

Research Article

Empirical Likelihood-Based Inference in Regressive Model with Moment Restrictions

Zhao Zhiwen ¹ and He Jinghua^{1,2}

¹College of Mathematics, Jilin Normal University, Siping 136000, China

²Zongai Senior High Middle School of Shouyang, Shanxi 045499, China

Correspondence should be addressed to Zhao Zhiwen; zhaozhiwenjilin@126.com

Received 23 August 2020; Revised 21 November 2020; Accepted 8 December 2020; Published 29 December 2020

Academic Editor: Krzysztof Puszynski

Copyright © 2020 Zhao Zhiwen and He Jinghua. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we consider the parameter estimation problem of linear regression model when the auxiliary information can be denoted by moment restrictions. We use the weighted least squares method to estimate the model parameters and to obtain the weights based on the auxiliary information by using the empirical likelihood method. The limiting distribution of the estimator is established, and the simulation studies are carried out to demonstrate the feasibility of our theoretical results.

1. Introduction

Suppose $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)\}$ is an independent and identically distributed random sample from the following regression model:

$$Y_i = \alpha^T X_i + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (1)$$

where X_i is $p \times 1$ random vector and Y_i is a scalar response variable; $\alpha = (\alpha_1, \dots, \alpha_p)^T$ is a $p \times 1$ vector of regression coefficients, and the error ε_i is a sequence of *i.i.d.* random variables with $E\varepsilon_i = 0$ and $E\varepsilon_i^2 = \sigma^2$. Moreover, we assume that $\{X_i, i = 1, 2, \dots\}$ is independent of $\{\varepsilon_i, i = 1, 2, \dots\}$.

Linear regression model is widely used in empirical work in economics, medicine, and many other disciplines due to its simple form. In this paper, we consider the parameter estimating problem of linear regression model when auxiliary information is available. We care about how to decrease the estimation bias by using the auxiliary information based on the empirical likelihood method. The customary design-based estimator does not make use of auxiliary population information at the estimation stage. Observe that auxiliary information can be used to increase the precision of estimators in sample surveys. Therefore, Chambers and Dunstan [1] propose a simple method for estimating the

population distribution functions which allows auxiliary population information to be directly incorporated into the estimation process. Rao et al. [2] obtain the ratio and difference estimators of a population distribution function under a general sampling design by using auxiliary population information. Chen and Qin [3] show that the empirical likelihood method can be naturally applied to finite population inference problems by making effective use of the auxiliary information. Zhang [4] employs the method of empirical likelihood to construct confidence intervals for M-functionals in the presence of auxiliary information. Zhong and Rao [5] show that making effective use of auxiliary population information can lead to efficient estimators for making inferences on finite population parameters. García and Cebrián [6] derive confidence intervals for medians in a finite population by using multiauxiliary information through a multivariate regression-type estimator of the population distribution function. By assuming auxiliary information on the unknown distribution of the data, Crudu and Porcu [7] introduce a weighted Z-estimator for moment condition models. Moreover, Tang and Leng [8] propose a weighted least squares estimation which incorporates auxiliary information to the efficiency of the estimator. Hellerstein and Imbens [9] analyze the estimation of

coefficients in regression models under moment restrictions in which the moment restrictions are derived from auxiliary data. But the moment restriction does not include nuisance parameter. We further consider the parameter estimation problem of linear regression model when the auxiliary information can be denoted by moment restrictions which include nuisance parameter based on the empirical likelihood method. Specifically, the auxiliary information can be expressed as moment constraint $E(g(X_i, \theta_0)) = 0$, where $g(X_i, \theta) = (g_1(X_i, \theta), \dots, g_r(X_i, \theta))$, $\theta \in R^d$ and $r \geq d$. It is worth mentioning that θ can be different from α . Therefore, moment constraint $E(g(X_i, \theta_0)) = 0$ can denote a broad class of information. We will use the least squares method to estimate the model parameter and apply empirical likelihood method to obtain the weights by using the auxiliary information.

As a nonparametric method, empirical likelihood method is widely used in statistical inference of various models. Empirical likelihood method is first proposed by Owen [10–12] and is further generalized by Qin and Lawless [13]. Empirical likelihood method is used for statistical inference of various models (see Chen and Keilegom [14] and Nordman and Lahiri [15]). In addition, some statisticians have also begun to pay attention to the statistical inference of semiparametric regression model with constraints (see Amini and Roozbeh [16], Roozbeh and Hamzah [17], and Roozbeh et al. [18]).

This paper proceeds as follows. We first introduce the methodology and the main results. Next, we will undertake a simulation study to illustrate the feasibility of our method. Lastly, we give the proofs of the main results.

