Ann Inst Stat Math (2017) 69:599–626
DOI 10.1007/s10463-016-0554-0

New non-parametric inferences for low-income proportions

Shan Luo[1] · **Gengsheng Qin**[1]

Received: 13 August 2013 / Revised: 14 September 2015 / Published online: 18 February 2016

Abstract Low-income proportion is an important index in describing the inequality of an income distribution. It has been widely used by governments in measuring social stability around the world. Established inferential methods for this index are based on the empirical estimator of the index. It may have poor finite sample performances when the real income data are skewed or has outliers. In this paper, based on a smooth estimator for the low-income proportion, we propose a smoothed jackknife empirical likelihood approach for inferences of the low-income proportion. Wilks theorem is obtained for the proposed jackknife empirical likelihood ratio statistic. Various confidence intervals based on the smooth estimator are constructed. Extensive simulation studies are conducted to compare the finite sample performances of the proposed intervals with some existing intervals. Finally, the proposed methods are illustrated by a public income dataset of the professors in University System of Georgia.

Keywords Bootstrap · Confidence interval · Cross-validation · Empirical likelihood · Jackknife · Low-income proportion

1 Introduction

Low-income proportion (LIP) is an important index in describing the inequality of an income distribution. It is often used to evaluate the social economic and poverty status of a population. A low-income proportion is defined as the proportion of the population income below a given fraction α $(0 < \alpha < 1)$ of the βth $(0 < \beta < 1)$ quantile

✉ Gengsheng Qin
gqin@gsu.edu

1 Department of Mathematics and Statistics, Georgia State University, 30 Pryor Street, Atlanta, GA 30303, USA

of an income distribution. Let $X \in [0, \infty)$ be an income variable with cumulative distribution function $F(x)$ and density function $f(x)$, and denote ξ_β as the βth quantile $F^{-1}(\beta)$ of F. Then, the α fraction of the βth quantile $\alpha\xi_\beta$ is the income line, and the low-income proportion is

$$\theta_{\alpha\beta} = P(X \leq \alpha\xi_\beta) = F(\alpha\xi_\beta).$$

As a general social economic indicator, the low-income proportion may provide meaningful information to government regulators, business owners, and individual researchers. Low-income proportion has been found in extensive applications by government. For example, based in Luxembourg, Eurostat (2000) provides EU member states with reliable statistics that allow comparisons across countries. It targets at offering a wide range of high-quality data and statistics at country level to government, commercial business, education institute, non-profit organization and other public department. According to Eurostat 2012 (Bezzina 2012), low-wage earners are defined as those employees who earn two thirds ($\alpha = \frac{2}{3}$) or less of the national median ($\beta = 0.5$) hourly earnings. For example, there are 17 % of employees in EU categorized as low-wage earners. The top five countries with the highest proportions of low-wage earners are Latvia (27.8 %), Lithuania (27.2 %), Romania (25.6 %), Poland (24.2 %), and Estonia (23.8 %), while the top five countries that own the lowest proportions of low-wage earners are Sweden (2.5 %), Finland (5.9 %), France (6.1 %), Belgium (6.4 %) and Denmark (7.7 %). A government should be on alert for a high value of low-income proportion because it indicates a potentially unstable social structure due to relative social wealth inequality.

Not only being widely used in government, the low-income proportion has also caused great interests of individual researchers. Preston (1995) discussed the reliability to estimate a low-income proportion based on simple random sample. Rongve (1997) proposed statistical inferences for the poverty index with fixed poverty lines. Zheng (2001) proved that the poverty estimators are asymptotically and normally distributed. Yves and Chris (2003) showed how a linearization method of variance estimation can be applied to low-income proportion based on Family Expenditure Survey data. However, most of the existing inferential methods are based on the simple empirical estimator and its asymptotic normal distribution. Most income data are highly right skewed due to a small percent of individuals having extremely high salaries. Established statistical inference methods based on the simple empirical estimator may have poor finite-sample performances because of the skewness of the real income data. In regard to this challenge, recent efforts have yielded new statistical inferences for a low-income proportion (Yang et al. 2011).

Empirical likelihood (EL), introduced by Owen (1988, 1990), has been shown to have diverse advantages in statistical inference. For example, the EL method can be used to construct a confidence interval without choosing a parametric distribution; the EL-based confidence region is shaped by samples, especially in higher order asymptotic analysis, while the normal approximation method would assume a symmetrical shape for a confidence region; the EL-based method is able to construct confidence interval without variance estimation. Thus, EL-based method may have advantages in developing statistical inferences with skewed data. For instance, Zhou et al. (2006)

developed a new EL-based inference method in censored cost regression models and showed that the EL method outperforms existing methods in analyzing highly skewed health care cost data. Some recent developments of empirical likelihood include inferences for risk measures (Wei et al. 2009; Wei and Zhu 2010; Li et al. 2011) and survey data (Rao and Wu 2010). Recently, Yang et al. (2011) developed a plug-in EL method for a low-income proportion and showed that the plug-in EL-based inference achieved good performance on skewed income data. However, their empirical likelihood ratio statistic follows a scaled Chi-square distribution, which requires estimation of an unknown scale constant. To bypass the estimation of the unknown scale constant, we propose a jackknife empirical likelihood (JEL) method for a low-income proportion in this paper. The JEL was originally proposed by Jing et al. (2009) for U-statistics. The general idea of the JEL is to construct a jackknife sample which is shown to be asymptotically independent, then to implement standard EL on these jackknife pseudo-values (Quenouille 1956). Gong et al. (2010) extended the JEL method with a smoothed ROC curve estimation, and showed that the JEL method results in shorter intervals than the naive bootstrap intervals in most cases.

The remaining sections are organized as follows: in Sect. 2, a kernel estimator is proposed for a low-income proportion, and this estimator is proved to be asymptotically normal. In Sect. 3, a smoothed jackknife empirical likelihood for a low-income proportion is defined, and Wilks theorem is proven to be held for the proposed jackknife empirical likelihood ratio statistic. In Sect. 4, multiple confidence intervals for a low-income proportion are constructed based on normal approximation, bootstrap and jackknife empirical likelihood methods, respectively. In Sect. 5, extensive simulation studies are conducted to evaluate the finite sample performances of the proposed methods, and an income data set in 2012 for professors at University System of Georgia is used to illustrate the recommended methods. Proofs of main theorems will be given in the Appendix.

2 The smoothed low-income proportion

2.1 The smoothed estimator

Let $X_1, X_2, \ldots, X_n$ be a simple random sample drawn from the population X with cumulative distribution function $F(x)$. The empirical estimate for $\theta_{\alpha\beta}$ is defined as

$$\hat{\theta}_{\alpha\beta} = F_n(\alpha\hat{\xi}_\beta) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le \alpha\hat{\xi}_\beta),$$

where $F_n(x)$ is the empirical distribution function of $X_1, X_2, \ldots, X_n$, and $\hat{\xi}_\beta = F_n^{-1}(\beta)$ is the βth quantile of $F_n(x)$.

Since the empirical estimator $\hat{\theta}_{\alpha\beta}$ is a non-smoothing estimator for $\theta_{\alpha\beta}$, while $\theta_{\alpha\beta} = F(\alpha\xi_\beta)$ is a function of the smoothing income distribution F in many applications. Instead of using the non-smoothing estimator $\hat{\theta}_{\alpha\beta}$, we apply kernel method to develop a smoothed estimator for $\theta_{\alpha\beta}$. Extensive literature has shown the advantage of kernel estimation. Falk (1983, 1985) concluded that for a distribution function $F(x)$, or its quantile function $F^{-1}(x)$, their corresponding kernel-based estimators asymptotically dominate the empirical estimators.

The kernel estimator for the low-income proportion $\theta_{\alpha\beta}$ is defined as follows:

$$\hat{T}_n(\alpha, \beta) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{\alpha\hat{\xi}_\beta - X_i}{h}\right), \tag{1}$$

where the kernel function $K(x) = \int_{-\infty}^{x} \omega(y)\mathrm{d}y$, $\omega(\cdot)$ is a known probability density function, and h is a pre-selected bandwidth.

Theorem 2.1 *Assume that the probability density function $\omega(\cdot)$ of the kernel function $K(\cdot)$ has a bounded support, its first derivative $\omega^{'}(\cdot)$ exists and is bounded on its supporting set, and $\int_{-\infty}^{\infty} |\omega^{'}(y)|\mathrm{d}y < \infty$. If $h = h(n) \to 0$, $\sqrt{n}h \to \infty$ as $n \to \infty$, then*

$$\sqrt{n}\{\hat{T}_n(\alpha, \beta) - \theta_{\alpha\beta}\} \xrightarrow{d} N(0, \sigma^2_{\alpha\beta}),$$

where $\sigma^2_{\alpha\beta} = \frac{\alpha^2\beta(1-\beta)f^2(\alpha\xi_\beta))}{f^2(\xi_\beta)} - 2\alpha(1-\beta)\theta_{\alpha\beta}\frac{f(\alpha\xi_\beta)}{f(\xi_\beta)} + \theta_{\alpha\beta}(1-\theta_{\alpha\beta}).$

Remark 1 It is noticed that this smoothed estimator and the empirical estimate $\hat{\theta}_{\alpha\beta}$ have the same asymptotic variance. However, with finite sample size, the smoothed estimator seems a good alternative to the empirical estimator for the low-income proportion. We observed that $\hat{T}_n(\alpha, \beta)$ has slightly smaller MSE than $\hat{\theta}_{\alpha\beta}$ (see Table 1). This smoothed estimator is also needed in the definition of the smoothed jackknife empirical likelihood for the low-income proportion. Our simulation results (not reported here) showed that the jackknife empirical likelihood does not work without smoothing. We believe that the main reason for the failure is that the empirical estimator $\hat{\theta}_{\alpha\beta}$ for the low-income proportion is a non-smoothing function of sample quantile $\hat{\xi}_\beta$. It is well known that the jackknife method can fail if the statistic is not smooth; one such example is the sample quantile (see Shao and Tu 1995).

