

A NOTE ON ESTIMATED COEFFICIENTS IN RANDOM EFFECTS PROBIT MODELS*

Wiji Arulampalam**
Department of Economics,
University of Warwick,
Coventry, CV4 7AL,
U.K.

email: wiji.arulampalam@warwick.ac.uk
tel: (+44) (01203) 523471
fax: (+44) (01203) 523032

August, 1996.
Revised October, 1998.

Abstract: This note points out to applied researchers what adjustments are needed to the coefficient estimates in a random effects probit model in order to make valid comparisons in terms of coefficient estimates and marginal effects across different specifications. These adjustments are necessary because of the normalisation that is used by standard software in order to facilitate easy estimation of the random effects probit model.

Keywords: Random effects probit.

JEL Classification: C13, C23.

* I should like to thank Alison Booth, Stephen Jenkins, Gordon Kemp and Chris Orme for helpful comments.

** nee Narendranathan.

I. INTRODUCTION

Random effects (RE) probit models have recently become very popular due to the availability of data such as the British Household Panel Survey (Arulampalam and Booth (1998), Booth et al. (1996)). This model imposes the restriction that the correlation between successive error terms for the same individual is a constant and thus known in the literature as the ‘equicorrelation’ model. But a static model can be estimated by pooling the data and ignoring this particular correlation structure to obtain *consistent* parameter estimates (Robinson (1982), Maddala (1987)). Hence any discussion between the pooled probit model estimates and a RE probit model estimates is not very illuminating under the maintained assumption that the RE probit specification is correct. Estimated coefficients between these models produced by popular software such as Limdep (Greene (1998)), Stata (1997) etc. look different because of different normalisations that are used by these programs to facilitate easy estimation. This note points out to applied researchers what adjustments are needed in order to be able to make valid comparisons in terms of coefficient estimates and marginal effects across different specifications.

Econometric model is presented and the issues discussed in the next section. Section III provides the necessary calculations for the marginal effects and is followed by an illustration using an artificially generated data. Final section concludes.

II. ECONOMETRIC MODEL AND ISSUES

Consider the following model,

$$y_{it}^* = \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it} \quad i=1,2,\dots,n \text{ and } t=1,\dots,T \quad (1)$$

$$v_{it} = \alpha_i + u_{it} \quad (2)$$

and

$$y_{it} = 1 \text{ if } y_{it}^* > 0 \text{ and } = 0 \text{ else,}$$

where, y^* denotes the unobservable variable, y is the observed outcome, \mathbf{x} is observable time varying and time invariant vector of strictly exogenous characteristics which influence y^* , β is the vector of coefficients associated with the \mathbf{x} , α_i denotes the individual specific unobservable effect and the u_{it} is a random error. In the case of probit random effects (RE) it is also assumed that $u_{it} \sim \text{IN}(0, \sigma_u^2)$. In order to marginalise the likelihood we assume that, conditional on the \mathbf{x}_{it} , α_i s are $\text{IN}(0, \sigma_\alpha^2)$ and are independent of the u_{it} s and the \mathbf{x}_{it} s. This implies that the correlation between two successive error terms for the same individual is a constant given by,

$$\rho = \text{corr}(v_{it}, v_{it-1}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2}. \quad (3)$$

As shown in Heckman (1981), the parameters of this model are easily estimated by noting that the distribution of y_{it}^* conditional on α_i are independent normal. We have,

$$\text{Prob}(y_{it} = 1 \mid \alpha_i, \mathbf{x}_{it}) = \text{Prob}\left(\frac{u_{it}}{\sigma_u} > \frac{-\mathbf{x}_{it}'\beta - \alpha_i}{\sigma_u}\right) = \Phi(z_{it}) \quad (4)$$

where

$$z_{it} = -(\mathbf{x}_{it}'\beta + \alpha_i) / \sigma_u \quad (5)$$

and Φ is the distribution function of the standard normal variate. We then marginalise the appropriate likelihood function with respect to α , which is given by