The symbols “ \xrightarrow{d} ” and “ \xrightarrow{p} ” denote convergence in distribution and convergence in probability, respectively. Convergence “almost surely” is written as “a.s.” Furthermore, $A \otimes B$ denotes the Kronecker product of matrices A and B , $\|\cdot\|$ denotes Euclidean norm of the matrix or vector, A^τ denotes the transposition of the matrix or the column vector A , $o_p(1)$ denotes a random variable that converges to zero in probability, and $O_p(1)$ denotes a random variable that is bounded in probability.

2. Methods and Main Results

In this section, we will introduce our method and give the main results of this paper. Firstly, we obtain the weights of the weighed least squares estimator based on the moment constraint $E(g(X_i, \theta_0)) = 0$ by using the empirical likelihood method (see Owen [12]). Specifically, let

$$L(\theta) = \sup \left\{ \prod_{i=1}^n \omega_i : \omega_i \geq 0, \sum_{i=1}^n \omega_i = 1, \sum_{i=1}^n \omega_i g(X_i, \theta) = 0 \right\}, \quad (2)$$

where θ is unknown. Combining with the least squares method, we can obtain the following weighed least squares estimator:

$$\hat{\alpha} = \arg \min_{\alpha} \sum_{i=1}^n \omega_i (Y_i - \alpha^\tau X_i)^2. \quad (3)$$

Introducing Lagrange multipliers $\lambda_{\theta_0} \in R^r$ in $L(\theta_0)$, we obtain the following weights:

$$\omega_i(\theta_0) = \frac{1}{n} \frac{1}{1 + \lambda_{\theta_0}^\tau g(X_i, \theta_0)}, \quad (4)$$

where λ_{θ_0} satisfies

$$\frac{1}{n} \sum_{i=1}^n \frac{g(X_i, \theta_0)}{1 + \lambda_{\theta_0}^\tau g(X_i, \theta_0)} = 0. \quad (5)$$

Therefore, by (3) and (4), we know that

$$\hat{\alpha} = \left(\sum_{i=1}^n \omega_i(\theta_0) X_i X_i^\tau \right)^{-1} \sum_{i=1}^n \omega_i(\theta_0) Y_i X_i. \quad (6)$$

In order to obtain the limiting distribution of $\hat{\alpha}$, we assume that the following conditions hold:

(A₁) $E(X_1 X_1^\tau) = W$, and $W > 0$.

(A₂) $E(g(X_i, \theta_0) g^\tau(X_i, \theta_0)) = \Sigma > 0$. There exists a neighborhood of θ_0 and an integrable function $\tilde{W}(x)$ such that, in this neighborhood, $\partial g(x, \theta)/\partial \theta$ is continuous, $\|\partial g(x, \theta)/\partial \theta\| \leq \tilde{W}(x)$, $\|g(x, \theta)\|^3 \leq \tilde{W}(x)$, and the rank of $E(\partial g(X_1, \theta)/\partial \theta)$ is d .

Condition (A₂) can be found in a study by Qin and Lawless [13]. The following theorem presents the asymptotic properties of $\hat{\alpha}$.

Theorem 1. Under (A₁) and (A₂) and assuming that $EX_1^4 < \infty$, if α_0 is the true value of α , then

$$\sqrt{n}(\hat{\alpha} - \alpha_0) \xrightarrow{d} N(0, W^{-1}(\Lambda - \Lambda_{12} \Sigma^{-1}(\theta_0) \Lambda_{12}^\tau) W^{-1}), \quad (7)$$

where $\Lambda = E(X_1 X_1^\tau (Y_1 - X_1^\tau \alpha_0)^2)$ and $\Lambda_{12} = E(X_1 g^\tau(X_1, \theta_0)(Y_1 - X_1^\tau \alpha_0))$.

We know that the asymptotic variance of the least squares estimator $\hat{\alpha} = (\sum_{i=1}^n X_i X_i^\tau)^{-1} \sum_{i=1}^n Y_i X_i$ is $W^{-1} \Lambda W^{-1}$. Note that matrix $\Sigma^{-1}(\theta_0)$ is nonnegative definite. Hence, Theorem 1 implies that the weighted least squares estimator has a smaller asymptotic variance compared with the least squares estimator. That is to say, using the moment restrictions does improve the effectiveness of the estimation.

$\hat{\alpha}$ contains unknown parameters. Therefore, we need to further estimate the unknown parameter θ . Let $\hat{\theta} = \arg \max_{\theta} L(\theta)$ in (2). Using results in a study by Qin and Lawless [13], we know that

$$\omega_i(\hat{\theta}) = \frac{1}{n} \frac{1}{1 + \lambda_{\hat{\theta}}^\tau g^\tau(X_i, \hat{\theta})}, \quad (8)$$

where $\lambda_{\hat{\theta}}$ satisfies

$$\frac{1}{n} \sum_{i=1}^n \frac{g^\tau(X_i, \hat{\theta})}{1 + \lambda_{\hat{\theta}}^\tau g^\tau(X_i, \hat{\theta})} = 0, \quad (9)$$

and $(\lambda_{\hat{\theta}}, \hat{\theta})$ solves

$$\frac{1}{n} \sum_{i=1}^n \frac{(\partial g^\tau(X_i, \hat{\theta})/\partial \theta^\tau) \lambda_{\hat{\theta}}}{1 + \lambda_{\hat{\theta}}^\tau g^\tau(X_i, \hat{\theta})} = 0. \quad (10)$$