2.2 Bandwidth selection

One of the difficulties in the calculation of the smoothed estimator $\hat{T}_n(\alpha, \beta)$ is to choose a bandwidth h for the kernel estimator. Extensive simulation analyses have shown that the choice of the kernel function K will not change the estimate much. However, as in many kernel methods, the choice of the bandwidth h may influence the performance of the proposed kernel estimate. Many methods have been proposed for selecting the bandwidth for kernel estimators (e.g., Bowman et al. 1998). In our study, we apply a cross-validation (CV) method for bandwidth selection. To ease the implementation, we utilize the twofold cross-validation method. The bandwidth h is suggested to be $h = cn^{-1/3}$, based on our simulation analyses. Then, the choice of h

Table 1 Bias and MSE of empirical estimator and kernel estimator for LIPs with $F = \chi_3^2$

Sample size	β	$\text{Bias}_{\hat{\theta}_{\alpha\beta}}$	$\text{Bias}_{\hat{T}_n(\alpha,\beta)}$	$\text{MSE}_{\hat{\theta}_{\alpha\beta}}$	$\text{MSE}_{\hat{T}_n(\alpha,\beta)}$
500	0.2	0.000027	0.001541	0.000094	0.000046
	0.3	0.000001	0.001877	0.000145	0.000085
	0.4	0.000411	0.002508	0.000210	0.000135
	0.5	0.000591	0.003359	0.000268	0.000185
	0.6	0.000180	0.004814	0.000324	0.000232
	0.7	0.000844	0.007711	0.000365	0.000286
	0.8	0.000595	0.015540	0.000424	0.000464
800	0.2	0.000229	0.001241	0.000059	0.000031
	0.3	0.000341	0.001527	0.000094	0.000057
	0.4	0.000304	0.001996	0.000129	0.000084
	0.5	0.000346	0.002718	0.000151	0.000114
	0.6	0.000569	0.003612	0.000194	0.000148
	0.7	0.000175	0.005004	0.000242	0.000187
	0.8	0.000976	0.009472	0.000265	0.000253
1000	0.2	0.000122	0.001060	0.000052	0.000028
	0.3	0.000950	0.001616	0.000080	0.000051
	0.4	0.000576	0.001870	0.000101	0.000072
	0.5	0.000046	0.001870	0.000132	0.000098
	0.6	0.000603	0.002295	0.000166	0.000122
	0.7	0.000644	0.003832	0.000191	0.000148
	0.8	0.000525	0.008100	0.000238	0.000210

is controlled by the constant c. Here and thereafter, we denote $\hat{T}_{n,c}(\alpha, \beta) = \hat{T}_n(\alpha, \beta)$. For a given β, we select c by minimizing the Mean Squared Error (MSE):

$$\text{MSE}(c) = E[\hat{T}_{n,c}(\alpha, \beta) - \theta_{\alpha\beta}]^2.$$

For this purpose, we randomly split the sample into two equal parts, where the first part is treated as the training sample, and the other part is as the validation sample. The kernel estimate $\hat{T}_{n,c}^{(1)}(\alpha, \beta)$ for the low-income proportion is constructed based on the training sample, while the empirical estimate $\hat{\theta}_{\alpha\beta}^{(2)}$ is constructed from the validation sample. After repeating this random split L times ($L \geq 30$ is suggested based on our extensive simulation studies), we obtain a set of kernel estimates and empirical estimates $\{(\hat{T}_{n,c}^{(1,l)}(\alpha, \beta), \hat{\theta}_{\alpha\beta}^{(2,l)}) : L = 1, \ldots, L\}$ for the low-income proportion, and the following cross-validation estimate of the MSE:

$$\text{CV}_c = \frac{1}{L} \sum_{l=1}^{L} [\hat{T}_{n,c}^{(1,l)}(\alpha, \beta) - \hat{\theta}_{\alpha\beta}^{(2,l)}]^2.$$

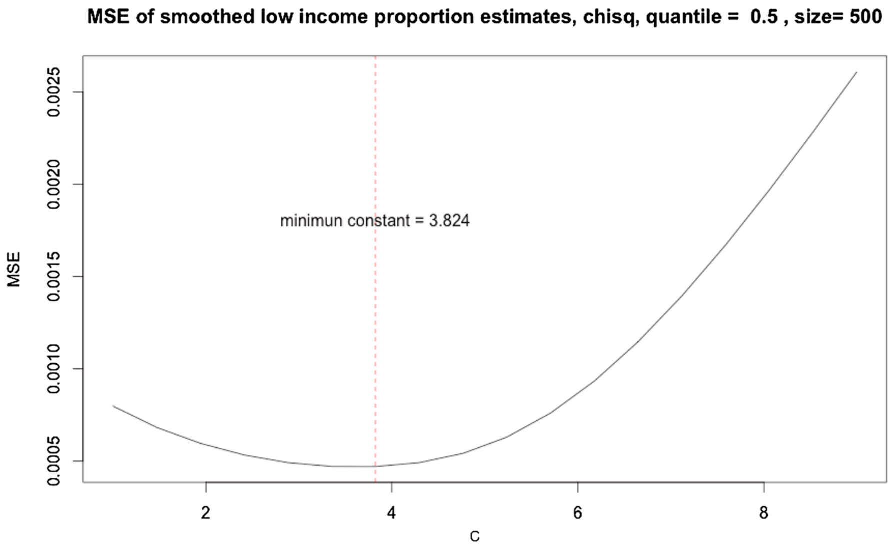

Fig. 1 Bandwidth selection by MSE

Then, the value of c is chosen as the constant that minimizes CV_c.

Figure 1 is a simulation example to illustrate the relationship between MSE and the constant c, which actually affects the bandwidth h. The value of h corresponding to the lowest point of MSE will be the optimal bandwidth.

Alternatively, if we focus on the overall performance of the smoothed estimator for low-income proportions across all β, we can use a similar cross-validation procedure for selecting c by minimizing the Average Mean Squared Error (AMSE):

$$\mathrm{AMSE}(c) = E\left\{\frac{1}{M}\sum_{m=1}^{M}[\hat{T}_{n,c}(\alpha, \beta_m) - \theta_{\alpha\beta_m}]^2\right\},$$

where β_m is a fine grid of $(0, 1)$, and M is the number of grid points.

Therefore, the cross-validation estimate of the AMSE is:

$$\mathrm{ACV}_c = \frac{1}{L}\frac{1}{M}\sum_{l=1}^{L}\sum_{m=1}^{M}[\hat{T}_{n,c}^{(1,l)}(\alpha, \beta_m) - \hat{\theta}_{\alpha\beta_m}^{(2,l)}]^2.$$

Again, c is chosen as the one that minimizes ACV_c.

Figure 2 illustrates the relationship between bandwidth and AMSE, and how we choose the constant c for bandwidth h.

Similarly, we choose the value of h corresponding to the lowest point of AMSE. This twofold cross-validation method by minimizing AMSE is applied in our study.

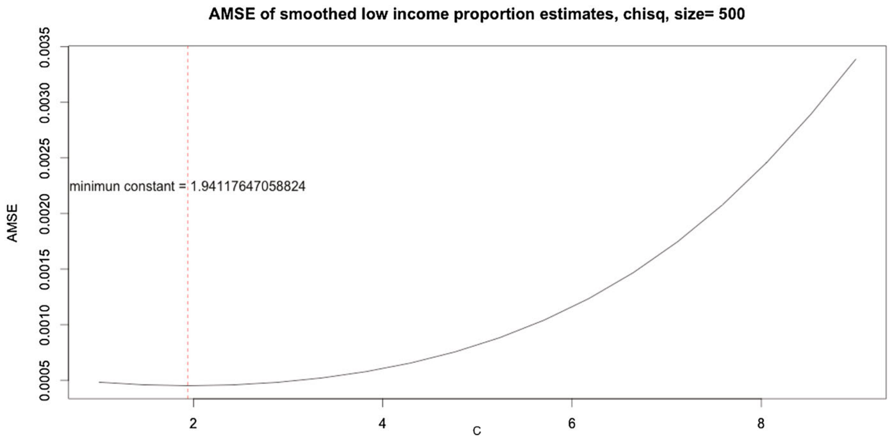

Fig. 2 Bandwidth selection by AMSE

3 Smoothed jackknife empirical likelihood for a low-income proportion

A smoothed version of jackknife empirical likelihood for the low-income proportion is defined in this section. Based on Tukey (1958), we define the jackknife pseudo-values for a low-income proportion as

$$\hat{V}_k(\alpha, \beta) = n\hat{T}_n(\alpha, \beta) - (n-1)\hat{T}_{n-1,k}(\alpha, \beta), \quad k = 1, 2, \ldots, n,$$

where $\hat{T}_{n-1,k}(\alpha, \beta) = \frac{1}{n-1}\sum_{j\neq k}^{n} K(\frac{\alpha\hat{\xi}_{\beta,-k} - X_j}{h})$ is the given statistics $\hat{T}_{n-1}(\alpha, \beta)$ but computed on $n-1$ observations $X_1, X_2, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n$, and $\hat{\xi}_{\beta,-k} = F_{n,-k}^{-1}(\beta)$ is the βth quantile of $F_{n,-k}(x) = \frac{1}{n-1}\sum_{j\neq k}^{n} I(X_j \leq x)$ which is the empirical distribution of these $n-1$ observations.

Then, the jackknife empirical likelihood for $\theta_{\alpha\beta}$ can be defined as follows:

$$L(\theta_{\alpha\beta}) = \sup\left\{\prod_{k=1}^{n} np_k : p_1 > 0, \ldots, p_n > 0, \sum_{k=1}^{n} p_k = 1, \sum_{k=1}^{n} p_k\hat{V}_k(\alpha, \beta) = \theta_{\alpha\beta}\right\}. \tag{2}$$

Using the Lagrange multiplier method, we obtain the maximization for (2) at

$$p_k = \frac{1}{n}\{1 + \lambda[\hat{V}_k(\alpha, \beta) - \theta_{\alpha\beta}]\}^{-1}, \quad k = 1, \ldots, n,$$

where $\lambda = \lambda(\alpha, \beta, \theta_{\alpha\beta})$ is the solution to

$$\frac{1}{n}\sum_{k=1}^{n}\frac{\hat{V}_k(\alpha, \beta) - \theta_{\alpha\beta}}{1 + \lambda(\hat{V}_k(\alpha, \beta) - \theta_{\alpha\beta})} = 0. \tag{3}$$

Since $\prod_{k=1}^{n} p_k$ is subject to $\sum_{k=1}^{n} p_k = 1, p_k \geq 0, k = 1, 2, \ldots, n$, $L(\theta_{\alpha\beta})$ will attain its maximum n^{-n} at $p_k = n^{-1}$. Thus, the jackknife empirical likelihood ratio statistic for $\theta_{\alpha\beta}$ can be defined as

$$L_n(\theta_{\alpha\beta}) = \prod_{k=1}^{n}(np_k) = \prod_{k=1}^{n}\{1 + \lambda(\hat{V}_k(\alpha, \beta) - \theta_{\alpha\beta})\}^{-1},$$

and the log jackknife empirical likelihood ratio statistic is

$$l_n(\theta_{\alpha\beta}) = -2\log L_n(\theta_{\alpha\beta}) = 2\sum_{k=1}^{n}\log\{1 + \lambda(\hat{V}_k(\alpha, \beta) - \theta_{\alpha\beta})\}. \tag{4}$$

We conjecture that the pseudo-values $\hat{V}_i(\alpha, \beta)$, $i = 1, \ldots, n$ could be treated as though they were i.i.d, and $\hat{V}_i(\alpha, \beta)$ has approximately the same variance as $\sqrt{n}\hat{T}_n(\alpha, \beta)$. Therefore, the variance of $\sqrt{n}\hat{T}_n(\alpha, \beta)$, denoted as $\mathrm{var}(\sqrt{n}\hat{T}_n(\alpha, \beta))$, can be estimated by the sample variance of $\{\hat{V}_1(\alpha, \beta), \ldots, \hat{V}_n(\alpha, \beta)\}$. The jackknife variance estimator of $\hat{T}_n(\alpha, \beta)$ is thus defined as follows:

$$\begin{aligned}
\upsilon_{\mathrm{JACK}}(\alpha, \beta) &= \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\hat{V}_i(\alpha, \beta) - \frac{1}{n}\sum_{j=1}^{n}\hat{V}_j(\alpha, \beta)\right)^2 \\
&= \frac{n-1}{n}\sum_{i=1}^{n}\left(\hat{T}_{n-1,i}(\alpha, \beta) - \frac{1}{n}\sum_{j=1}^{n}\hat{T}_{n-1,j}(\alpha, \beta)\right)^2.
\end{aligned}$$

The following theorem shows that this jackknife variance estimator is a consistent estimator for the asymptotic variance $\sigma^2_{\alpha\beta}$.