$$\prod_i \left\{ \int_{-\infty}^{\infty} \prod_{t=1}^T \left[1 - \Phi(x_{it}'\beta^* + \sqrt{\frac{\rho}{1-\rho}} \alpha^*) \right]^{1-y_{it}} \left[\Phi(x_{it}'\beta^* + \sqrt{\frac{\rho}{1-\rho}} \alpha^*) \right]^{y_{it}} \phi(\alpha^*) d\alpha^* \right\} \quad (6)$$

where $\beta^* = \beta / \sigma_u$ and $\alpha^* = \alpha / \sigma_\alpha$. Software programs such as Limdep and Stata return estimates of β^* and ρ . But note the fact that the coefficient estimates here are normalised on σ_u .

In the above model, one could obtain consistent parameter estimates of the β coefficients by ignoring the correlation structure (Robinson (198), Maddala (1987)).¹ Thus, a simple pooled probit will provide consistent parameter estimates. But, because of the binary nature of the data, a simple probit model estimation does not allow one to estimate the scale parameter. One only obtains parameter estimates of β / σ_v . Hence it is clear that pooled probit parameter estimates will be equal to the RE probit model parameter estimates only when $\sigma_\alpha^2=0$. When $\sigma_\alpha^2 \neq 0$ the consistent pooled probit estimates will not look similar to the RE probit estimates because of the normalisation. A simple calculation will show that to convert RE probit coefficients from β / σ_u to β / σ_v , RE probit coefficient estimates need to be multiplied by the factor $\sqrt{1-\hat{\rho}}$. This will enable one to make valid comparisons not only between univariate and RE probit estimated effects, but also across estimates obtained from different specifications of the vector x_{it} .

¹ But the standard errors will be wrongly calculated. The covariance matrix of the estimated parameters produced by a standard software will be biased. As discussed in Guilkey and Murphy (1993), a consistent estimate of the covariance matrix is obtained as $H^{-1}GH^{-1}$ where, H is the hessian and the G the outer product of the score matrix. For example, one can obtain a consistent estimate of the covariance matrix in Limdep using the 'choice based sampling' option with weights set equal to one for each observation.

III MARGINAL EFFECTS

But how do we interpret the various estimated effects in these models with unobservable individual specific components? In standard cross-sectional univariate probits, it is customary to provide expected changes in the outcome probability when particular characteristics are changed one at a time, known as marginal effects.

Consider the mean effect of changing one particular continuous covariate x_j , $j=1, \dots, k$, by a small amount, on the outcome probability. In the pooled probit, under the normalisation $\sigma_v=1$, this effect will be given by

$$\frac{\partial [Prob(y_{it} = 1 | \mathbf{x}_{it})]}{\partial x_{jit}} = \frac{\partial [E(y_{it})]}{\partial x_{jit}} = \frac{\partial [\Phi(\mathbf{x}_{it}'\boldsymbol{\beta})]}{\partial x_{jit}} = \phi(\mathbf{x}_{it}'\boldsymbol{\beta}) \beta_j \quad j=1, \dots, k \quad (7)$$

As this varies with the values of \mathbf{x} we can either evaluate this at the means of the regressors or evaluate this separately for each individual in the sample and then average over the sample. But remember the implicit normalisation of $\sigma_v=1$ in the above evaluation.

