Economics Letters 127 (2015) 47–50

Contents lists available at ScienceDirect

Economics Letters

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

An LM test based on generalized residuals for random effects in a nonlinear model

William Greene[a,*], Colin McKenzie[b]

[a] New York University, USA
[b] Keio University, Japan

HIGHLIGHTS

- We derive the LM test statistic for random effects in a panel probit model.
- We obtain a reparameterization of the statistic that produces a feasible calculation.
- We show that the statistic can be computed using generalized residuals.
- We demonstrate that the results generalize to other models.
- The test is employed in a substantive application.

ARTICLE INFO

Article history:
Received 29 November 2014
Received in revised form
18 December 2014
Accepted 22 December 2014
Available online 31 December 2014

JEL classification:
C23
C25
I11

Keywords:
LM test
Panel data
Probit
Random effects

ABSTRACT

We obtain an LM test for the random effects probit model. In the natural parameterization of the model the necessary derivatives are identically zero under the null hypothesis. After a reparameterization, the feasible LM test is based on generalized residuals.

1. Introduction

Empirical analyses involving limited dependent variables and random individual effects in panel data sets have become quite common. By far the leading application of the random effects (RE) model after the linear regression is the binary probit model. The estimated coefficients from RE and pooled probit models are different because of the different normalizations implied by the models. As Arulampalam (1999) discusses, in the presence of random effects, an adjustment of the estimated partial effects is needed to remove an ambiguity in the interpretation of the estimates.

Our interest is in the Lagrange Multiplier (LM) test for random effects in the probit model. The LM test has provided a standard means of testing parametric restrictions for a variety of settings. Its primary advantage among the trinity of tests (LM, Likelihood Ratio (LR), Wald) is that it is based on the null, restricted model, which is usually simpler to estimate than the alternative, unrestricted one. Breusch and Pagan's (1980) LM test for random effects in a linear model that is based on pooled OLS residuals is the leading example.

Testing for random effects in the probit model is an example of a problem that emerges when the parametric restriction in the null hypothesis puts the value of a variance parameter on the boundary of the parameter space. The restriction is that the standard deviation of the random effect equals zero. When RE probit models are estimated, popular computer packages automatically produce LR and Wald-type tests of the null hypothesis of no random effects, but would appear to use the $\chi^2_{(1)}$ (or standard normal)

* Correspondence to: Stern School of Business, New York University, 44 West 4, St., New York, NY, 10012, USA. Tel.: +1 212 998 0876.
E-mail address: wgreene@stern.nyu.edu (W. Greene).

http://dx.doi.org/10.1016/j.econlet.2014.12.031

distribution to compute the p-values for these tests. If, under the null hypothesis, the parameter being tested lies on the boundary of the parameter space, an additional advantage of the LM test is that it will still have standard distributional properties, whereas the LR and Wald tests will not (see Andrews, 2001). In fact, in testing for random effects in the probit model, the LR and Wald tests will be distributed as a $(1/2)\chi^2_{(1)}$ distribution under the null hypothesis (see Gourieroux et al., 1987). This means the correct critical values for these two tests at the 5% and 10% significance level are 5.02 and 3.84 respectively, rather than the commonly used values of 3.84 and 2.71 from the $\chi^2_{(1)}$ distribution.

The RE probit model is, after the linear regression model, by far the leading application of the more general class of random effects models. Despite the obvious simplicity of the restricted model, the LM test for this model does not appear in the existing literature. One reason for this is that the usual parameterization of the model has the inconvenient feature that the score vector is identically zero at the restricted ML estimates. The received literature (e.g., Chesher, 1984, Lee and Chesher, 1986 and Kiefer, 1982) identifies a handful of specific cases in which the score vector needed to compute the LM statistic is identically zero at the restricted estimates, which would seem to preclude using the LM test. We find that the RE probit model represents an entire class of such models.

Lee and Chesher (1986) discuss a general theory of how to deal with score vectors that are zero under the null hypothesis. Despite what would seem to be its broad application, we have not found any applications in the subsequent 30 years of literature. We will provide a useful expression for the LM test statistic and illustrate its use with an empirical application on hospitalization behavior.

2. The random effects probit model

The random effects probit model is

$$\begin{aligned}
&y^*_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + \sigma_u u_i + \varepsilon_{it}; \quad i = 1, \ldots, n; \; t = 1, \ldots, T_i, \\
&y_{it} = 1[y^*_{it} > 0], \\
&\varepsilon_{it} \sim N[0, 1^2], \qquad u_i \sim N[0, 1^2], \\
&E[\varepsilon_{it}\varepsilon_{js}] = 0, \quad i \neq j, \; t \neq s; \\
&E[u_i u_j] = 0, \quad i \neq j, \qquad E[\varepsilon_{it} u_s] = 0 \quad \forall i, t, s,
\end{aligned} \tag{1}$$

where $\boldsymbol{\beta}$ and $\mathbf{x}_{it}$ are $K \times 1$ vectors. (The analysis to follow is not dependent on normality for u_i, though that is the natural case to consider.) The log likelihood for a sample of n observations, conditioned on the unobserved heterogeneity, $u_1, u_2, \ldots, u_n$, is

$$\log L(\boldsymbol{\beta}, \sigma_u | u_1, \ldots, u_n) = \sum_{i=1}^{n} \log \prod_{t=1}^{T_i} \Phi[q_{it}(\boldsymbol{\beta}'\mathbf{x}_{it} + \sigma_u u_i)], \tag{2}$$

where $\Phi(t)$ is the standard normal CDF and $q_{it} = 2y_{it} - 1$. ML estimation is based on the unconditional log likelihood given by

$$\begin{aligned}
\log L(\boldsymbol{\beta}, \sigma_u) &= \sum_{i=1}^{n} \log \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \Phi[q_{it}(\boldsymbol{\beta}'\mathbf{x}_{it} + \sigma_u u_i)] \right\} \phi(u_i)\, du_i \\
&= \sum_{i=1}^{n} \log L_i(\boldsymbol{\beta}, \sigma_u),
\end{aligned} \tag{3}$$

where $\phi(t)$ is the standard normal PDF. Butler and Moffitt's (1982) estimation method based on Hermite quadrature is generally used in contemporary applications.