Theorem 3.1 *Under the conditions of Theorem 2.1, we have*

$$\upsilon_{\mathrm{JACK}}(\alpha, \beta) \xrightarrow{p} \sigma^2_{\alpha\beta},$$

where $\sigma^2_{\alpha\beta}$ is defined in Theorem 2.1.

Then, the Wilks theorem for $l_n(\theta_{\alpha\beta})$ is obtained in the following theorem.

Theorem 3.2 *Under the conditions of Theorem 3.1, if $\sqrt{n}h^2 \longrightarrow \infty$, we have*

$$l_n(\theta_{\alpha\beta}) \xrightarrow{d} \chi^2(1).$$

Detailed proofs for Theorems 3.1 and 3.2 will be given in Appendix. In the next section, we will discuss methods for constructing confidence intervals of a low-income proportion.

4 Confidence intervals for a low-income proportion

4.1 Normal approximation-based confidence intervals

One of the most popular methods to construct a confidence interval for an unknown parameter is normal approximation. To construct a normal approximation-based confidence interval for $\theta_{\alpha\beta}$, we need to first obtain an appropriate estimator for $\theta_{\alpha\beta}$, and then derive its asymptotic normal distribution. Based on Preston (1995), the empirical estimate $\hat{\theta}_{\alpha\beta}$ for the low-income proportion is asymptotically normal with variances σ_v^2, i.e., $\sqrt{n}(\hat{\theta}_{\alpha\beta} - \theta_{\alpha\beta}) \longrightarrow N(0, \sigma_v^2)$, where $\sigma_v^2 = \theta_{\alpha\beta}(1-\theta_{\alpha\beta}) - 2\alpha(1-\beta)\theta_{\alpha\beta}\frac{f(\alpha\xi_\beta)}{f(\xi_\beta)} + \alpha^2\beta(1-\beta)[\frac{f(\alpha\xi_\beta)}{f(\xi_\beta)}]^2$. Therefore, the first $(1-\alpha)$ level normal approximation-based (NA1) confidence interval for $\theta_{\alpha\beta}$ can be constructed as

$$(l_1, u_1) = \left(\hat{\theta}_{\alpha\beta} - \frac{z_{1-\frac{\alpha}{2}}\hat{\sigma}_v}{\sqrt{n}},\ \hat{\theta}_{\alpha\beta} + \frac{z_{1-\frac{\alpha}{2}}\hat{\sigma}_v}{\sqrt{n}}\right),$$

where $z_{1-\frac{\alpha}{2}}$ is the $(1-\frac{\alpha}{2})$th quantile of the standard normal distribution. $\hat{\sigma}_v^2$ is a consistent estimate for σ_v^2 and is defined as

$$\hat{\sigma}_v^2 = \hat{\theta}_{\alpha\beta}(1-\hat{\theta}_{\alpha\beta}) - 2\alpha(1-\beta)\hat{\theta}_{\alpha\beta}\frac{\hat{f}(\alpha\xi_\beta)}{\hat{f}(\xi_\beta)} + \alpha^2\beta(1-\beta)\left[\frac{\hat{f}(\alpha\xi_\beta)}{\hat{f}(\xi_\beta)}\right]^2,$$

with $\hat{f}(.)$ being the kernel density function estimate defined in Preston (1995).

As in Theorem 3.1, the jackknife variance estimator $\upsilon_{\mathrm{JACK}}(\alpha, \beta)$ is a consistent estimator for $\sigma_{\alpha\beta}^2$. Thus, the second $(1-\alpha)$ level normal approximation-based (NA2) confidence interval for $\theta_{\alpha\beta}$ can be constructed as

$$(l_2, u_2) = \left(\hat{T}_n(\alpha, \beta) - \frac{z_{1-\frac{\alpha}{2}}\sqrt{\upsilon_{\mathrm{JACK}}(\alpha, \beta)}}{\sqrt{n}},\ \hat{T}_n(\alpha, \beta) + \frac{z_{1-\frac{\alpha}{2}}\sqrt{\upsilon_{\mathrm{JACK}}(\alpha, \beta)}}{\sqrt{n}}\right).$$

4.2 Bootstrap-based confidence intervals

The normal approximation-based confidence intervals may have poor performance since the income data are skewed or has outliers. Introduced by Efron (1979), bootstrap is a powerful non-parametric approach for constructing confidence intervals when the asymptotic variance of an estimator is unknown and of a complex form. Although $\upsilon_{\mathrm{JACK}}(\alpha, \beta)$ can be used to estimate the asymptotic variance of the kernel estimator $\hat{T}_n(\alpha, \beta)$, we would also like to compare it with the bootstrap method that can estimate $\sigma_{\alpha\beta}^2$. Inspired by the bootstrap intervals based on the empirical estimator by Yang et al. (2011), we construct bootstrap intervals for $\theta_{\alpha\beta}$ based on the kernel estimator.

Let $\{X_1^*, \ldots, X_n^*\}$ be a bootstrap sample drawn from the original data $\{X_1, \ldots, X_n\}$. The bootstrap version of $\hat{T}_n(\alpha, \beta)$ is

$$\hat{T}^*(\alpha, \beta) = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{\alpha\hat{\xi}_\beta^* - X_i^*}{h}\right).$$

After repeating this bootstrap procedure B ($B \geq 500$) times, B bootstrap copies of $\hat{T}_n(\alpha, \beta)$ are obtained, denoted as $\{\hat{T}_b^*, b = 1, 2, \ldots, B\}$. The bootstrap sample variance of $\hat{T}_b^*$s is

$$V_T^* = \frac{1}{B-1} \sum_{b=1}^{B} (\hat{T}_b^* - \bar{T}^*)^2,$$

where $\bar{T}^* = \frac{1}{B} \sum_{b=1}^{B} \hat{T}_b^*$. V_T^* can be used to estimate the asymptotic variance of $\hat{T}_n(\alpha, \beta)$. Two bootstrap confidence intervals based on the kernel estimator can be constructed as follows:

1. BT1 interval:

$$(l_3, u_3) = \left(\hat{T}_n(\alpha, \beta) - z_{1-\alpha/2}\sqrt{V_T^*},\ \hat{T}_n(\alpha, \beta) + z_{1-\alpha/2}\sqrt{V_T^*}\right).$$

2. BT2 interval:

$$(l_4, u_4) = \left(\bar{T}^* - z_{1-\alpha/2}\sqrt{V_T^*},\ \bar{T}^* + z_{1-\alpha/2}\sqrt{V_T^*}\right).$$

   Another non-parametric method to construct a confidence interval for $\theta_{\alpha\beta}$ is the bootstrap bias correction and acceleration (BCa) method, which does not need a variance estimation.

3. BCa interval:

$$(l_5, u_5) = \left(\hat{T}^*_{([B\beta_1])},\ \hat{T}^*_{([B\beta_2])}\right),$$

   where

$$\beta_1 = \Phi\left(b + \frac{b + z_{\alpha/2}}{1 - a(b + z_{\alpha/2})}\right), \quad \beta_2 = \Phi\left(b + \frac{b + z_{1-\alpha/2}}{1 - a(b + z_{1-\alpha/2})}\right)$$

   with correction constants a and b defined by

$$a = \frac{1}{6} \sum_{i=1}^{n} \varphi_i^3 \Big/ \left(\sum_{i=1}^{n} \varphi_i^2\right)^{\frac{3}{2}}, \quad b = \Phi^{-1}\left(\frac{1}{B} \sum_{b=1}^{B} I(\hat{T}_b^* \leq \hat{T}_n(\alpha, \beta))\right)$$

   where $\varphi_i = \hat{T}_{(.)} - \hat{T}_{(-i)}$, and $\hat{T}_{(-i)}$ is the $\hat{T}_n(\alpha, \beta)$ computed by deleting the ith observation in original data, and $\hat{T}_{(.)} = \frac{1}{n} \sum_{i=1}^{n} \hat{T}_{(-i)}$.

4.3 Smoothed jackknife empirical likelihood-based confidence interval

The proposed smoothed jackknife empirical likelihood (SJEL) based on the kernel estimator can be used to make inference for $\theta_{\alpha\beta}$. Based on Theorem 3.2, an SJEL-based confidence interval for $\theta_{\alpha\beta}$ can be constructed as

$$(l_6, u_6) = \{\theta : l_n(\theta) \le \chi^2_{1,1-\alpha}\},$$

where $\chi^2_{1,1-\alpha}$ is the $(1-\alpha)$th quantile of χ^2_1.

5 Simulation studies and a real example

In this section, we first compare the proposed kernel estimator with the empirical estimator in terms of Bias and Mean Square Error (MSE). Then, we present results for the coverage probability and the average length of NA1, NA2, BT1, BT2, BCa, EL and SJEL intervals for the low-income proportion discussed in previous section, where EL is the plug-in EL method proposed by Yang et al. (2011). Finally, the proposed methods are illustrated by a real example.

5.1 Simulation studies

5.1.1 Point estimator evaluation

It is interesting to compare the finite sample performances of the kernel estimator $\hat{T}_n(\alpha, \beta)$ with those of the empirical estimator $\hat{\theta}_{\alpha\beta}$. The evaluation criteria used here are Bias and Mean Square Error (MSE). The MSE of $\hat{\theta}_{\alpha\beta}$ is $\text{MSE}_{\hat{\theta}_{\alpha\beta}} = E[\hat{\theta}_{\alpha\beta} - \theta_{\alpha\beta}]^2$, and the MSE of $\hat{T}_n(\alpha, \beta)$ is $\text{MSE}_{\hat{T}_n(\alpha,\beta)} = E[\hat{T}_n(\alpha, \beta) - \theta_{\alpha\beta}]^2$. MSE can be composed by two parts, the square of bias, which measures the accuracy, and the variance, which measures the precision of the estimator. Minimizing MSE can achieve the balance between the bias and the variance.

The evaluation of the kernel estimator and empirical estimator is conducted by simulation studies. The simulation setting is as follows: the fraction α of the low-income proportion is fixed at 0.5. To see how the comparisons perform across different quantiles, $\beta = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$, and 0.8 are considered. Monte Carlo simulations are employed to simulate samples from the Chi-square distribution with degree of freedom 3. The sample sizes n are chosen to be 500, 800, and 1000. One thousand random samples are generated from the above distribution. Bias and MSE are calculated based on the simulated samples.

