Statistics and Probability Letters 104 (2015) 153–162

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

A moment-based test for individual effects in the error component model with incomplete panels

Jianhong Wu [a,b,*], Jinxu Qin [a], Qing Ding [a]

[a] School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China
[b] Center for Applied Statistics, Renmin University of China, Beijing, China

ARTICLE INFO

Article history:
Received 5 February 2015
Received in revised form 15 April 2015
Accepted 15 May 2015
Available online 3 June 2015

MSC:
62F05
62H10
62H15

Keywords:
Incomplete panel
Individual effect
Moment-based test

ABSTRACT

This paper proposes a robust test statistic for individual effects in the error component model with incomplete panel data. Specifically, on the base of the difference of variance estimators of the idiosyncratic errors at different levels, we construct a statistic to test for the existence of individual effects, which can be shown to asymptotically normally distributed under the null hypothesis. Power study shows that the test can detect local alternatives distinct at the parametric rate from the null hypothesis and has a larger asymptotic power than the corresponding ANOVA F test when the individual heterogeneity effects are correlated with regressors. Monte Carlo evidence shows that the test statistic has desired finite sample properties. A real data example is analyzed for illustration.

1. Introduction

In practice, one often find that some data are incomplete or unbalanced, which can be led by many reasons. For example, in the econometric analysis of consumer panel data, some individuals' data can be dropped out after some time points because the corresponding households moved out. Up to our knowledge, however, the incomplete panels do not receive the attention they deserve in theory. Note that, the incomplete data include the monotone missing patterns and the random missing patterns. In this paper, for the sake of simplicity, we mainly focus on the monotone missing panel data patterns, and leave the latter as a future research topic.

Consider the following one-way error component regression model with monotone missing panel data,

$$y_{it} = \alpha + X_{it}'\beta + \mu_i + \nu_{it}, \quad i = 1, 2, \ldots, N,\ t = 1, 2, \ldots, T_i, \tag{1}$$

where α is a scalar, X_{it} is the it-th observation on K explanatory variables and β is the vector of coefficients of the explanatory variables. The individual effect μ_i is used to capture the heterogeneity of individuals, which is seen as a random variable with mean zero and variance σ_μ^2 in this paper, and the idiosyncratic errors $\{\nu_{it}\}$ are assumed independent and identically distributed with mean zero and positive variance σ_ν^2. Throughout this paper, the missing pattern is set to monotone, which means that once $\{y_{it}, X_{it}\}$ misses then $\{y_{is}, X_{is}\}$ will miss for each $s > t$. As Baltagi (2008) argued, misspecification of individual effects in the error component will lead to seriously biased standard errors and inefficient statistical inference. It is therefore important to test for the existence of individual effects in the error component. The main purpose of this paper is to develop some robust test methods for individual effects in the error component model with incomplete panels.

* Corresponding author at: School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China.
E-mail address: wjhstat1@163.com (J. Wu).

http://dx.doi.org/10.1016/j.spl.2015.05.013
0167-7152/

In the literature, there are some relevant work on testing for the individual effects in the error component models with incomplete panel data. We give simple reviews as follows. Baltagi and Li (1990) extended the traditional Breusch and Pagan (1980) Lagrange multiplier test to the error component regression model with incomplete data. The test is based on the assumption that the alternative hypothesis is two-sided while in fact it should be one-sided due to the nonnegativity of variance. Moulton (1987) extended the uniformly mostly powerful test (UMPT) of Honda (1985) to the case with incomplete panels. It is reported by Moulton and Randolph (1989) that the asymptotic $N(0, 1)$ approximation of the test can be very poor even in large samples. Moulton and Randolph (1989) further suggested a test statistic which has better critical value approximations by standardized the test statistic in Moulton (1987). Similar to the aforementioned LM tests, this test has a simple form that is easy to compute from OLS residuals. However, all the tests mentioned above require the assumption of normality of idiosyncratic errors and independence among the regressors, individual effects and the idiosyncratic errors.

In this paper, we propose a test statistic for individual effects in the error component model with incomplete panel data, which can be robust to the normally distributional assumption of the idiosyncratic errors and the independence among the regressors, effects and the idiosyncratic errors. Specifically, on the basis of the difference of variance estimators of the idiosyncratic errors at different levels, we construct a test statistic which can be shown to asymptotically normally distributed under the null hypothesis, and as the extension of the moment-based test methods of Wu and Li (2014) to the case with incomplete panels. Power study shows that the test can detect local alternatives distinct at the parametric rate from the null hypothesis and has a larger asymptotic power than the corresponding ANOVA F test when the individual heterogeneity effects are correlated with regressors. Monte Carlo evidence shows that the test statistic has desired finite sample properties. A real data example is also analyzed for illustration.