2.1. LM test for random effects

To form the LM statistic for the test of the null hypothesis of no random effects, $\sigma_u = 0$, we require $\partial \log L_i(\boldsymbol{\beta}, \sigma_u)/\partial \sigma_u$;

$$\frac{\partial \log L_i(\boldsymbol{\beta}, \sigma_u)}{\partial \sigma_u} = \frac{\int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \Phi[q_{it}\ a_{it}] \right\} \left\{ \sum_{t=1}^{T_i} \frac{q_{it}\phi[q_{it}\ a_{it}]}{\Phi[q_{it}\ a_{it}]} \right\} u_i \phi(u_i)\, du_i}{\int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \Phi[q_{it}\ a_{it}] \right\} \phi(u_i)\, du_i}, \tag{4}$$

where $a_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + \sigma_u u_i$. In order to compute the LM statistic, we need to evaluate this expression at $\sigma_u = 0$. Moving all terms not involving u_i outside the integrals produces

$$\frac{\partial \log L_i(\boldsymbol{\beta}, 0)}{\partial \sigma_u} = \frac{\left\{ \prod_{t=1}^{T_i} \Phi[q_{it}(\boldsymbol{\beta}'\mathbf{x}_{it})] \right\} \left\{ \sum_{t=1}^{T_i} \frac{q_{it}\phi[q_{it}(\boldsymbol{\beta}'\mathbf{x}_{it})]}{\Phi[q_{it}(\boldsymbol{\beta}'\mathbf{x}_{it})]} \right\} \int_{-\infty}^{\infty} u_i \phi(u_i)\, du_i}{\left\{ \prod_{t=1}^{T_i} \Phi[q_{it}(\boldsymbol{\beta}'\mathbf{x}_{it})] \right\} \int_{-\infty}^{\infty} \phi(u_i)\, du_i}. \tag{5}$$

The integrals in the numerator and denominator are $E[u_i] = 0$ and 1, respectively. Regardless of the value of $\boldsymbol{\beta}'\mathbf{x}_{it}$, each term in $\partial \log L(\boldsymbol{\beta}, \sigma_u)/\partial \sigma_u$ is identically zero when σ_u equals zero. The terms in $\partial \log L_i(\boldsymbol{\beta}, \sigma_u)/\partial \boldsymbol{\beta}$ are also zero. The score vector under the null hypothesis is identically zero. The result (and the derivation to follow) will extend generally to other single index models with random effects. (Surprisingly, it also holds for the linear regression model for which Breusch and Pagan's LM test has been used since 1980. See Chesher, 1984.)

2.2. LM test based on a reparameterization

Chesher (1984), Lee and Chesher (1986) and Cox and Hinkley (1974) suggested reparameterization of the model as a strategy for obtaining the LM test. We use $\gamma = \sigma_u^2$. The log likelihood becomes

$$\begin{aligned}
\log L(\boldsymbol{\beta}, \gamma) &= \sum_{i=1}^{n} \log L_i(\boldsymbol{\beta}, \gamma) \\
&= \sum_{i=1}^{n} \log \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \Phi\left[q_{it}(\boldsymbol{\beta}'\mathbf{x}_{it} + u_i\sqrt{\gamma})\right] \right\} \phi(u_i)\, du_i.
\end{aligned} \tag{6}$$

Then,

$$\frac{\partial \log L_i(\boldsymbol{\beta}, \gamma)}{\partial \gamma} = \frac{\frac{1}{2\sqrt{\gamma}} \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \Phi_{it} \right\} \left\{ \sum_{t=1}^{T_i} g_{it} \right\} u_i \phi(u_i)\, du_i}{\int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \Phi_{it} \right\} \phi(u_i)\, du_i}, \tag{7}$$

where $b_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + u_i\sqrt{\gamma}$, $\phi_{it} = \phi(q_{it}b_{it})$, $\Phi_{it} = \Phi(q_{it}b_{it})$ and $g_{it} = q_{it}\phi_{it}/\Phi_{it}$. Note that $g_{it}u_i$ is $\partial \log \Phi_{it}/\partial(\sqrt{\gamma})$. Evaluated at $\gamma = 0$, the numerator now takes the form $0/0$. We use L'Hôpital's rule, taking the limits as γ approaches zero from above. Then,

$$\frac{\partial \log L_i(\boldsymbol{\beta}, 0)}{\partial \gamma} = \frac{\lim_{\gamma \downarrow 0} \frac{1}{2\frac{1}{2\sqrt{\gamma}}} \int_{-\infty}^{\infty} L_i \left[\left(\sum_{t=1}^{T_i} h_{it} \right) + \left(\sum_{t=1}^{T_i} g_{it} \right)^2 \right] \frac{1}{2\sqrt{\gamma}} u_i^2 \phi(u_i)\, du_i}{\int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \Phi_{it} \right\} \phi(u_i)\, du_i}, \tag{8}$$