Table 1 lists Bias and MSE for the kernel estimator and the empirical estimator. $\text{Bias}_{\hat{\theta}_{\alpha\beta}}$ is the bias calculated for the empirical estimator, while $\text{Bias}_{\hat{T}_n(\alpha,\beta)}$ is the bias for the kernel estimator. From this table, we observe that the proposed kernel estimator has smaller MSE than the empirical estimator, although the bias of the kernel estimator is larger than the bias of the empirical estimator. This observation is as expected because the kernel estimator is not an unbiased estimator for the low-income proportion. This point estimation comparison results show that our proposed kernel estimator is a competitive estimator in terms of MSE.

5.1.2 Interval estimation evaluation

In this section, we will evaluate NA1, NA2, BT1, BT2, BCa, EL and SJEL confidence intervals proposed in Sect. 4 under the same simulation settings used for the point estimation evaluation, except that Monte Carlo samples are generated from a Chi-square distribution with degree of freedom 3, and a standard Log normal distribution $\log N(0, 1)$. 90 and 95 % confidence intervals for $\theta_{\alpha\beta}$ are constructed with $\beta = 0.5, 0.6, 0.7, 0.8$. Triweight kernel density function $\omega(t) = \frac{35}{32}(1-t^2)^2 I(|t| \leq 1)$ is selected for the kernel estimator, and the constant c for the bandwidth $h = cn^{-1/3}$ is chosen via the proposed cross-validation method. For the bootstrap variance estimates, 500 bootstrap samples are drawn with replacement from the original sample.

Tables 2 and 3 display coverage probabilities and average lengths of various confidence intervals for low-income proportions with $F = \chi_3^2$ at 90 and 95 % confidence

Table 2 Coverage probabilities and average lengths of 90 % level confidence intervals for LIPs with $F = \chi_3^2$

Size	Method	$\beta = 50$ %		$\beta = 60$ %		$\beta = 70$ %		$\beta = 80$ %	
		Coverage	Length	Coverage	Length	Coverage	Length	Coverage	Length
500	NA1	0.872	0.039	0.923	0.047	0.936	0.055	0.926	0.049
	NA2	0.890	0.038	0.910	0.042	0.906	0.044	0.923	0.045
	BT1	0.911	0.039	0.899	0.042	0.885	0.045	0.897	0.045
	BT2	0.918	0.039	0.906	0.042	0.897	0.045	0.900	0.045
	BCa	0.715	0.038	0.728	0.041	0.703	0.044	0.706	0.044
	EL	0.906	0.045	0.898	0.046	0.864	0.051	0.866	0.041
	SJEL	0.908	0.040	0.896	0.043	0.905	0.046	0.908	0.046
800	NA1	0.872	0.034	0.867	0.039	0.864	0.040	0.889	0.038
	NA2	0.893	0.031	0.920	0.034	0.916	0.036	0.906	0.039
	BT1	0.906	0.031	0.898	0.034	0.911	0.036	0.902	0.037
	BT2	0.922	0.031	0.905	0.034	0.912	0.036	0.910	0.037
	BCa	0.734	0.031	0.763	0.034	0.752	0.036	0.721	0.037
	EL	0.906	0.033	0.918	0.036	0.894	0.038	0.852	0.038
	SJEL	0.898	0.032	0.894	0.035	0.899	0.037	0.904	0.038
1000	NA1	0.878	0.032	0.875	0.037	0.871	0.038	0.871	0.041
	NA2	0.923	0.030	0.903	0.031	0.927	0.033	0.876	0.034
	BT1	0.894	0.028	0.892	0.031	0.896	0.033	0.902	0.034
	BT2	0.892	0.028	0.898	0.031	0.898	0.033	0.910	0.034
	BCa	0.738	0.028	0.756	0.031	0.748	0.033	0.732	0.033
	EL	0.892	0.028	0.908	0.031	0.886	0.033	0.872	0.033
	SJEL	0.902	0.029	0.904	0.032	0.901	0.034	0.908	0.035

Table 3 Coverage probabilities and average lengths of 95 % level confidence intervals for LIPs with $F = \chi^2_3$

Size	Method	$\beta = 50\ \%$		$\beta = 60\ \%$		$\beta = 70\ \%$		$\beta = 80\ \%$	
		Coverage	Length	Coverage	Length	Coverage	Length	Coverage	Length
500	NA1	0.892	0.046	0.911	0.059	0.929	0.060	0.967	0.065
	NA2	0.953	0.046	0.946	0.051	0.936	0.054	0.950	0.053
	BT1	0.951	0.046	0.936	0.050	0.948	0.053	0.943	0.054
	BT2	0.955	0.046	0.946	0.050	0.956	0.053	0.950	0.054
	BCa	0.791	0.044	0.809	0.049	0.789	0.052	0.785	0.052
	EL	0.956	0.048	0.958	0.049	0.932	0.059	0.916	0.058
	SJEL	0.950	0.047	0.945	0.051	0.944	0.054	0.953	0.055
800	NA1	0.934	0.044	0.935	0.043	0.930	0.044	0.920	0.048
	NA2	0.940	0.038	0.940	0.041	0.950	0.045	0.970	0.045
	BT1	0.965	0.037	0.946	0.041	0.955	0.043	0.946	0.044
	BT2	0.966	0.037	0.949	0.041	0.958	0.043	0.949	0.044
	BCa	0.818	0.036	0.835	0.040	0.829	0.042	0.801	0.043
	EL	0.952	0.041	0.952	0.043	0.918	0.048	0.914	0.047
	SJEL	0.947	0.038	0.950	0.042	0.956	0.044	0.951	0.045
1000	NA1	0.924	0.039	0.932	0.040	0.929	0.041	0.929	0.045
	NA2	0.963	0.034	0.960	0.037	0.933	0.040	0.940	0.041
	BT1	0.948	0.034	0.932	0.037	0.950	0.039	0.954	0.040
	BT2	0.950	0.034	0.936	0.037	0.950	0.039	0.962	0.040
	BCa	0.794	0.033	0.828	0.036	0.840	0.039	0.820	0.038
	EL	0.952	0.036	0.952	0.039	0.928	0.042	0.912	0.041
	SJEL	0.949	0.035	0.945	0.038	0.950	0.040	0.954	0.041

levels, and Tables 4 and 5 display coverage probabilities and average lengths for low-income proportions with $F = \log N(0, 1)$ at 90 and 95 % confidence levels, respectively. According to these simulation results, we observe that all the confidence intervals perform better when sample size increases. The coverage probabilities of the newly proposed NA2 intervals based on the kernel estimator are closer to the nominal confidence levels than those of the NA1 intervals based on the empirical estimator. Among the three bootstrap-based BT1, BT2 and BCa intervals, BT1 and BT2 intervals have similar performances, and perform much better than the BCa interval. Meanwhile, compared with the plug-in empirical likelihood (EL) interval proposed by Yang et al. (2011), SJEL has comparable coverage probabilities and interval lengths in most cases, and has better coverage accuracy in some cases. Out of the 7 confidence intervals, the proposed SJEL-based confidence intervals are observed to achieve the best performance in terms of coverage probability in most cases considered here. The smoothed bootstrap-based confidence intervals (BT1 and BT2) have good finite sample performances next to the EL and SJEL intervals. Since the SJEL interval has a great advantage over the EL interval to avoid estimating the unknown scale constant,

Table 4 Coverage probabilities and average lengths of 90 % level confidence intervals for LIPs with $F = \log N(0, 1)$

Size	Method	$\beta = 50$ %		$\beta = 60$ %		$\beta = 70$ %		$\beta = 80$ %	
		Coverage	Length	Coverage	Length	Coverage	Length	Coverage	Length
500	NA1	0.842	0.048	0.888	0.057	0.857	0.060	0.883	0.061
	NA2	0.886	0.051	0.896	0.057	0.920	0.061	0.883	0.062
	BT1	0.897	0.050	0.912	0.056	0.904	0.061	0.886	0.061
	BT2	0.905	0.050	0.919	0.056	0.912	0.061	0.895	0.061
	BCa	0.778	0.050	0.764	0.056	0.763	0.060	0.723	0.060
	EL	0.894	0.055	0.898	0.055	0.892	0.061	0.884	0.062
	SJEL	0.892	0.051	0.901	0.057	0.892	0.061	0.892	0.062
800	NA1	0.869	0.038	0.856	0.047	0.865	0.042	0.864	0.054
	NA2	0.910	0.041	0.910	0.045	0.900	0.049	0.896	0.050
	BT1	0.952	0.046	0.952	0.052	0.948	0.058	0.939	0.060
	BT2	0.954	0.046	0.953	0.052	0.954	0.058	0.946	0.060
	BCa	0.882	0.048	0.860	0.053	0.876	0.057	0.813	0.058
	EL	0.882	0.043	0.894	0.048	0.896	0.058	0.882	0.057
	SJEL	0.901	0.041	0.900	0.046	0.910	0.050	0.901	0.051
1000	NA1	0.878	0.033	0.865	0.031	0.871	0.032	0.881	0.031
	NA2	0.913	0.036	0.936	0.041	0.883	0.044	0.906	0.045
	BT1	0.914	0.033	0.894	0.041	0.890	0.039	0.904	0.041
	BT2	0.918	0.033	0.902	0.041	0.900	0.039	0.912	0.041
	BCa	0.800	0.036	0.808	0.040	0.792	0.044	0.752	0.044
	EL	0.892	0.037	0.894	0.036	0.908	0.035	0.888	0.032
	SJEL	0.902	0.029	0.904	0.032	0.901	0.034	0.908	0.035

we recommend the smoothed jackknife empirical likelihood (SJEL) interval for the interval estimation of low-income proportions.

5.2 Georgia Public University employee income data example

Georgia Department of Audits and Accounts compiled annually updated salary information for all employees from each department, office, institution, board, commission, authority and agency of the State government; every university or college in the University System of Georgia; any regional educational service agency; the General Assembly including all legislative offices and agencies; offices of the Judicial Branch; local boards of education, etc. Each record has ending periods in June 30, 2008, June 30, 2009, June 30, 2010, June 30, 2011 and June 30, 2012. These income data include a list of employee's name, title or functional area, salary and travel reimbursement. The purpose of these income files is to strengthen the transparency of the Georgia government. Our analysis will be based on these individual's annual income data.