The rest of this paper is organized as follows. In Section 2, we introduce an orthogonal-based estimator for the parameter coefficient, construct a moment-based test statistic and study its asymptotic properties under the null and the alternatives respectively. Monte Carlo simulation experiments and the corresponding results are stated in Section 3. Section 4 applies our testing method to a real data example. Section 5 contains a short discussion and conclusion. Technique proofs are relegated to the Appendix.

2. Testing for individual effect

2.1. Parameter estimation

Model (1) can be rewritten into the vector form as

$$y_i = \alpha\iota_{T_i} + X_i\beta + \mu_i\iota_{T_i} + \nu_i, \quad i = 1, 2, \ldots, N, \tag{2}$$

where $y_i = (y_{i1}, y_{i2}, \ldots, y_{iT_i})'$, $X_i = (X_{i1}, X_{i2}, \ldots, X_{iT_i})'$, $\nu_i = (\nu_{i1}, \nu_{i2}, \ldots, \nu_{iT_i})'$ and ι_{T_i} is a vector of ones of dimension T_i. We can find a matrix Q_{T_i} such that $(\frac{\iota_{T_i}}{\sqrt{T_i}}, Q_{T_i})$ is a $T_i \times T_i$ orthogonal matrix and then $Q'_{T_i}\iota_{T_i} = 0$, although the matrix Q_{T_i} is not unique. Conducting a transformation on model (2) with the matrix Q'_{T_i}, we have

$$Q'_{T_i}y_i = Q'_{T_i}X_i\beta + Q'_{T_i}\nu_i, \quad i = 1, 2, \ldots, N. \tag{3}$$

Denote $Q_{T_i} = (q_{i,1}, q_{i,2}, \ldots, q_{i,T_i-1})$ and $q_{i,l} = (q_{i1l}, q_{i2l}, \ldots, q_{iT_il})'$, $l = 1, 2, \ldots, T_i - 1$. In practice, we recommend a simple form of Q_{T_i} with $q_{i,l} = \frac{1}{\sqrt{l(l+1)}}[le_{i,l+1} - \sum_{k=1}^{l} e_{i,k}]$ for $l = 1, 2, \ldots, T_i - 1$, where $e_{i,k}$ is a T_i-dimensional vector with the kth element being one and other elements zero. We consider the ordinary least squares (OLS) method for parameter β in model (3)

$$\hat{\beta} = \operatorname{argmin}_{\beta} \sum_{i=1}^{N} \|Q'_{T_i}y_i - Q'_{T_i}X_i\beta\|^2 = \left(\sum_{i=1}^{N} X_i' P_{\iota_{T_i}^{\perp}} X_i\right)^{-1} \sum_{i=1}^{N} X_i' P_{\iota_{T_i}^{\perp}} y_i,$$

where $\|\cdot\|$ is the Euclidean norm and $P_{\iota_{T_i}^{\perp}} = Q_{T_i}Q'_{T_i} = I_{T_i} - \iota_{T_i}(\iota'_{T_i}\iota_{T_i})^{-1}\iota'_{T_i}$ is independent of the matrix Q_{T_i}. Note that the above estimator is the so-called orthogonality-based estimator (OBE) in the paper of Wu and Zhu (2010). We can further obtain that $\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{N} \iota'_{T_i}(y_i - X_i\hat{\beta})$ with $n = \sum_{i=1}^{N} T_i$. Under some mild conditions, $\hat{\alpha}$ and $\hat{\beta}$ can be shown to $\sqrt{N}$-consistent and asymptotically normal, see Wu and Zhu (2010) and Baltagi (2008) for more details on the parameter estimation.

2.2. Construction of statistics and power study

In this subsection, we consider to construct a statistic to test for individual effects in model (1). And we need to point out that all the asymptotic results in this paper are based on the setting that the individual number N tends to infinity and the numbers of time periods $\{T_i\}$ are fixed.

The hypotheses of the test can be formalized as

$$H_0 : \sigma_\mu^2 = 0 \quad \text{vs.} \quad H_1 : \sigma_\mu^2 > 0,$$

where σ_μ^2 is the variance of individual effects. For model (3), it holds that

$$E\|Q'_{T_i}\nu_i\|^2 = E\left[\sum_{l=1}^{T_i-1}(q'_{i,l}\nu_i)^2\right] = (T_i - 1)\sigma_\nu^2.$$