)\}$. We have

$$
\begin{aligned}
-2\log J(\rho_0(X,Y)) &= 2\sum_{i=1}^n \log(1+\gamma_i) \\
&= 2n\lambda\left\{\frac{1}{n}\sum_{i=1}^n \hat{V}_i - \rho_0(X,Y)\right\} - nS_n\lambda^2 + o_p(1) \\
&= \frac{n\left\{n^{-1}\sum_{i=1}^n \hat{V}_i - \rho_0(X,Y)\right\}^2}{S_n} + o_p(1).
\end{aligned}
$$

Therefore, we have $-2\log J(\rho_0(X, Y)) \xrightarrow{\mathcal{D}} \chi_1^2$ from Slutsky's theorem. □

Proof of Theorem 2.2 Following the similar arguments on proofs of Theorem 2 in Zhao et al. (2015), we prove Theorem 2.2. We omit the details. □

Proof of Theorem 2.3 Following the same lines on proofs of Theorem 3.1 in Liang et al. (2019), we prove Theorem 2.3 similar as Huang et al. (2024) did.

Proof of Theorem 2.4 Combining the results of Theorems 2.2 and 2.3, we prove Theorem 2.4 like Huang et al. (2024). □

Acknowledgements The authors would like to thank the Editor-in-Chief, Professor Jose Angulo, the Associate Editor and two reviews for their helpful comments and constructive suggestions. Yichuan Zhao is partly supported by the Simons Foundation (638679) and the National Science Foundation (DMS-2409876).

References

Alemdjrodo K, Zhao Y (2019) Reduce the computation in jackknife empirical likelihood for comparing two correlated Gini indices. J Nonparametr Stat 31:849–866
Bouzebda S, Keziou A (2024) Empirical likelihood based confidence regions for functional of copulas. J Nonparametr Stat. https://doi.org/10.1080/10485252.2024.2312396
Chen Y, Ning W (2016) Adjusted jackknife empirical likelihood. arXiv:1603.04093
Chen J, Variyath A, Abraham B (2008) Adjusted empirical likelihood and its properties. J Comput Graph Stat 17:426–443
Cheng Y, Zhao Y (2019) Bayesian jackknife empirical likelihood. Biometrika 106(4):981–988
Cheng G, Zhao Y, Li B (2012) Empirical likelihood inferences for the semiparametric additive isotonic regression. J Multivar Anal 112:172–182
Feng Z, Zhang J, Chen Q (2020) Statistical inference for linear regression models with additive distortion measurement errors. Stat Pap 61(6):2483–2509
Fu L, Hu S, Li J (2024) Robust penalized empirical likelihood in high dimensional longitudinal data analysis. J Stat Plann Inference 228:11–22
Harrison DJ, Rubinfeld DL (1978) Hedonic housing prices and demand for clean air. J Environ Econ Manage 5:81–108
Hewage S, Sang Y (2024) Jackknife empirical likelihood confidence intervals for the categorical Gini correlation. J Stat Plann Inference 231:106123
Huang H, Zhao Y (2018) Empirical likelihood for the bivariate survival function under univariate censoring. J Stat Plann Inference 194:32–46
Huang L, Zhang L, Zhao Y (2024) Jackknife empirical likelihood for the lower-mean ratio. J Nonparametr Stat 36:287–312
Jing BY, Yuan J, Zhou W (2009) Jackknife empirical likelihood. J Am Stat Assoc 104:1224–1232
Lazar N (2003) Bayesian empirical likelihood. Biometrika 90:319–326
Lazar N (2021) A review of empirical likelihood. Ann Rev Stat Appl 8:329–344
Liang W, Dai H, He S (2019) Mean empirical likelihood. Comput Stat Data Anal 138:155–169
Lin H, Li Z, Wang D, Zhao Y (2017) Jackknife empirical likelihood for the error variance in linear models. J Nonparametr Stat 29:151–166
Liu AA, Liang HY (2017) Jackknife empirical likelihood of error variance in partially linear varying-coefficient errors-in-variables models. Stat Pap 58:95–122
Liu P, Zhao Y (2023) A review of recent advances in empirical likelihood. WIREs Comput Stat 15(3):e1599e1599
Lock R (1993) New car data. J Stat Educ 1(1)
Owen AB (1988) Empirical likelihood ratio confidence intervals for a single functional. Biometrika 75:237–249
Owen A (1990) Empirical likelihood ratio confidence regions. Ann Stat 18:90–120

Owen AB (2001) Empirical likelihood. Chapman & Hall, London
Pidgeon B, Liu P, Zhao Y (2024) Jackknife empirical likelihood for the correlation coefficient with multiplicative distortion measurement error. J Nonparametr Stat. https://doi.org/10.1080/10485252.2024.2342304
Sang Y, Dang X, Zhao Y (2019) Jackknife empirical likelihood methods for Gini correlations and their equality testing. J Stat Plann Inference 199:45–59
Sentürk D, Müller H (2005) Covariate adjusted correlation analysis via varying coefficient models. J Stat 32:365–383
Sentürk D, Müller H (2008) Generalized varying coefficient models for longitudinal data. Biometrika 95:653–666
Silverman BW (1986) Density estimation for statistics and data analysis. Chapman & Hall, London
Wang D, Wu T, Zhao Y (2019) Penalized empirical likelihood for the sparse Cox regression model. J Stat Plann Inference 201:71–85
Xu HX, Fan GL, Wang JF (2022) Jackknife empirical likelihood for the error variance in linear errors-in-variables models with missing data. Commun Stat Theory Methods 51:4827–4840
Yang H, Zhao Y (2017) Smoothed jackknife empirical likelihood for the one-sample difference of quantiles. Comput Stat Data Anal 120:58–69
Yu W, Bondell H (2024) Variational Bayes for fast and accurate empirical likelihood inference. J Am Stat Assoc 119:1089–1101
Yu X, Zhao Y (2019a) Empirical likelihood inference for semi-parametric transformation models with length-biased sampling. Comput Stat Data Anal 132:115–125
Yu X, Zhao Y (2019b) Jackknife empirical likelihood inference for the accelerated failure time model. TEST 28:269–288
Zhang J, Feng Z, Zhou B (2014) A revisit to correlation analysis for distortion measurement error data. J Multivar Anal 124:116–129
Zhang J, Chen Q, Zhou N (2017a) Correlation analysis with additive distortion measurement errors. J Stat Comput Simul 87:664–688
Zhang J, Zhou Y, Lin B, Yu Y (2017b) Estimation and hypothesis test on partial linear models with additive distortion measurement errors. Comput Stat Data Anal 112:114–128
Zhang J, Zhu J, Zhou Y, Cui X, Lu T (2020) Multiplicative regression models with distortion measurement errors. Stat Pap 61(5):2031–2057
Zhao Y, Meng X, Yang H (2015) Jackknife empirical likelihood inference for the mean absolute deviation. Comput Stat Data Anal 91:92–101
Zhao Y, Moss A, Yang H, Zhang Y (2018) Jackknife empirical likelihood for the skewness and kurtosis. Stat Interface 11(4):709–719
Zhao P, Haziza D, Wu C (2022) Sample empirical likelihood and the design-based oracle variable selection theory. Stat Sin 32:435–457
Zheng M, Yu W (2013) Empirical likelihood method for multivariate Cox regression. Comput Stat 28:1241–1267