The salary information for part-time or temporary employee does not meet our annual salary definition. To minimize the downward bias introduced by those types

Table 5 Coverage probabilities and average lengths of 95 % level confidence intervals for LIPs with $F = \log N(0, 1)$

Size	Method	$\beta = 50$ %		$\beta = 60$ %		$\beta = 70$ %		$\beta = 80$ %	
		Coverage	Length	Coverage	Length	Coverage	Length	Coverage	Length
500	NA1	0.903	0.067	0.896	0.076	0.930	0.078	0.930	0.087
	NA2	0.930	0.061	0.963	0.067	0.953	0.072	0.950	0.073
	BT1	0.940	0.060	0.957	0.067	0.950	0.072	0.940	0.073
	BT2	0.943	0.060	0.963	0.067	0.956	0.072	0.949	0.073
	BCa	0.849	0.059	0.842	0.066	0.832	0.071	0.805	0.070
	EL	0.961	0.068	0.932	0.076	0.904	0.073	0.942	0.052
	SJEL	0.947	0.061	0.948	0.068	0.941	0.073	0.944	0.074
800	NA1	0.897	0.045	0.890	0.054	0.901	0.048	0.881	0.068
	NA2	0.963	0.049	0.956	0.054	0.953	0.058	0.950	0.059
	BT1	0.952	0.048	0.952	0.054	0.948	0.058	0.939	0.060
	BT2	0.954	0.048	0.953	0.054	0.954	0.058	0.946	0.060
	BCa	0.882	0.048	0.860	0.053	0.876	0.057	0.813	0.058
	EL	0.942	0.050	0.952	0.056	0.946	0.060	0.956	0.060
	SJEL	0.949	0.049	0.949	0.055	0.951	0.059	0.946	0.060
1000	NA1	0.883	0.040	0.932	0.040	0.929	0.041	0.929	0.045
	NA2	0.940	0.043	0.943	0.048	0.940	0.052	0.923	0.053
	BT1	0.954	0.043	0.952	0.042	0.942	0.042	0.956	0.044
	BT2	0.958	0.043	0.958	0.042	0.950	0.042	0.956	0.044
	BCa	0.888	0.043	0.882	0.048	0.852	0.052	0.850	0.052
	EL	0.930	0.046	0.940	0.046	0.948	0.041	0.955	0.050
	SJEL	0.949	0.045	0.945	0.038	0.950	0.040	0.954	0.041

of employee, a homogeneous income group with relatively evenly distributed income is thus created for all faculty positions of universities and colleges in Georgia. In our analysis, we limit the income data to all titles with Professor, Associate Professor and Assistant Professor from Units of University System and Georgia Military College from 2012 fiscal year. There are 10,332 individuals obtained initially. However, we observe some records having abnormally low salary, and we infer that those types of professors are not working full time during 2012. It may be caused by several reasons. First of all, there are some newly hired professors in 2012, who did not work for the whole 2012 fiscal year. After dropping those professors who did not have salary record in 2011, 9229 observations are kept. Second of all, some professors may not be in full-time service in 2012, who may possibly either take leave or transfer to another organization. We filter this type of records out by dropping those whose income in 2012 is far less than that in 2011. Therefore, 6195 observations are kept. Then, we drop the part-time professors whose salary is less than $20,000. Finally, there remain 5921 observations in the analysis. By taking above steps, we create a relatively homogeneous income group by retaining professors who are more likely to provide full-time service during 2012 fiscal year. All the real example analyses are based on these 5921 observations.

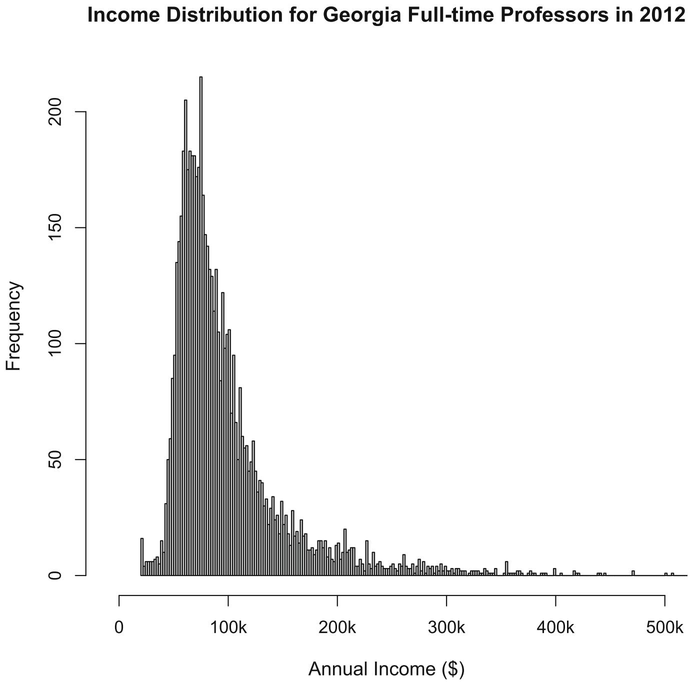

Fig. 3 Income distribution for Georgia Professors in 2012

We plot the histogram of 2012 annual salary for these professors in Fig. 3. It is observed that the income data are highly right skewed. Next, we present some basic statistics of annual income by job title. There are 2266 Assistant Professors recorded in 2012 with median salary \$68,618 and mean salary \$81,055. The number of recorded Associate Professors in Georgia is 1891. Their median salary is \$80,675 and mean salary is \$91,760. While for the 1764 Professors, the median salary is \$109,044 and mean salary is \$129,640. The maximum recorded salary for Assistant Professors, Associate Professors and Professors is \$507,500, \$633,260, and \$949,419, respectively.

To evaluate annual income by school, out of the 5921 observations, University of Georgia (UGA), Georgia State University (GSU), Georgia Institute of Technology (GIT) and Georgia Health Sciences University (GHSU) are the top 4 universities that have the largest number of recorded professors. According to the low-wage definition (i.e., $\alpha = \frac{2}{3}, \beta = 0.5$) by Eurostat 2012, we calculated low-income lines, empirical estimates and kernel estimates of low-income proportions for UGA, GSU, GIT and GHSU in Table 6. It is observed that GSU has smaller proportion of low-wage earners compared with the other three Georgia public universities.

To obtain an interval estimate for the low-income proportion of Georgia professors, we fix the fraction α at $\frac{2}{3}$, and choose $\beta = 0.5$ based on the low-wage definition by Eurostat 2012. At 95 % confidence level, there are 9.0 to 10.8 % of Georgia professors in 2012 who can be categorized as low-wage earners based on our recommended SJEL interval. This finding will provide meaningful information for the government.

Table 6 Empirical estimate $\hat{\theta}_{\alpha\beta}$ and kernel estimate $\hat{T}_n(\alpha, \beta)$ for LIP, and the estimated low-income line $\alpha\xi_\beta$ in 2012 for four Georgia research universities

Organization	Empirical estimate (%)	Kernel estimate (%)	Low-income line
UGA	8.16	8.14	\$62,987
GSU	7.25	7.34	\$62,188
GIT	13.52	13.52	\$88,030
GHSU	22.73	22.59	\$91,731

6 Discussion

Development of accurate and robust inferences for low-income proportions is increasingly important. In this paper, we have proposed a kernel estimator for a low-income proportion, and obtained the asymptotic normality of the kernel estimator. Later, the jackknife empirical likelihood for a low-income proportion is defined, and the log-jackknife empirical likelihood ratio statistic is proved to be asymptotically a standard Chi-square distribution. We applied a cross-validation method for bandwidth selection in our study, and there are still many other interesting bandwidth selection methods. For example, Bowman et al. (1998) selected bandwidth by introducing a cross-validation method for the smoothing estimation of distribution function. Our extensive simulation studies have showed that the proposed smoothed estimator has smaller MSE than the traditional empirical estimator for the low-income proportion. Simulation studies also indicate that the proposed smoothed jackknife empirical likelihood-based (SJEL) interval performs better than other intervals considered in this paper, which is not surprising because the proposed SJEL method combines the power of both jackknife and empirical likelihood. While the proposed bootstrap-based confidence intervals (BT1 and BT2) have good finite sample performances, they are computationally expensive, particularly when sample size is getting larger. Compared with the existing empirical likelihood-based intervals (EL) proposed by Yang et al. (2011), the SJEL interval bypasses the estimation of the scale parameter. It can be directly calculated by implementing the algorithm for computing the standard empirical likelihood interval (Hall and La Scala 1990). Based on this study, we recommend the use of the proposed SJEL confidence interval for a low-income proportion.

Appendix: Proof of theorems

The Proof of Theorem 2.1 We have the following decomposition

$$\sqrt{n}\{\hat{T}_n(\alpha, \beta) - \theta_{\alpha\beta}\} = \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} K\left(\frac{\alpha\hat{\xi}_\beta - X_i}{h}\right) - \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{\alpha\xi_\beta - X_i}{h}\right)\right]$$
$$+\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} K\left(\frac{\alpha\xi_\beta - X_i}{h}\right) - F(\alpha\xi_\beta)\right] \equiv I_1 + I_2. \quad (5)$$

I_1 from (5) can be written as

$$I_1 = \int_{-\infty}^{\infty}\left[K\left(\frac{\alpha\hat{\xi}_\beta - x}{h}\right) - K\left(\frac{\alpha\xi_\beta - x}{h}\right)\right]\mathrm{d}\left[\sqrt{n}(F_n(x) - F(x))\right]$$
$$+\sqrt{n}\int_{-\infty}^{\infty}\left[K\left(\frac{\alpha\hat{\xi}_\beta - x}{h}\right) - K\left(\frac{\alpha\xi_\beta - x}{h}\right)\right]\mathrm{d}F(x) \equiv I_{11} + I_{12}. \quad (6)$$

Then using Taylor expansion and the Bahadur's representation for sample quantile (Bahadur 1966)

$$\hat{\xi}_\beta - \xi_\beta = \frac{\beta - \frac{1}{n}\sum_{i=1}^n I(X_i \le \xi_\beta)}{f(\xi_\beta)} + o_p(n^{-\frac{1}{2}}),$$

I_{12} from (6) can be written as

$$\begin{aligned} I_{12} &= \sqrt{n}\int_{-\infty}^{\infty}\left[K\left(\frac{\alpha\hat{\xi}_\beta - x}{h}\right) - K\left(\frac{\alpha\xi_\beta - x}{h}\right)\right]\mathrm{d}F(x) \\ &= \sqrt{n}\int_{-\infty}^{\infty}\omega\left(\frac{\alpha\xi_\beta - x}{h}\right)\frac{\alpha\hat{\xi}_\beta - \alpha\xi_\beta}{h}\,\mathrm{d}F(x) + o_p(1) \\ &= -\sqrt{n}\int_{-\infty}^{\infty}\omega\left(\frac{\alpha\xi_\beta - x}{h}\right)\frac{\alpha}{h}\,\frac{\frac{1}{n}\sum_{i=1}^n I(X_i \le \xi_\beta) - \beta}{f(\xi_\beta)}\,\mathrm{d}F(x) + o_p(1) \\ &= -\frac{\alpha f(\alpha\xi_\beta)U_n(\beta)}{f(\xi_\beta)} + o_p(1), \end{aligned} \quad (7)$$

where $U_n(\beta) = \sqrt{n}[\frac{1}{n}\sum_{i=1}^n I(X_i \le \xi_\beta) - \beta] = \sqrt{n}[\frac{1}{n}\sum_{i=1}^n I(F(X_i) \le \beta) - \beta]$.

Since $\sqrt{n}[F_n(x) - F(x)] \to B(x)$, which is a Gaussian process, $\sqrt{n}(\hat{\xi}_\beta - \xi_\beta) = O_p(1)$, and $\sqrt{n}h \to \infty$, we get $I_{11} = o_p(1)$. Therefore, $I_1 = -\frac{\alpha f(\alpha\xi_\beta)U_n(\beta)}{f(\xi_\beta)} + o_p(1)$.

