Economics Letters 164 (2018) 19–23

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Robust estimation and empirical likelihood inference with exponential squared loss for panel data models

Shaomin Li [a], Kangning Wang [b], Yanyan Ren [a,*]

[a] School of Economics, Shandong University, Jinan, China
[b] School of Statistics, Shandong Technology and Business University, Yantai, China

HIGHLIGHTS

- A new robust and efficient empirical likelihood method (ESL-EL) for panel data is proposed.
- The asymptotic normality and consistency of the proposed estimator are proved under some appropriate conditions.
- The influence function of the proposed estimator is bounded, which means our estimator is robust.
- Simulations show that our estimator is mildly affected by the contaminations while the GEE method becomes invalid.

ARTICLE INFO

Article history:
Received 1 June 2017
Received in revised form 19 December 2017
Accepted 20 December 2017
Available online 28 December 2017

JEL classification:
C13
C15
C18
C23
C40

Keywords:
Exponential squared loss
Empirical likelihood
Panel data
Robust estimation

ABSTRACT

This paper introduces a robust estimation for panel data models using the exponential squared loss function. We propose the method by constructing the robust empirical likelihood ratio function. The Monte Carlo simulations show that the proposed estimator is robust in the fixed and random effects models.

1. Introduction

During the last two decades, panel data analysis has been well developed and widely used in many fields, such as economics, finance, biology, engineering and medical studies. For an overview of theories and applications of parametric panel data models, one can refer to the books by Baltagi (2008) and Hsiao (2014). The estimations and inferences for panel data models are usually based either on the maximum likelihood or the generalized estimating equation (GEE, Liang and Zeger, 1986). Both methods, however, are sensitive to outliers. Therefore, a number of robust approaches have been proposed, such as Cantoni (2004), He et al. (2005), Qin et al. (2009a, b), Baltagi and Bresson (2012), Qin et al. (2012), Baltagi and Bresson (2015) and Dhaene and Zhu (2016). Most of these papers adopt Huber's loss function (Huber, 1981). Although Huber's method is robust, it has limitations in terms of efficiency. To achieve better robustness and efficiency, Wang et al. (2013) proposed a class of robust estimators based on the exponential squared loss function $\Phi_\gamma(r) = 1 - \exp(-r^2/\gamma)$ which is widely used in boost algorithm (Friedman et al., 2000). For example, the parameter of the linear model $y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i$ can be estimated by minimizing $\sum_{i=1}^n \Phi_{\gamma_n}(r_i)$, where $r_i = y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}$ represents the residual of the ith observation, and $\gamma_n > 0$ controls the degree of robustness and efficiency. For a large γ, $1 - \exp(-r^2/\gamma) \approx r^2/\gamma$, which means the proposed estimator is similar to the least squares estimator in this case. When γ is small, observations with large values of $|r_i|$ will result in large losses of $\Phi_{\gamma_n}(r_i)$ and therefore have a small impact on the estimation. Hence, a smaller γ would limit the influence of outliers on the estimators. The optimal choice of

* Corresponding author.
E-mail address: yanyanren@sdu.edu.cn (Y. Ren).

https://doi.org/10.1016/j.econlet.2017.12.029
0165-1765/

γ is proposed in Section 3.3.2 in Wang et al. (2013). Minimizing $\sum_{i=1}^{n}\Phi_{\gamma_n}(r_i)$ is equivalent to solving the following equations

$$\sum_{i=1}^{n}\boldsymbol{x}_i\varphi_{\gamma_n}(r_i)=\mathbf{0}, \tag{1}$$

where $\varphi_\gamma(r)=-\frac{2r}{\gamma}\exp(-r^2/\gamma)$ is the derivative of $\Phi_\gamma(r)$. Note that $\varphi_\gamma(r)$ is also a bounded score since $\lim_{r\to\infty}\varphi_\gamma(r)=0$. Wang et al. (2013) pointed out that their method is more robust than other existing robust methods, e.g., Huber's estimator (Huber, 1981), quantile regression estimator (Koenker and Bassett, 1978), composite quantile regression estimator (Zou and Yuan, 2008), etc. However, this approach was only considered when the data are independent.

In this paper, by using the exponential squared loss, we introduce a new robust and efficient empirical likelihood method (ESL-EL) for correlated data. Specifically, we replace the score function in generalized estimating equations with $\varphi_\gamma(r)$, and then construct the empirical likelihood (Owen, 1988) based on the robust GEE. By selecting the additional tuning parameter γ automatically, ESL-EL can achieve a better balance between robustness and efficiency.