Let

$$\hat{\alpha}' = \arg \min_{\beta} \sum_{i=1}^n \omega_i(\hat{\theta}) (Y_i - \alpha^\tau X_i)^2. \quad (11)$$

There are no unknown parameters in $\hat{\alpha}'$, so it can be used to estimate unknown parameters α in practice. Next, we investigate the limiting properties of $\hat{\alpha}'$. For convenience sake, let $\Gamma(\theta_0) = E(\partial g^\tau(X_i, \theta_0)/\partial \theta)$, $\Omega(\theta_0) = (\Gamma(\theta_0)\Sigma^{-1}(\theta_0)\Gamma^\tau(\theta_0))^{-1}$ and $B = \Sigma^{-1}(\theta_0)(I - \Gamma(\theta_0)\Omega(\theta_0)\Gamma^\tau(\theta_0)\Sigma^{-1}(\theta_0))$, where I is the unit matrix. The following theorem gives the limiting property of $\hat{\alpha}'$.

Theorem 2. Under (A_1) and (A_2) and assuming that $EX_1^4 < \infty$, if α_0 is the true value of α , then

$$\sqrt{n}(\hat{\alpha}' - \alpha_0) \xrightarrow{d} N(0, W^{-1}(\Lambda - \Lambda_{12}B\Lambda_{12}^\tau)W^{-1}). \quad (12)$$

Similarly, since matrix B is nonnegative definite, the weighted least squares estimator $\hat{\alpha}'$ also has a smaller asymptotic variance compared with the least squares estimator.

3. Simulation Studies

In this section, we study the finite sample properties of the above estimator by simulation. We consider the following models:

Model 1. $Y_i = \alpha_1 + \alpha_2 X_i + \varepsilon_i$, $i = 1, 2, \dots, n$, where the distribution of $\{X_i\}$ is exponential with unit mean and the distribution of $\{\varepsilon_i\}$ is standard normal.

Model 2. $Y_i = \alpha_1 + \alpha_2 X_i + \varepsilon_i$, $i = 1, 2, \dots, n$, where the distribution of $\{X_i\}$ is exponential with unit mean. $\varepsilon_i = \lambda_i \xi_i + (1 - \lambda_i) \zeta_i$, where the distribution of ξ_i is standard normal, the distribution function of ζ_i is $N(0, 9)$, and the probability distribution of λ_i is $P(\lambda_i = 1) = 1 - P(\lambda_i = 0) = 1 - p$.

Model 3. $Y_i = \alpha_1 X_{1i} + \alpha_2 X_{2i} + \varepsilon_i$, $i = 1, 2, \dots, n$, where the distribution of X_{1i} is exponential with mean 0.5, the distribution of X_{2i} is exponential with unit mean, and the distribution of $\{\varepsilon_i\}$ is standard normal.

In order to illustrate the feasibility of our proposed method, we compute the mean absolute deviation ratio of the estimator given by (11) to the least squares estimator. For model 1 and model 2, we use $g(X_i, \theta) = (\theta_1(X_i - 1), \theta_1(X_i^2 - 2), \theta_2(X_i - 1))$ as moment restrictions. For model 3, we use $g(X_{2i}, \theta) = (\theta_1(X_{2i} - 1), \theta_1(X_{2i}^2 - 2), \theta_2(X_{2i} - 1))$ as moment restrictions.

For each experiment, we conduct 1000 repetitions, and three alternative sample sizes ($n = 30, 100$ and 300) are considered. The simulation results for model 1 are summarized in Table 1. For model 2, we take different pollution level $p = 0.2, 0.4$, and the simulation results are presented in Tables 2 and 3, respectively. Moreover, the simulation results for model 3 are summarized in Table 4.

It can be seen from Tables 1–4 that the ratio of the mean absolute deviations of the least squares estimation with moment restrictions to that of the ordinary least squares estimation is very small. This implies that, for the different sample sizes, the different error, and the different parameters, the least squares estimation with moment restrictions is a more precise estimator. Hence, utilizing the auxiliary information indeed improves the efficiency of estimation.

4. Proofs of the Main Results

In order to prove Theorem 1, we first present several lemmas.