Next, let us consider I_2 from (5). We are going to prove

$$EK\left(\frac{\alpha\xi_\beta - X}{h}\right) \longrightarrow \theta_{\alpha\beta}, \quad \text{and} \quad EK^2\left(\frac{\alpha\xi_\beta - X}{h}\right) \longrightarrow \theta_{\alpha\beta}, \qquad \text{as } h \to 0.$$

Notice that

$$\begin{aligned} \lim_{h\to 0} EK\left(\frac{\alpha\xi_\beta - X}{h}\right) &= \lim_{h\to 0}\int_{-\infty}^{\infty}K\left(\frac{\alpha\xi_\beta - x}{h}\right)f(x)\,\mathrm{d}x \\ &= \lim_{h\to 0}\int_{-\infty}^{\infty}\int_{-\infty}^{\frac{\alpha\xi_\beta - x}{h}}\omega(y)\,\mathrm{d}y f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\left[\lim_{h\to 0}\int_{-\infty}^{\frac{\alpha\xi_\beta - x}{h}}\omega(y)\,\mathrm{d}y\right]f(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\{0 * I[\alpha\xi_\beta < x] + \int_{-\infty}^{0}\omega(y)\,\mathrm{d}y * I[\alpha\xi_\beta = x] + 1 * I[\alpha\xi_\beta > x]\}f(x)\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} I[\alpha\xi_\beta > x] f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} I[F(x) < F(\alpha\xi_\beta)]\,\mathrm{d}F(x) \\
&= F(\alpha\xi_\beta) = \theta_{\alpha\beta}. \qquad (8)
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&\lim_{h\to 0} EK^2\left(\frac{\alpha\xi_\beta - X}{h}\right) = \lim_{h\to 0}\int_{-\infty}^{\infty} K^2\left(\frac{\alpha\xi_\beta - x}{h}\right) f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \{0 * I[\alpha\xi_\beta < x] + \int_{-\infty}^{0} \omega(y)\mathrm{d}y * I[\alpha\xi_\beta = x] + 1 * I[\alpha\xi_\beta > x]\}^2 f(x)\,\mathrm{d}x \\
&= F(\alpha\xi_\beta) = \theta_{\alpha\beta}. \qquad (9)
\end{aligned}
$$

Let $W_i = K(\frac{\alpha\xi_\beta - X_i}{h})$. So I_2 from (5) can be rewritten as

$$
\begin{aligned}
I_2 &= \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} K\left(\frac{\alpha\xi_\beta - X_i}{h}\right) - EK\left(\frac{\alpha\xi_\beta - X}{h}\right)\right] + o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(W_i - EW_i) + o_p(1). \qquad (10)
\end{aligned}
$$

Let $\{U_1, U_2, \ldots, U_n\}$ be an i.i.d. sample from $U(0,1)$ (uniform distribution on $[0,1]$) and independent of $\{X_1, X_2, \ldots, X_n\}$. Since $U_i \overset{d}{=} F(X_i) \overset{\text{i.i.d.}}{\sim} U(0,1)$ for any continuous distribution function F, then

$$
\begin{aligned}
I_1 &= -\frac{\alpha f(\alpha\xi_\beta)}{f(\xi_\beta)}\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} I(F(X_i) \le \beta) - \beta\right] + o_p(1) \\
&\overset{d}{=} -\frac{\alpha f(\alpha\xi_\beta)}{f(\xi_\beta)}\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} I(U_i \le \beta) - \beta\right] + o_p(1), \qquad (11)
\end{aligned}
$$

where $\overset{d}{=}$ means that two statistics asymptotically have the same distribution. Therefore,

$$
I_1 + I_2 \overset{d}{=} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[-\frac{\alpha f(\alpha\xi_\beta)}{f(\xi_\beta)}(I(U_i \le \beta) - \beta) + (W_i - EW_i)\right] + o_p(1). \qquad (12)
$$

Since

$$\begin{aligned}
&\lim_{h\to 0} E[(I(U_i \le \beta) - \beta)(W_i - EW_i)] \\
&\quad = \int_{-\infty}^{\infty} [I(F(x) \le \beta) - \beta] \left[\lim_{h\to 0} K\left(\frac{\alpha\xi_\beta - x}{h}\right) - \lim_{h\to 0} EW_i\right] f(x)\,\mathrm{d}x \\
&\quad = \int_{-\infty}^{\infty} [I(F(x) \le \beta) - \beta][I[\alpha\xi_\beta > x] - \theta_{\alpha\beta}] f(x)\,\mathrm{d}x \\
&\quad = \theta_{\alpha\beta}(1-\beta),
\end{aligned}$$

$\mathrm{Var}[I(U_i \le \beta) - \beta] = \beta(1-\beta)$, $\mathrm{Var}(W_i) = EK^2(\frac{\alpha\xi_\beta - x}{h}) - [EK(\frac{\alpha\xi_\beta - x}{h})]^2 = \theta_{\alpha\beta}(1-\theta_{\alpha\beta}) + o(1)$, by the central limit theorem applicable to a triangular array setting (see Shao 2003), we get that

$$I_1 + I_2 \xrightarrow{d} N(0, \sigma_{\alpha\beta}^2), \tag{13}$$

where $\sigma_{\alpha\beta}^2 = \frac{\alpha^2\beta(1-\beta)f^2(\alpha\xi_\beta)}{f^2(\xi_\beta)} - 2\alpha(1-\beta)\theta_{\alpha\beta}\frac{f(\alpha\xi_\beta)}{f(\xi_\beta)} + \theta_{\alpha\beta}(1-\theta_{\alpha\beta})$. The proof of Theorem 2.1 is complete.

We need Lemmas 1 and 2 to prove Theorem 3.1.

Lemma 1 *Under the conditions in Theorem 2.1, we have*

$$\sqrt{n}\left\{\frac{1}{n}\sum_{k=1}^{n} \hat{V}_k(\alpha,\beta) - \theta_{\alpha\beta}\right\} \xrightarrow{d} N(0, \sigma_{\alpha\beta}^2), \tag{14}$$

where $\sigma_{\alpha\beta}^2$ is defined in Theorem 2.1.

Proof Note that $\frac{1}{n}\sum_{k=1}^{n} \hat{V}_k(\alpha,\beta)$ can be decomposed into

$$\frac{1}{n}\sum_{k=1}^{n} \hat{V}_k(\alpha,\beta) = \frac{n-1}{n}\sum_{k=1}^{n}[\hat{T}_n(\alpha,\beta) - \hat{T}_{n-1,k}(\alpha,\beta)] + \hat{T}_n(\alpha,\beta), \tag{15}$$

while

$$\begin{aligned}
&\hat{T}_n(\alpha,\beta) - \hat{T}_{n-1,k}(\alpha,\beta) \\
&\quad = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{\alpha\hat{\xi}_\beta - X_i}{h}\right) - \frac{1}{n-1}\sum_{j\ne k}^{n} K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_j}{h}\right) \\
&\quad = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{\alpha\hat{\xi}_\beta - X_i}{h}\right) - \frac{1}{n}\sum_{j\ne k}^{n} K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_j}{h}\right) \\
&\qquad + \frac{1}{n}\sum_{j\ne k}^{n} K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_j}{h}\right) - \frac{1}{n-1}\sum_{j\ne k}^{n} K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_j}{h}\right).
\end{aligned} \tag{16}$$

So

$$
\begin{aligned}
&\sum_{k=1}^{n}(\hat{T}_n(\alpha,\beta)-\hat{T}_{n-1,k}(\alpha,\beta))\\
&\quad=\left\{\frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)-K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_i}{h}\right)\right]\right\}\\
&\qquad+\left\{\frac{1}{n}\sum_{k=1}^{n}\sum_{j=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_j}{h}\right)\right]-\frac{1}{n-1}\sum_{k=1}^{n}\sum_{j=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_j}{h}\right)\right]\right.\\
&\qquad\left.+\frac{1}{n-1}\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)\right\}\equiv I_1+I_2. \qquad (17)
\end{aligned}
$$

Using the Bahadur representation for sample quantile (Bahadur 1966), we get that

$$
\begin{aligned}
\hat{\xi}_{\beta,-k}-\hat{\xi}_\beta &= (\hat{\xi}_{\beta,-k}-\xi_\beta)-(\hat{\xi}_\beta-\xi_\beta)\\
&=\left[\frac{\beta-\frac{1}{n-1}\sum_{j\neq k}^{n}I(X_j\le\xi_\beta)}{f(\xi_\beta)}\right]-\left[\frac{\beta-\frac{1}{n}\sum_{i=1}^{n}I(X_i\le\xi_\beta)}{f(\xi_\beta)}\right]\\
&\quad+o_p(n^{-1/2})\\
&=\frac{\frac{1}{n-1}[I(X_k\le\xi_\beta)-F_n(\xi_\beta)]}{f(\xi_\beta)}+o_p(n^{-1/2}), \qquad (18)
\end{aligned}
$$

and $(\frac{\alpha\hat{\xi}_{\beta,-k}-\alpha\hat{\xi}_\beta}{h})^2=O_p(\frac{1}{n^2h^2})$. Under conditions of Theorem 2.1, using Taylor expansion, we get that

$$
\begin{aligned}
I_1&=\frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)-K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_i}{h}\right)\right]\\
&=\frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{n}\left[-\omega\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)\frac{\alpha\hat{\xi}_{\beta,-k}-\alpha\hat{\xi}_\beta}{h}\right.\\
&\quad\left.-\frac{1}{2}\omega'\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)\left(\frac{\alpha\hat{\xi}_{\beta,-k}-\alpha\hat{\xi}_\beta}{h}\right)^2\right]+o_p\left(\frac{1}{nh}\right)\\
&=\frac{1}{n}\sum_{i=1}^{n}\left\{-\omega\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)\sum_{k=1}^{n}\frac{\alpha\hat{\xi}_{\beta,-k}-\alpha\hat{\xi}_\beta}{h}\right.\\
&\quad\left.-\frac{1}{2}\sum_{k=1}^{n}\omega'\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)\left(\frac{\alpha\hat{\xi}_{\beta,-k}-\alpha\hat{\xi}_\beta}{h}\right)^2\right\}+o_p\left(\frac{1}{nh}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{1}{2}\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{n}\omega^{'}\left(\frac{\alpha\hat{\xi}_\beta - X_i}{h}\right)\left(\frac{\alpha\hat{\xi}_{\beta,-k} - \alpha\hat{\xi}_\beta}{h}\right)^2 + o_p(n^{-1/2}) + o_p\left(\frac{1}{nh}\right)\\
&= O_p\left(\frac{1}{nh^2}\right)\int_{-\infty}^{\infty}\omega^{'}\left(\frac{\alpha\hat{\xi}_\beta - x}{h}\right)\mathrm{d}F_n(x) + o_p(n^{-1/2}) + o_p\left(\frac{1}{nh}\right)\\
&= O_p\left(\frac{1}{nh^2}\right)\int_{-\infty}^{\infty}\omega^{'}\left(\frac{\alpha\hat{\xi}_\beta - x}{h}\right)\mathrm{d}F(x) + o_p(n^{-1/2}) + o_p\left(\frac{1}{nh}\right)\\
&= O_p\left(\frac{1}{nh}\right)\int_{-\infty}^{\infty}\omega^{'}(y)f(\alpha\hat{\xi}_\beta - yh)\mathrm{d}y + o_p(n^{-1/2}) + o_p\left(\frac{1}{nh}\right)\\
&= O_p\left(\frac{1}{nh}\right) + o_p(n^{-1/2}).
\end{aligned}
\tag{19}
$$