The outline of the paper is as follows. Section 2 introduces our approach in details and presents the asymptotic properties of the proposed approach. Section 3 conducts several numerical simulations to compare the performance of the proposed method with GEE. Section 4 concludes.

2. A robust estimator and empirical likelihood inference

Consider the model

$$y_{it}=\boldsymbol{x}_{it}^\top\boldsymbol{\beta}+\alpha_i+\varepsilon_{it},\quad i=1,\dots,n,t=1,\dots,m_i,$$

where $\boldsymbol{x}_{it}$ is a p-dimensional vector of regressors, $\{\alpha_i\}$ reflect unobserved individual effects, $\{\varepsilon_{it}\}$ are independent and identically distributed error terms with mean 0 and variance σ_ε^2. Let $N=\sum_{i=1}^{n}m_i$, $\boldsymbol{y}_i=(y_{i1},y_{i2},\dots,y_{im_i})^\top$, $\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\dots,\boldsymbol{x}_{im_i})$, $\boldsymbol{\varepsilon}_i=(\varepsilon_{i1},\varepsilon_{i2},\dots,\varepsilon_{im_i})^\top$, $\mathbf{1}=(1,1,\dots,1)^\top$, then the model can be rewritten as

$$\boldsymbol{y}_i=\boldsymbol{X}_i^\top\boldsymbol{\beta}+\alpha_i\mathbf{1}+\boldsymbol{\epsilon}_i, i=1,\dots,n. \tag{2}$$

In the following, we will introduce our estimators for fixed and random effects models, respectively.

2.1. Fixed effects models

For fixed effects models, individual effects $\{\alpha_i\}$ can be viewed as intercepts of each group. In this paper, we treat $\{\alpha_i\}$ as nuisance parameters[1] and eliminate them by difference: $y_{it}-y_{i1}=(\boldsymbol{x}_{it}-\boldsymbol{x}_{i1})^\top\boldsymbol{\beta}+\varepsilon_{it}-\varepsilon_{i1}, t=2,3,\dots,m_i, i=1,2,\dots,n$. Let $\boldsymbol{y}_i^*=(y_{i2}-y_{i1},y_{i3}-y_{i1},\dots,y_{im_i}-y_{i1})^\top$, $\boldsymbol{X}_i^*=(\boldsymbol{x}_{i2}-\boldsymbol{x}_{i1},\boldsymbol{x}_{i3}-\boldsymbol{x}_{i1},\dots,\boldsymbol{x}_{im_i}-\boldsymbol{x}_{i1})$, $\boldsymbol{\varepsilon}_i^*=(\varepsilon_{i2}-\varepsilon_{i1},\varepsilon_{i3}-\varepsilon_{i1},\dots,\varepsilon_{im_i}-\varepsilon_{i1})^T$, model (2) is transformed to

$$\boldsymbol{y}_i^*=\boldsymbol{X}_i^{*\top}\boldsymbol{\beta}+\boldsymbol{\varepsilon}_i^*,\quad i=1,\dots,n. \tag{3}$$

Then we construct the p-dimensional auxiliary random vector

$$\boldsymbol{Z}_i^*(\boldsymbol{\beta})=\boldsymbol{X}_i^{*\top}\boldsymbol{V}_i^{*-1}\varphi_{\gamma_N}(\boldsymbol{r}_i^*(\boldsymbol{\beta})),$$

where $\boldsymbol{r}_i^*(\boldsymbol{\beta})=\boldsymbol{y}_i^*-\boldsymbol{X}_i^{*\top}\boldsymbol{\beta}$, $\varphi_\gamma(r)=-(2r/\gamma)\exp(-r^2/\gamma)$ and $\boldsymbol{V}_i^*$ represents covariance matrix of $\boldsymbol{\varepsilon}_i^*$. Obviously, $\boldsymbol{V}_i^*=2\sigma_\varepsilon^2\boldsymbol{R}_i^*$ where $\boldsymbol{R}_i^*$ is the correlation matrix with diagonal elements being 1 and others being 0.5. The nuisance parameter σ_ε^2 can be eliminated in the robust generalized estimating equations (RGEE)

$$\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Z}_i^*(\boldsymbol{\beta})=\mathbf{0}. \tag{4}$$

Based on (4), we construct the empirical log-likelihood ratio function of β

$$l^*(\boldsymbol{\beta})=-\sum_{i=1}^{n}\log(1+\boldsymbol{\lambda}^\top\boldsymbol{Z}_i^*(\boldsymbol{\beta})),$$

where $\boldsymbol{\lambda}$ satisfies

$$\sum_{i=1}^{n}\frac{\boldsymbol{Z}_i^*(\boldsymbol{\beta})}{1+\boldsymbol{\lambda}^\top\boldsymbol{Z}_i^*(\boldsymbol{\beta})}=\mathbf{0}.$$