Lemma 1. If (A_1) and (A_2) hold, then

$$\max_{1 \leq i \leq n} \|g(X_i, \theta_0)\| = o_p(n^{(1/2)}). \quad (13)$$

Proof. Let

$$\Sigma_n(\theta_0) = \frac{1}{n} \sum_{i=1}^n g(X_i, \theta_0) g^\tau(X_i, \theta_0). \quad (14)$$

From the law of large numbers, we know that

$$\Sigma_n(\theta_0) \xrightarrow{a.s.} \Sigma(\theta), \quad \text{as } n \rightarrow \infty. \quad (15)$$

Using the method of the proof of Lemma 2 in [19], we can prove that (13) holds. Lemma 1 is established. \square

Lemma 2. Under (A_1) and (A_2) and assuming that $EX_1^4 < \infty$, if α_0 is the true value of α , then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i^\tau \alpha_0, g^\tau(X_i, \theta_0))^\tau \xrightarrow{d} N(0, M), \quad (16)$$

where

$$M = \begin{pmatrix} \Lambda & \Lambda_{12} \\ \Lambda_{12}^\tau & \Sigma \end{pmatrix}. \quad (17)$$

Proof. According to the central limit theorem of independent and identical distribution, it is easy to know that Lemma 2 holds. \square

Lemma 3. Under (A_1) and (A_2) and assuming that $EX_1^4 < \infty$, we have

$$\lambda_{\theta_0} = O_p(n^{-(1/2)}). \quad (18)$$

TABLE 1: The simulation results for model 1.

(α_1, α_2)	$n = 30$	$n = 100$	$n = 300$
(0, 1)	$(0.1503, 0.0962) \times 10^{-3}$	$(0.4429, 0.0805) \times 10^{-3}$	$(0.3574, 0.0265) \times 10^{-3}$
(0, 2)	(0.0013, 0.0001)	$(0.5059, 0.1128) \times 10^{-3}$	$(0.6640, 0.0125) \times 10^{-3}$
(0, 3)	(0.0018, 0.0001)	(0.0014, 0.0001)	(0.0018, 0.0000)
(0, 4)	$(0.3198, 0.0327) \times 10^{-3}$	(0.0066, 0.0002)	(0.8846, 0.0121)
(0, 5)	(0.0234, 0.0012)	(0.0094, 0.0001)	(0.2887, 0.0035)
(3, 1)	(0.0016, 0.0037)	$(0.1341, 0.4235) \times 10^{-3}$	$(0.5841, 0.8164) \times 10^{-3}$
(3, 2)	$(0.6284, 0.6922) \times 10^{-3}$	(0.0013, 0.0016)	(0.0009, 0.0012)
(3, 3)	(0.0622, 0.1554)	$(0.6129, 0.5704) \times 10^{-3}$	$(0.2794, 0.2578) \times 10^{-3}$
(3, 4)	(0.0028, 0.0018)	$(0.2891, 0.1785) \times 10^{-3}$	(0.0013, 0.0009)
(3, 5)	(0.0027, 0.0014)	$(0.1241, 0.0512) \times 10^{-3}$	$(0.1882, 0.1078) \times 10^{-3}$
(6, 1)	$(0.0589, 0.9486) \times 10^{-3}$	(0.0003, 0.0021)	$(0.0692, 0.4257) \times 10^{-3}$
(6, 2)	(0.0003, 0.0011)	$(0.0529, 0.2192) \times 10^{-3}$	$(0.1971, 0.6406) \times 10^{-3}$
(6, 3)	$(0.1244, 0.2126) \times 10^{-3}$	$(0.3027, 0.7651) \times 10^{-3}$	$(0.2588, 0.9244) \times 10^{-3}$
(6, 4)	$(0.0604, 0.1164) \times 10^{-3}$	$(0.2382, 0.3704) \times 10^{-3}$	$(0.2163, 0.3758) \times 10^{-3}$
(6, 5)	$(0.2337, 0.3729) \times 10^{-3}$	$(0.3247, 0.5159) \times 10^{-4}$	$(0.1075, 0.1288) \times 10^{-3}$

TABLE 2: The simulation results for model 2 ($p = 0.2$).