Meanwhile, I_2 from (17) can be written to

$$
\begin{aligned}
I_2 &= \sum_{k=1}^{n}\left[\left(\frac{1}{n} - \frac{1}{n-1}\right)\sum_{j=1}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_j}{h}\right) + \frac{1}{n-1}K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_k}{h}\right)\right]\\
&= \frac{-1}{n(n-1)}\sum_{k=1}^{n}\sum_{j=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_j}{h}\right) - K\left(\frac{\alpha\hat{\xi}_\beta - X_j}{h}\right)\right]\\
&\quad + \frac{-1}{n(n-1)}\sum_{k=1}^{n}\sum_{j=1}^{n}K\left(\frac{\alpha\hat{\xi}_\beta - X_j}{h}\right)\\
&\quad - \frac{-1}{n-1}\sum_{k=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_k}{h}\right) - K\left(\frac{\alpha\hat{\xi}_\beta - X_k}{h}\right)\right]\\
&\quad - \frac{-1}{n-1}\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_\beta - X_k}{h}\right)\\
&= O_p\left(\frac{1}{n(n-1)h}\right) - \frac{1}{n(n-1)}\sum_{k=1}^{n}\sum_{j=1}^{n}K\left(\frac{\alpha\hat{\xi}_\beta - X_j}{h}\right)\\
&\quad - O_p\left(\frac{1}{(n-1)^2h}\right) + \frac{1}{n-1}\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_\beta - X_k}{h}\right)\\
&= O_p\left(\frac{1}{n^2h}\right).
\end{aligned}
\tag{20}
$$

From (19) and (20), we get $I_1 + I_2 = O_p(\frac{1}{nh}) + o_p(n^{-1/2})$, which implies that

$$\frac{1}{n}\sum_{k=1}^{n}\hat{V}_k(\alpha,\beta) = \frac{n-1}{n}\sum_{k=1}^{n}[\hat{T}_n(\alpha,\beta) - \hat{T}_{n-1,k}(\alpha,\beta)] + \hat{T}_n(\alpha,\beta)$$
$$= \hat{T}_n(\alpha,\beta) + O_p\left(\frac{1}{nh}\right) + o_p(n^{-1/2}). \tag{21}$$

Therefore,

$$\sqrt{n}\left\{\frac{1}{n}\sum_{k=1}^{n}\hat{V}_k(\alpha,\beta) - \theta_{\alpha\beta}\right\} = \sqrt{n}[\hat{T}_n(\alpha,\beta) - \theta_{\alpha\beta}]$$
$$+ O_p\left(\frac{1}{\sqrt{n}h}\right) + o_p(1) \xrightarrow{d} N(0, \sigma^2_{\alpha\beta}).$$

Lemma 2 *Under the conditions in Theorem 2.1, we have that*

$$\frac{1}{n}\sum_{k=1}^{n}\{\hat{V}_k(\alpha,\beta) - \theta_{\alpha\beta}\}^2 \xrightarrow{p} \sigma^2_{\alpha\beta}. \tag{22}$$

Proof From Lemma 1, it follows that

$$\frac{1}{n}\sum_{k=1}^{n}\{\hat{V}_k(\alpha,\beta) - \theta_{\alpha\beta}\}^2 = \frac{1}{n}\sum_{k=1}^{n}\hat{V}_k^2(\alpha,\beta) - 2\theta_{\alpha\beta}\frac{1}{n}\sum_{k=1}^{n}\hat{V}_k(\alpha,\beta) + \frac{1}{n}\sum_{k=1}^{n}\theta^2_{\alpha\beta}$$
$$\xrightarrow{p} \frac{1}{n}\sum_{k=1}^{n}\hat{V}_k^2(\alpha,\beta) - 2\theta_{\alpha\beta}\theta_{\alpha\beta} + \frac{1}{n}n\theta^2_{\alpha\beta} = \frac{1}{n}\sum_{k=1}^{n}\hat{V}_k^2(\alpha,\beta) - \theta^2_{\alpha\beta}. \tag{23}$$

By the definition of the jackknife pseudo-values for the low-income proportion, we have that

$$\hat{V}_k(\alpha,\beta) = n\hat{T}_n(\alpha,\beta) - (n-1)\hat{T}_{n-1,k}(\alpha,\beta),$$
$$= \sum_{i=1}^{n}K\left(\frac{\alpha\hat{\xi}_\beta - X_i}{h}\right) - \sum_{j\neq k}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_j}{h}\right)$$
$$= \sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta - X_i}{h}\right) - K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_i}{h}\right)\right] + K\left(\frac{\alpha\hat{\xi}_{\beta,-k} - X_k}{h}\right), \tag{24}$$

and

$$\hat{V}_k^2(\alpha,\beta)=\left\{\sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)-K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_i}{h}\right)\right]\right\}^2+K^2\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)$$
$$+2K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)\sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)-K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_i}{h}\right)\right]. \tag{25}$$

Therefore,

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}\hat{V}_k^2(\alpha,\beta)&=\frac{1}{n}\sum_{k=1}^{n}\left\{\sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)-K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_i}{h}\right)\right]\right\}^2\\&\quad+\frac{1}{n}\sum_{k=1}^{n}K^2\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)\\&\quad+\frac{2}{n}\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)\\&\quad\times\sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)-K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_i}{h}\right)\right]\\&\equiv J_1+J_2+J_3.\end{aligned} \tag{26}$$

Using Taylor expansion and the Bahadur representation for sample quantile, the first term J_1 in (26) can be written as

$$\begin{aligned}J_1&=\frac{1}{n}\sum_{k=1}^{n}\left\{\sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)-K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_i}{h}\right)\right]\right\}^2\\&=\frac{1}{n}\sum_{k=1}^{n}\left\{\sum_{i=1}^{n}\omega\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)\frac{\alpha\hat{\xi}_{\beta,-k}-\alpha\hat{\xi}_\beta}{h}\right\}^2+o_p(1)\\&=\frac{1}{n}\sum_{k=1}^{n}\left(\frac{\alpha\hat{\xi}_{\beta,-k}-\alpha\hat{\xi}_\beta}{h}\right)^2\left\{\sum_{i=1}^{n}\omega\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)\right\}^2+o_p(1)\\&=\frac{1}{n}\frac{\alpha^2}{h^2}\sum_{k=1}^{n}(\hat{\xi}_{\beta,-k}-\hat{\xi}_\beta)^2\left\{n\int_{-\infty}^{\infty}\omega\left(\frac{\alpha\hat{\xi}_\beta-x}{h}\right)\mathrm{d}F_n(x)\right\}^2+o_p(1)\\&=\frac{\alpha^2}{nh^2}\sum_{k=1}^{n}(\hat{\xi}_{\beta,-k}-\hat{\xi}_\beta)^2\left\{n\int_{-\infty}^{\infty}\omega\left(\frac{\alpha\hat{\xi}_\beta-x}{h}\right)\mathrm{d}F(x)\right\}^2+o_p(1)\\&=\frac{n\alpha^2}{h^2}\sum_{k=1}^{n}(\hat{\xi}_{\beta,-k}-\hat{\xi}_\beta)^2\left[\int_{-\infty}^{\infty}\omega(z)\mathrm{d}F(\alpha\xi_\beta-zh)\right]^2+o_p(1)\end{aligned}$$

$$
\begin{aligned}
&= \frac{n\alpha^2}{f^2(\xi_\beta)h^2}\sum_{k=1}^{n}\frac{1}{(n-1)^2}\Bigg[I(x_k\le\xi_\beta)I(x_k\le\xi_\beta)-2I(x_k\le\xi_\beta)\frac{1}{n}\sum_{i=1}^{n}I(X_i\le\xi_\beta)\\
&\quad+\frac{1}{n^2}\sum_{i=1}^{n}I(X_i\le\xi_\beta)\sum_{i=1}^{n}I(X_i\le\xi_\beta)\Bigg]\left[\int_{-\infty}^{\infty}\omega(z)\mathrm{d}F(\alpha\xi_\beta-zh)\right]^2+o_p(1)\\
&= \frac{n^2\alpha^2(-h)^2}{f^2(\xi_\beta)h^2(n-1)^2}[F_n(\xi_\beta)-F_n^2(\xi_\beta)]\left[\int_{-\infty}^{\infty}\omega(z)f(\alpha\xi_\beta-zh)\,\mathrm{d}z\right]^2+o_p(1)\\
&\xrightarrow{p} \frac{\alpha^2\beta(1-\beta)f^2(\alpha\xi_\beta)}{f^2(\xi_\beta)}. \qquad (27)
\end{aligned}
$$

J_2 from (26) can be written as

$$
\begin{aligned}
J_2 &= \frac{1}{n}\sum_{k=1}^{n}K^2\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)\\
&= \frac{1}{n}\sum_{k=1}^{n}\left[K^2\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)-K^2\left(\frac{\alpha\xi_\beta-X_k}{h}\right)\right]+\frac{1}{n}\sum_{k=1}^{n}K^2\left(\frac{\alpha\xi_\beta-X_k}{h}\right)\\
&= \frac{1}{n}\sum_{k=1}^{n}K^2\left(\frac{\alpha\xi_\beta-X_k}{h}\right)+o_p(1)\\
&= EK^2\left(\frac{\alpha\xi_\beta-x}{h}\right)+o_p(1)=\theta_{\alpha\beta}+o_p(1). \qquad (28)
\end{aligned}
$$