We can obtain the estimator of $\boldsymbol{\beta}$ by maximizing the log-likelihood ratio:

$$\hat{\boldsymbol{\beta}}_{ESL-EL}=\underset{\boldsymbol{\beta}}{\operatorname{argmax}}\, l^*(\boldsymbol{\beta}).$$

Remark. The tuning parameter γ is very important since it controls the degree of robustness and efficiency of the proposed estimator. We use the procedure of Wang et al. (2013) to select γ and set the MM-estimator (Yohai, 1987) as the initial estimate of β. For fixed effects models, the MM-estimation can be directly used since the errors are independent. For random effects models, we ignore the dependence among errors when using the MM-estimation.

2.2. Random effect models

For random effect models, we assume that $\{\alpha_i\}$ are independent and identically distributed with mean 0 and variance σ_α^2 and are uncorrelated with $\{\varepsilon_{it}\}$ and $\{\boldsymbol{X}_i\}$. Let $e_{it}=\alpha_i+\varepsilon_{it}$ and $\boldsymbol{e}_i=(e_{i1},e_{i2},\dots,e_{im_i})^\top$, $i=1,\dots,n$, $t=1,\dots,m_i$, then model (2) can be rewritten as

$$\boldsymbol{y}_i=\boldsymbol{X}_i^\top\boldsymbol{\beta}+\boldsymbol{e}_i,\quad i=1,\dots,n. \tag{5}$$

Similar to model (3), we construct the auxiliary random vector $\boldsymbol{Z}_i(\boldsymbol{\beta},\rho)=\boldsymbol{X}_i^\top\boldsymbol{V}_i^{-1}\varphi_{\gamma_N}(\boldsymbol{r}_i(\boldsymbol{\beta}))$. Note that unlike $\boldsymbol{V}_i^*$, the covariance matrix $\boldsymbol{V}_i$ in this situation has unknown parameters in its correlation matrix. Specifically, $\boldsymbol{V}_i=(\sigma_\varepsilon^2+\sigma_\alpha^2)\boldsymbol{R}_i$, where $\boldsymbol{R}_i$ is an $m_i\times m_i$ matrix with diagonal elements being 1 and others being the unknown parameter $\rho=\sigma_\alpha^2/(\sigma_\varepsilon^2+\sigma_\alpha^2)$. The empirical log-likelihood ratio function is

$$l(\boldsymbol{\beta},\rho)=-\sum_{i=1}^{n}\log(1+\boldsymbol{\lambda}^\top\boldsymbol{Z}_i(\boldsymbol{\beta},\rho)),$$

where $\boldsymbol{\lambda}$ satisfies

$$\sum_{i=1}^{n}\frac{\boldsymbol{Z}_i(\boldsymbol{\beta},\rho)}{1+\boldsymbol{\lambda}^\top\boldsymbol{Z}_i(\boldsymbol{\beta},\rho)}=\mathbf{0}.$$

According to Wang et al. (2005), ρ can be estimated by

$$\hat{\rho}=\frac{1}{nS}\sum_{i=1}^{n}\frac{1}{m_i(m_i-1)}\sum_{j\neq k}\varphi_{\gamma_N}(r_{ij})\varphi_{\gamma_N}(r_{ik}),$$

where $S=\frac{1}{N}\sum_{i,j}\varphi_{\gamma_N}^2(r_{ij})$. After replacing ρ with $\hat{\rho}$ in $l(\boldsymbol{\beta},\rho)$, we can obtain the estimator $\hat{\boldsymbol{\beta}}_{ESL-EL}$ by maximizing $l(\boldsymbol{\beta},\hat{\rho})$. According to Qin et al. (2009a), the influence function of $\hat{\boldsymbol{\beta}}_{ESL-EL}$ is $IF(\boldsymbol{D}_0,\boldsymbol{\beta}_0)=(E[\frac{\partial\boldsymbol{Z}}{\partial\boldsymbol{\beta}}]\big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0})^{-1}\boldsymbol{Z}(\boldsymbol{D}_0,\boldsymbol{\beta}_0)$, where $\boldsymbol{D}_0=(\boldsymbol{X}_0,y_0)$ represents the outlier. Since $\varphi_\gamma(r)$ is bounded, the influence function is bounded as well. Therefore, $\hat{\boldsymbol{\beta}}_{ESL-EL}$ is robust.

The analysis of the properties of the random effects estimator is more general because of the additional parameter ρ in Z_i. Thus in

[1] Put $\hat{\beta}$ into the model and we can estimate fixed effects $\{\alpha_i\}$ individually.