(α_1, α_2)	$n = 30$	$n = 100$	$n = 300$
(0, 1)	$(0.7218, 0.1631) \times 10^{-3}$	$(0.2068, 0.3868) \times 10^{-3}$	$(0.4816, 0.0390) \times 10^{-3}$
(0, 2)	(0.0023, 0.0004)	(0.0730, 0.0198)	(0.0016, 0.0002)
(0, 3)	$(0.5612, 0.2035) \times 10^{-4}$	$(0.4688, 0.0697) \times 10^{-3}$	(0.2909, 0.0131)
(0, 4)	(0.0044, 0.0003)	$(0.7965, 0.0492) \times 10^{-3}$	(0.0090, 0.0002)
(0, 5)	$(0.4604, 0.1192) \times 10^{-3}$	(0.4088, 0.0280)	(0.1139, 0.0025)
(3, 1)	(0.0015, 0.0068)	(0.0044, 0.0139)	$(0.1750, 0.0806) \times 10^{-3}$
(3, 2)	(0.0129, 0.0143)	(0.0013, 0.0020)	$(0.5347, 0.9881) \times 10^{-3}$
(3, 3)	(0.0011, 0.0009)	(0.0095, 0.0068)	(0.0025, 0.0036)
(3, 4)	$(0.2909, 0.3911) \times 10^{-3}$	(0.0011, 0.0009)	$(0.8104, 0.8536) \times 10^{-3}$
(3, 5)	$(0.1006, 0.0220) \times 10^{-3}$	$(0.7710, 0.6039) \times 10^{-3}$	(0.0047, 0.0028)
(6, 1)	(0.0005, 0.0032)	$(0.0283, 0.3442) \times 10^{-3}$	(0.0002, 0.0001)
(6, 2)	$(0.0262, 0.3567) \times 10^{-3}$	$(0.0566, 0.3088) \times 10^{-3}$	(0.0006, 0.0021)
(6, 3)	$(0.1006, 0.0679) \times 10^{-3}$	(0.0009, 0.0023)	(0.0009, 0.0022)
(6, 4)	$(0.2790, 0.5487) \times 10^{-3}$	$(0.1418, 0.2085) \times 10^{-3}$	(0.0029, 0.0045)
(6, 5)	(0.0007, 0.0011)	$(0.4147, 0.5056) \times 10^{-3}$	$(0.0713, 0.3029) \times 10^{-3}$

TABLE 3: The simulation results for model 2 ($p = 0.4$).

(α_1, α_2)	$n = 30$	$n = 100$	$n = 300$
(0, 1)	$(0.0097, 0.4296) \times 10^{-3}$	(0.0010, 0.0008)	$(0.1758, 0.1850) \times 10^{-3}$
(0, 2)	$(0.9817, 0.3201) \times 10^{-3}$	$(0.4836, 0.0195) \times 10^{-3}$	(0.0011, 0.0000)
(0, 3)	(0.0072, 0.0065)	(0.0027, 0.0010)	(0.0037, 0.0001)
(0, 4)	(0.0116, 0.0014)	(0.0012, 0.0002)	(0.3652, 0.0319)
(0, 5)	(0.0027, 0.0003)	(0.0064, 0.0007)	(0.0059, 0.0001)
(3, 1)	$(0.1768, 0.5159) \times 10^{-3}$	(0.0020, 0.0059)	(0.0303, 0.0777)
(3, 2)	(0.0081, 0.0105)	(0.0025, 0.0038)	(0.0104, 0.0211)
(3, 3)	(0.0101, 0.0088)	(0.0011, 0.0008)	$(0.0517, 0.1450) \times 10^{-3}$
(3, 4)	$(0.6036, 0.3940) \times 10^{-3}$	(0.0017, 0.0012)	(0.0036, 0.0027)
(3, 5)	(0.0019, 0.0011)	(0.0033, 0.0020)	(0.0058, 0.0028)
(6, 1)	(0.0008, 0.0020)	(0.0007, 0.0056)	(0.0007, 0.0056)
(6, 2)	$(0.0726, 0.1935) \times 10^{-3}$	$(0.3039, 0.8903) \times 10^{-3}$	(0.0006, 0.0020)
(6, 3)	$(0.3542, 0.1108) \times 10^{-3}$	$(0.4838, 0.2409) \times 10^{-4}$	$(0.0006, 0.0011) \times 10^{-3}$
(6, 4)	(0.0008, 0.0012)	(0.0056, 0.0102)	(0.0013, 0.0020)
(6, 5)	$(0.1707, 0.2671) \times 10^{-3}$	$(0.3853, 0.4861) \times 10^{-3}$	$(0.2360, 0.2885) \times 10^{-3}$

TABLE 4: The simulation results for model 3.