The term J_3 from (26) can be written as

$$
\begin{aligned}
J_3 &= \frac{2}{n}\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)\sum_{i=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)-K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_i}{h}\right)\right]\\
&= \frac{-2}{n}\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)\left(\frac{\alpha\hat{\xi}_{\beta,-k}-\alpha\hat{\xi}_\beta}{h}\right)\sum_{i=1}^{n}\omega\left(\frac{\alpha\hat{\xi}_\beta-X_i}{h}\right)+o_p(1)\\
&= \frac{-2\alpha}{nh}\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)(\hat{\xi}_{\beta,-k}-\hat{\xi}_\beta)n\int_{-\infty}^{\infty}\omega\left(\frac{\alpha\hat{\xi}_\beta-x}{h}\right)\mathrm{d}F(x)+o_p(1)\\
&= -2\alpha\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)(\hat{\xi}_{\beta,-k}-\hat{\xi}_\beta)\int_{-\infty}^{\infty}\omega(z)f(\alpha\xi_\beta-zh)\,\mathrm{d}z+o_p(1)\\
&= -2\alpha\sum_{k=1}^{n}K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)\frac{[I(X_k\le\xi_\beta)-F_n(\xi_\beta)]}{(n-1)f(\xi_\beta)}f(\alpha\xi_\beta)+o_p(1)\\
&= \frac{-2\alpha f(\alpha\xi_\beta)}{f(\xi_\beta)}\frac{1}{n-1}\sum_{k=1}^{n}\left[K\left(\frac{\alpha\hat{\xi}_{\beta,-k}-X_k}{h}\right)-K\left(\frac{\alpha\hat{\xi}_\beta-X_k}{h}\right)\right.
\end{aligned}
$$

$$
\begin{aligned}
&+K\left(\frac{\alpha\hat{\hat{\xi}}_\beta - X_k}{h}\right)\Bigg][I(X_k \le \xi_\beta) - F_n(\xi_\beta)] + o_p(1) \\
&= \frac{-2\alpha f(\alpha\xi_\beta)}{f(\xi_\beta)}\left\{\frac{1}{n-1}\sum_{k=1}^{n}\left[-\omega\left(\frac{\alpha\hat{\hat{\xi}}_\beta - X_k}{h}\right)\frac{\alpha\hat{\hat{\xi}}_{\beta,-k} - \alpha\hat{\hat{\xi}}_\beta}{h}\right]\right. \\
&\quad\times [I(X_k \le \xi_\beta) - F_n(\xi_\beta)] \\
&\quad\left.+\frac{1}{n-1}\sum_{k=1}^{n} K\left(\frac{\alpha\hat{\hat{\xi}}_\beta - X_k}{h}\right)[I(X_k \le \xi_\beta) - F_n(\xi_\beta)]\right\} + o_p(1) \\
&\equiv \frac{-2\alpha f(\alpha\xi_\beta)}{f(\xi_\beta)}\{M_1 + M_2\} + o_p(1), \qquad (29)
\end{aligned}
$$

and

$$
\begin{aligned}
M_1 &= \frac{1}{n-1}\sum_{k=1}^{n}\left[-\omega\left(\frac{\alpha\hat{\hat{\xi}}_\beta - X_k}{h}\right)\frac{\alpha\hat{\hat{\xi}}_{\beta,-k} - \alpha\hat{\hat{\xi}}_\beta}{h}\right][I(X_k \le \xi_\beta) - F_n(\xi_\beta)] \\
&= \frac{-\alpha}{(n-1)h}\sum_{k=1}^{n}\omega\left(\frac{\alpha\hat{\hat{\xi}}_\beta - X_k}{h}\right)\frac{[I(X_k \le \xi_\beta) - F_n(\xi_\beta)]^2}{(n-1)f(\xi_\beta)} + o_p(1) \\
&= \frac{-\alpha}{(n-1)^2 h f(\xi_\beta)}\sum_{k=1}^{n}\omega\left(\frac{\alpha\hat{\hat{\xi}}_\beta - X_k}{h}\right)[I(X_k \le \xi_\beta) - F_n(\xi_\beta)]^2 + o_p(1) \\
&= \frac{-n\alpha}{(n-1)^2 h f(\xi_\beta)}\int_{-\infty}^{\infty}\omega\left(\frac{\alpha\hat{\hat{\xi}}_\beta - x}{h}\right)[I(F(x) \le \beta) - \beta]^2\,\mathrm{d}F(x) + o_p(1) \\
&= o_p(1), \qquad (30) \\
M_2 &= \frac{1}{n-1}\sum_{k=1}^{n} K\left(\frac{\alpha\hat{\hat{\xi}}_\beta - X_k}{h}\right)[I(X_k \le \xi_\beta) - F_n(\xi_\beta)] \\
&= \frac{n}{n-1}\int_{-\infty}^{\infty} K\left(\frac{\alpha\xi_\beta - x}{h}\right)[I(F(x) \le \beta) - \beta]\,\mathrm{d}F(x) + o_p(1) \\
&= \frac{n}{n-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\frac{\alpha\xi_\beta - x}{h}}\omega(y)\,\mathrm{d}y[I(F(x) \le \beta) - \beta]\mathrm{d}F(x) + o_p(1) \\
&= \int_{-\infty}^{\infty} I(F(x) \le F(\alpha\xi_\beta))[I(F(x) \le \beta) - \beta]\,\mathrm{d}F(x) + o_p(1) \\
&= \theta_{\alpha\beta}(1-\beta) + o_p(1). \qquad (31)
\end{aligned}
$$

From (29), (30), and (31), we get $J_3 = -2\alpha(1-\beta)\theta_{\alpha\beta}\frac{f(\alpha\xi_\beta)}{f(\xi_\beta)} + o_p(1)$. Therefore,

$$
\frac{1}{n}\sum_{k=1}^{n}\hat{V}_k^2(\alpha,\beta) \xrightarrow{p} \frac{\alpha^2\beta(1-\beta)f^2(\alpha\xi_\beta)}{f^2(\xi_\beta)} + \theta_{\alpha\beta} - 2\alpha(1-\beta)\theta_{\alpha\beta}\frac{f(\alpha\xi_\beta)}{f(\xi_\beta)}. \qquad (32)
$$

In sum, we have that

$$\frac{1}{n}\sum_{k=1}^{n}\{\hat{V}_k(\alpha,\beta)-\theta_{\alpha\beta}\}^2 \xrightarrow{p} \sigma^2_{\alpha\beta}.$$

The Proof of Theorem 3.1 It follows immediately from Lemmas 1 and 2.

The Proof of Theorem 3.2 Let $g(\lambda)=\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta}}{1+\lambda(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})}$. It is easy to check that

$$\begin{aligned}
0=|g(\lambda)| &= \frac{1}{n}\left|\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})-\lambda\sum_{i=1}^{n}\frac{(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})^2}{1+\lambda(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})}\right| \\
&\geq \left|\frac{\lambda}{n}\sum_{i=1}^{n}\frac{(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})^2}{1+\lambda(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})}\right|-\left|\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta}\right| \\
&\geq \frac{|\lambda|S_n}{1+|\lambda|Z_n}-\left|\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})\right|,
\end{aligned} \tag{33}$$

where $S_n=\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})^2$ and $Z_n=\max_{1\leq i\leq n}|\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta}|$.

From Lemmas 1 and 2, we have $|\lambda|=O_p(n^{-\frac{1}{2}})$. Put $\gamma_i=\lambda(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})$, then we have $\max_{1\leq i\leq n}|\gamma_i|=o_p(1)$, and

$$\begin{aligned}
0=g(\lambda) &= \frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})\left(1-\gamma_i+\frac{\gamma_i^2}{1+\gamma_i}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})-S_n\lambda+\frac{\lambda^2}{n}\sum_{i=1}^{n}\frac{(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})^3}{1+\gamma_i} \\
&= \frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})-S_n\lambda+o_p(n^{-1/2}),
\end{aligned} \tag{34}$$

which implies that $\lambda=S_n^{-1}\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})+\beta_n$, where $\beta_n=o_p(n^{-1/2})$.

Therefore,

$$\begin{aligned}
l_n(\theta_{\alpha\beta}) &= 2\sum_{i=1}^{n}\log\{1+\lambda(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})\} \\
&= 2\sum_{i=1}^{n}\gamma_i-\sum_{i=1}^{n}\gamma_i^2+o_p(1) \\
&= 2n\lambda\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})-nS_n\lambda^2+o_p(1)
\end{aligned}$$

$$= \frac{n\{\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})\}^2}{S_n} - nS_n\beta_n^2 + o_p(1)$$
$$= \frac{n\{\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i(\alpha,\beta)-\theta_{\alpha\beta})\}^2}{S_n} + o_p(1) \xrightarrow{d} \chi^2(1). \quad (35)$$

References

Bahadur, R. (1966). A note on quantiles in large samples. *Annals of Mathematical Statistics*, *37*, 577–580.

Bezzina, E. (2012). In 2010, 17% of Employees in the EU were Low-Wage Earners. Eurostat Statistics in Focus. Issue number 48. http://ec.europa.eu/eurostat/web/products-statistics-in-focus/-/KS-SF-12-048.

Bowman, A. W., Hall, P., Prvan, T. (1998). Cross-validation for the smoothing of distribution functions. *Biometrika*, *85*, 799–808.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, *7*, 1–26.

Eurostat. (2000). Low-wage employees in EU countries. In *Statistics in focus: Population and social conditions*. Luxembourg: Office for Official Publications of EC.

Falk, M. (1983). Relative efficiency and deficiency of kernel type estimators of smooth distribution functions. *Statistica Neerlandica*, *37*, 73–83.

Falk, M. (1985). Asymptotic normality of the kernel quantile estimator. *The Annals of Statistics*, *13*, 428–433.

Gong, Y., Peng, L., Qi, Y. C. (2010). Smoothed jackknife empirical likelihood method for ROC curve. *Journal of Multivariate Analysis*, *101*, 1520–1531.

Hall, P., La Scala, B. (1990). Methodology and algorithms of empirical likelihood. *International Statistical Review*, *58*, 109–127.

Jing, B., Yuan, J., Zhou, W. (2009). Jackknife empirical likelihood. *Journal of American Statistical Association*, *104*, 1224–1232.

Li, Z., Gong, Y., Peng, L. (2011). Empirical likelihood intervals for conditional value-at-risk in heteroscedastic regression models. *Scandinavian Journal of Statistics*, *38*, 781–787.

Owen, A. (1988). Empirical likelihood ratio confidence intervals for single functional. *Biometrika*, *75*, 237–249.

Owen, A. (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics*, *18*, 90–120.

Preston, I. (1995). Sampling distributions of relative poverty statistics. *Applied Statistics*, *44*, 91–99 [Correction, *45*, 399 (1996).

Quenouille, M. (1956). Note on bias in estimation. *Biometrika*, *43*, 353–360.

Rao, J. N. K., Wu, C. (2010). Pseudo-empirical likelihood inference for multiple frame surveys. *Journal of American Statistical Association*, *105*, 1494–1503.

Rongve, I. (1997). Statistical inference for poverty indices with fixed poverty lines. *Applied Economics*, *29*, 387–392.

Shao, J. (2003). *Mathematical statistics*. New York: Springer Science & Business Media.

Shao, J., Tu, D. (1995). *The jackknife and bootstrap*. New York: Springer.

Turkey, J. W. (1958). Bias and confidence in not quite large samples (Abstract). *The Annals of Mathematical Statistics*, *29*, 614.

Wei, Z., Zhu, L. (2010). Evaluation of value at risk: An empirical likelihood approach. *Statistica Sinica*, *20*, 455–468.

Wei, Z., Wen, S., Zhu, L. (2009). Empirical likelihood-based evaluations of value at risk models. *Science in China Series A: Mathematics*, *52*, 1995–2006.

Yang, B. Y., Qin, G. S., Qin, J. (2011). Empirical likelihood-based inferences for a low income proportion. *The Canadian Journal of Statistics*, *39*, 1–16.

Yves, G. B., Chris, J. S. (2003). Variance estimation for a low income proportion. *Applied Statistics*, *52*, 457–468.

Zheng, B. (2001). Statistical inference for poverty measures with relative poverty lines. *Journal of Econometrics*, *101*, 337–356.

Zhou, X. H., Qin, G. S., Lin, H. Z., Li, G. (2006). Inferences in censored cost regression models with empirical likelihood. *Statistica Sinica*, *16*, 1213–1232.