(α_1, α_2)	$n = 30$	$n = 100$	$n = 300$
(0, 1)	$(0.5517, 0.0218) \times 10^{-3}$	$(0.7855, 0.1065) \times 10^{-3}$	$(0.6022, 0.0350) \times 10^{-3}$
(0, 2)	(0.0015, 0.0001)	(0.0028, 0.0001)	$(0.6271, 0.0008) \times 10^{-3}$
(0, 3)	(0.0024, 0.0000)	$(0.4016, 0.0078) \times 10^{-3}$	(0.0041, 0.0000)
(0, 4)	(0.0042, 0.0001)	(0.0073, 0.0001)	(0.0118, 0.0002)
(0, 5)	(0.0026, 0.0000)	(0.0088, 0.0001)	(0.0131, 0.0001)
(3, 1)	$(0.0113, 0.5320) \times 10^{-3}$	$(0.0175, 0.2675) \times 10^{-3}$	(0.0306, 0.5866)
(3, 2)	$(0.0671, 0.1866) \times 10^{-3}$	$(0.0302, 0.3476) \times 10^{-3}$	$(0.0799, 0.6932) \times 10^{-3}$
(3, 3)	$(0.1395, 0.5421) \times 10^{-3}$	$(0.0174, 0.1121) \times 10^{-3}$	$(0.0232, 0.2970) \times 10^{-3}$
(3, 4)	$(0.1309, 0.3677) \times 10^{-3}$	$(0.1905, 0.1747) \times 10^{-3}$	(0.0129, 0.0251)
(3, 5)	$(0.0390, 0.2233) \times 10^{-3}$	$(0.0205, 0.3749) \times 10^{-3}$	$(0.6382, 0.1322) \times 10^{-4}$
(6, 1)	$(0.0057, 0.2006) \times 10^{-3}$	$(0.1223, 0.2038) \times 10^{-3}$	$(0.0544, 0.1760) \times 10^{-3}$
(6, 2)	$(0.0140, 0.3297) \times 10^{-3}$	$(0.0864, 0.4524) \times 10^{-3}$	$(0.0568, 0.7259) \times 10^{-3}$
(6, 3)	$(0.0060, 0.2511) \times 10^{-3}$	$(0.0181, 0.2149) \times 10^{-3}$	$(0.0092, 0.2824) \times 10^{-3}$
(6, 4)	$(0.0143, 0.2344) \times 10^{-3}$	$(0.0378, 0.2712) \times 10^{-3}$	$(0.2272, 0.0901) \times 10^{-3}$
(6, 5)	$(0.0078, 0.1755) \times 10^{-3}$	$(0.0235, 0.2951) \times 10^{-3}$	$(0.7367, 0.3097) \times 10^{-3}$

Proof. Let $\lambda_{\theta_0} = \xi\beta_0$, where $\|\beta_0\| = 1, \xi = \|\lambda_{\theta_0}\|$. By (5), we know that

$$\begin{aligned}
 0 &= \frac{\beta_0^\tau}{n} \sum_{i=1}^n \frac{g(X_i, \theta_0)}{1 + \xi\beta_0^\tau g(X_i, \theta_0)} \\
 &= \frac{\beta_0^\tau}{n} \sum_{i=1}^n g(X_i, \theta_0) - \frac{\xi}{n} \sum_{i=1}^n \frac{(\beta_0^\tau g(X_i, \theta_0))^2}{1 + \xi\beta_0^\tau g(X_i, \theta_0)} \\
 &\leq \frac{\beta_0^\tau}{n} \sum_{i=1}^n g(X_i, \theta_0) \\
 &\quad - \frac{\xi}{1 + \xi \max_{1 \leq i \leq n} \|g(X_i, \theta_0)\|} \beta_0^\tau \Sigma_n(\theta_0) \beta_0.
 \end{aligned} \tag{19}$$

Therefore, we have

$$\begin{aligned}
 &\xi\beta_0^\tau \Sigma_n(\theta_0) \beta_0 - \max_{1 \leq i \leq n} \|g(X_i, \theta_0)\| \frac{\beta_0^\tau}{n} \sum_{i=1}^n g(X_i, \theta_0) \\
 &\leq \frac{\beta_0^\tau}{n} \sum_{i=1}^n g(X_i, \theta_0).
 \end{aligned} \tag{20}$$

Note that

$$\begin{aligned}
 \left| \frac{\beta_0^\tau}{n} \sum_{i=1}^n g(X_i, \theta_0) \right| &\leq \left\| \frac{1}{n} \sum_{i=1}^n g(X_i, \theta_0) \right\| \\
 &= O_P(n^{-(1/2)}).
 \end{aligned} \tag{21}$$

Furthermore, by Lemma 1, we obtain

$$\max_{1 \leq i \leq n} \|g(X_i, \theta_0)\| \frac{\beta_0^\tau}{n} \sum_{i=1}^n g(X_i, \theta_0) = o_P(1). \tag{22}$$

By the law of large numbers for independent and identically distributed random variables, we have

$$\beta_0^\tau \Sigma_n(\theta_0) \beta_0 \xrightarrow{a} .s. \beta_0^\tau \Sigma(\theta_0) \beta_0. \tag{23}$$

Combining with (21) and (22), we know that Lemma 3 holds. \square

Proof of Theorem 1.

By (5), we know that

$$\lambda_{\theta_0} = (\Sigma_n(\theta_0))^{-1} \frac{1}{n} \sum_{i=1}^n g(X_i, \theta_0) + (\Sigma_n(\theta_0))^{-1} R_n(\theta_0), \tag{24}$$

where

$$R_n(\theta_0) = \frac{1}{n} \sum_{i=1}^n g^\tau(X_i, \theta_0) \frac{(\lambda_{\theta_0}^\tau g(X_i, \theta_0))^2}{1 + \lambda_{\theta_0}^\tau g(X_i, \theta_0)}. \tag{25}$$

By Lemmas 1–3, we know that

$$R_n(\theta_0) = o_P(n^{-(1/2)}). \tag{26}$$

By using Taylor's formula for $(1/1 + \lambda_{\theta_0}^\tau g(X_i, \theta_0))$, we have

$$\frac{1}{1 + \lambda_{\theta_0}^\tau g(X_i, \theta_0)} = 1 - \lambda_{\theta_0}^\tau g(X_i, \theta_0) + o(\lambda_{\theta_0}^\tau g(X_i, \theta_0)). \tag{27}$$

So, we have

$$\omega_i = \frac{1}{n} (1 - \lambda_{\theta_0}^\tau g(X_i, \theta_0) + o(\lambda_{\theta_0}^\tau g(X_i, \theta_0))). \tag{28}$$

By Lemma 1 and Lemma 3, we have

$$\max_{1 \leq i \leq n} |\lambda_{\theta_0}^\tau g(X_i, \theta_0)| = o_P(1). \tag{29}$$

Therefore, we have that uniformly for i ,

$$\frac{o(\lambda_{\theta_0}^\tau g(X_i, \theta_0))}{\lambda_{\theta_0}^\tau g(X_i, \theta_0)} = o_P(1). \tag{30}$$

Note that

$$\sqrt{n}(\hat{\alpha} - t\alpha_0) = T_n^{-1} S_n, \tag{31}$$

where

$$\begin{aligned} T_n &= \sum_{i=1}^n \omega_i(\theta_0) X_i X_i^\tau, \\ S_n &= \sqrt{n} \sum_{i=1}^n \omega_i(\theta_0) X_i (Y_i - \alpha_0^\tau X_i). \end{aligned} \quad (32)$$

Firstly, we prove that

$$T_n \xrightarrow{P} W. \quad (33)$$

After simple algebra calculation, we have

$$\begin{aligned} T_n &= \frac{1}{n} \sum_{i=1}^n (1 - \lambda_{\theta_0}^\tau g(X_i, \theta_0)) (1 + o_p(1)) X_i X_i^\tau \\ &= \frac{1}{n} \sum_{i=1}^n X_i X_i^\tau - \left(\frac{1}{n} \sum_{i=1}^n X_i X_i^\tau \right) \\ &\quad \otimes \left((\Sigma_n(\theta_0))^{-1} R_n(\theta_0) \right)^\tau g(X_i, \theta_0) (1 + o_p(1)) \\ &\quad - \frac{1}{n} \sum_{i=1}^n X_i X_i^\tau \otimes (\Sigma_n(\theta_0))^{-1} \times \\ &\quad \frac{1}{n} \sum_{i=1}^n (g(X_i, \theta_0))^\tau g(X_i, \theta_0) (1 + o_p(1)) \\ &\triangleq U_1 - U_2 - U_3. \end{aligned} \quad (34)$$

By the law of large numbers for independent and identically distributed random variables, we have

$$U_1 \xrightarrow{P} W. \quad (35)$$

Furthermore, by (26) and Lemma 2, we know that

$$U_2 = o_p(1), \quad (36)$$

$$U_3 = o_p(1). \quad (37)$$

Combined with (36) and (37), this establishes (33).

In the following, we consider S_n . By (26) and (26), we know that

$$\begin{aligned} S_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^n (1 - \lambda_{\theta_0}^\tau g(X_i, \theta_0)) \\ &\quad \times (1 + o_p(1)) X_i (Y_i - \alpha_0^\tau X_i) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i (Y_i - \alpha_0^\tau X_i) \\ &\quad - \frac{1}{n} \sum_{i=1}^n (Y_i - \alpha_0^\tau X_i) X_i g^\tau(X_i, \theta_0) (\Sigma_n(\theta_0))^{-1} \\ &\quad \times \frac{1}{\sqrt{n}} \sum_{i=1}^n g(X_i, \theta_0) \\ &\quad + o_p(1). \end{aligned} \quad (38)$$

By the law of large numbers for independent and identically distributed random variables, we know that

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \alpha_0^\tau X_i) X_i g^\tau(X_i, \theta_0) \xrightarrow{a} .s. \Lambda_{12}. \quad (39)$$

Furthermore, by Lemma 2, we know that

$$S_n \xrightarrow{d} N(0, \Lambda_{12} \Sigma^{-1} \Lambda_{12}^\tau). \quad (40)$$

Combined with (31), we know that Theorem 1 holds. \square

Proof of Theorem 2

Using the results in [13], we know that

$$\lambda_{\hat{\theta}} = B \frac{1}{n} \sum_{i=1}^n g^\tau(X_i, \theta_0) + o_p(n^{-(1/2)}). \quad (41)$$

After simple algebra calculation, we have

$$\sqrt{n}(\hat{\alpha}' - \alpha_0) = \tilde{T}_n^{-1} \tilde{S}_n, \quad (42)$$

where $\tilde{T}_n = \sum_{i=1}^n \omega_i(\hat{\theta}) X_i X_i^\tau$, $\tilde{S}_n = \sqrt{n} \sum_{i=1}^n \omega_i(\hat{\theta}) X_i (Y_i - \alpha_0^\tau X_i)$. Similar to the proof of Theorem 1, we know that

$$\tilde{T}_n \xrightarrow{P} W, \tilde{S}_n \xrightarrow{d} N(0, \Lambda - \Lambda_{12} B \Lambda_{12}^\tau). \quad (43)$$

Thus, Theorem 2 is established. \square

5. Conclusions

In this paper, we discuss how to use auxiliary information to improve the efficiency of regression model parameter estimation when auxiliary information exists. First of all, we use the auxiliary information to establish the estimation equation. Based on this estimation equation, we use empirical likelihood method to obtain the weight of weighted least squares estimation of regression model parameters and then give the weighted least squares estimation of model parameters. Secondly, it is noticed that the auxiliary information contains unknown parameters, so the weighted least squares estimator also contains unknown parameters. Therefore, we use the maximum empirical likelihood estimation to further estimate the unknown parameters contained in the auxiliary information. Therefore, we then obtain the weighted least squares estimation that can be used in practice. Finally, we use simulation analysis to illustrate the feasibility of our method.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. 11571138, 11671054, 11301137, 11271155, 11371168, J1310022, and 11501241) and the National Social Science Fund of China (no. 16BTJ020).

References

- [1] R. L. Chambers and R. Dunstan, "Estimating distribution functions from survey data," *Biometrika*, vol. 73, no. 3, pp. 597–604, 1986.
- [2] J. N. K. Rao, J. G. Kovar, and H. J. Mantel, "On estimating distribution functions and quantiles from survey data using auxiliary information," *Biometrika*, vol. 77, no. 2, pp. 365–375, 1990.
- [3] J. Chen and J. Qin, "Empirical likelihood estimation for finite populations and the effective usage of auxiliary information," *Biometrika*, vol. 80, no. 1, pp. 107–116, 1993.
- [4] B. Zhang, "Empirical likelihood confidence intervals for M-functionals in the presence of auxiliary information," *Statistics & Probability Letters*, vol. 32, no. 1, pp. 87–97, 1997.
- [5] B. Zhong and J. Rao, "Empirical likelihood inference under stratified random sampling using auxiliary population information," *Biometrika*, vol. 87, no. 4, pp. 929–938, 2000.
- [6] M. Rueda García and A. Arcos Cebrián, "On estimating the median from survey data using multiple auxiliary information," *Metrika*, vol. 54, no. 1, pp. 59–76, 2001.
- [7] F. Crudu and E. Porcu, "Z-estimators and auxiliary information for strong mixing processes," *Stochastic Environmental Research and Risk Assessment*, vol. 33, no. 1, pp. 1–11, 2019.
- [8] C. Y. Tang and C. Leng, "An empirical likelihood approach to quantile regression with auxiliary information," *Statistics & Probability Letters*, vol. 82, no. 1, pp. 29–36, 2012.
- [9] J. K. Hellerstein and G. W. Imbens, "Imposing moment restrictions from auxiliary data by weighting," *Review of Economics and Statistics*, vol. 81, no. 1, pp. 1–14, 1999.
- [10] A. B. Owen, "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, vol. 75, no. 2, pp. 237–249, 1988.
- [11] A. Owen, "Empirical likelihood ratio confidence regions," *The Annals of Statistics*, vol. 18, no. 1, pp. 90–120, 1990.
- [12] A. B. Owen, *Empirical Likelihood*, Chapman and Hall, New York, NY, USA, 2001.
- [13] J. Qin and J. Lawless, "Empirical likelihood and general estimating equations," *The Annals of Statistics*, vol. 22, no. 1, pp. 300–325, 1994.
- [14] S. X. Chen and I. Van Keilegom, "A review on empirical likelihood methods for regression," *Test*, vol. 18, no. 3, pp. 415–447, 2009.
- [15] D. J. Nordman and S. N. Lahiri, "A review of empirical likelihood methods for time series," *Journal of Statistical Planning and Inference*, vol. 155, pp. 1–18, 2014.
- [16] M. Amini and M. Roozbeh, "Optimal partial ridge estimation in restricted semiparametric regression models," *Journal of Multivariate Analysis*, vol. 136, pp. 26–40, 2015.
- [17] M. Roozbeh and N. A. Hamzah, "Feasible robust estimator in restricted semiparametric regression models based on the LTS approach," *Communications in Statistics - Simulation and Computation*, vol. 46, no. 9, pp. 7332–7350, 2017.
- [18] M. Roozbeh, G. Hesamian, and M. G. Akbari, "Ridge estimation in semi-parametric regression models under the stochastic restriction and correlated elliptically contoured errors," *Journal of Computational and Applied Mathematics*, vol. 378, pp. 1–19, 2020.
- [19] Z. W. Zhao and D. H. Wang, "Statistical inference for generalized random coefficient autoregressive model," *Mathematical & Computer Modelling*, vol. 56, no. 7-8, pp. 152–166, 2012.