The associated covariance matrix may be calculated using the so-called ‘Delta-Method’ (see Greene (1997)) which uses a first order Taylor-series expansion to calculate the covariance matrix in the case of non-linear functions of random variables. This is provided in Greene (1997) for the probit model,

$$Asy. Var [\phi(\mathbf{x}'\hat{\boldsymbol{\beta}}) \hat{\boldsymbol{\beta}}] = Asy. Var[\hat{\boldsymbol{\gamma}}] = \begin{bmatrix} \mathcal{J}\hat{\boldsymbol{\gamma}} \\ \mathcal{J}\hat{\boldsymbol{\beta}} \end{bmatrix} \mathbf{V} \begin{bmatrix} \mathcal{J}\hat{\boldsymbol{\gamma}} \\ \mathcal{J}\hat{\boldsymbol{\beta}} \end{bmatrix}' \quad (8)$$

where $\mathbf{V}=Asy.var[\hat{\boldsymbol{\beta}}]$ and the marginal effects $\hat{\boldsymbol{\gamma}}$ are evaluated at the means of the variables and hence $\phi(\mathbf{x}'\hat{\boldsymbol{\beta}}) = \hat{\phi}$ is a scalar. The matrix of derivatives is

$$\left[\frac{\partial \hat{\gamma}}{\partial \hat{\beta}} \right] = \hat{\phi} \mathbf{I} - \hat{\phi} (\hat{\beta} \mathbf{x}')$$

(9)

where, \mathbf{I} is a k dimensional Identity matrix.

But in the case of a RE probit model, as shown above, taking the normalisation into consideration implies that in the calculation of the marginal effects, the coefficient estimates need to be multiplied by $\sqrt{1-\hat{\rho}}$ prior to using (7) giving

$$\frac{\partial \left[\text{Prob} (y_{it} = 1 | \mathbf{x}_{it}) \right]}{\partial x_{jit}} = \phi(\mathbf{x}_{it}' \boldsymbol{\beta}) (\beta_j) = \phi(\mathbf{x}_{it}' \boldsymbol{\beta}^* \sqrt{1-\rho}) (\sqrt{1-\rho} \beta_j^*) \quad (10)$$

Software will return estimates of $\boldsymbol{\beta}^*$ and ρ , and their covariance matrix $\boldsymbol{\Omega}$. But

$\hat{\boldsymbol{\beta}} = \sqrt{1-\hat{\rho}} \hat{\boldsymbol{\beta}}^*$ which implies that

$$\text{Covar}(\hat{\boldsymbol{\beta}}) = \mathbf{A} \boldsymbol{\Omega} \mathbf{A}' \quad (11)$$

$$\text{where } \mathbf{A} = \begin{bmatrix} \sqrt{1-\hat{\rho}} & 0 & \cdot & \cdot & 0 & -\hat{\beta}_1 / (2\sqrt{1-\hat{\rho}}) \\ 0 & \sqrt{1-\hat{\rho}} & 0 & \cdot & \cdot & -\hat{\beta}_2 / (2\sqrt{1-\hat{\rho}}) \\ \cdot & 0 & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & 0 & \sqrt{1-\hat{\rho}} & -\hat{\beta}_k / (2\sqrt{1-\hat{\rho}}) \end{bmatrix}, \quad (12)$$

\mathbf{k} by $(\mathbf{k}+1)$ matrix which contains in its j 'th row the derivative of $\hat{\beta}_j$ with respect to all the $\hat{\beta}_j^*$ s as well as $\hat{\rho}$. To calculate the covariance matrix for the correct marginal effects given by (10), we can still use equation (8) but we need to replace the covariance matrix \mathbf{V} by $\mathbf{A} \boldsymbol{\Omega} \mathbf{A}'$ (see (11)).

IV ILLUSTRATION

The above points are illustrated using artificially generated data. The data are generated from the model,

$$y_{it} = 1 \text{ if } (1.0 - 0.5 * x_{it} + \alpha_i + u_{it}) > 0$$

$$= 0 \text{ else, } \quad I=1, \dots, 500, \text{ and } t=1, 2,$$

and x_{it} , α_i and u_{it} are all drawn as $iidN(0,1)$. This implies that ρ given in (3) is 0.5 and thus the theoretical correction factor $\sqrt{1-\rho}$ is 0.7071. Results from the pooled probit and the RE probit model estimation using the above data, along with the marginal effects calculated with and without the corrections, are given in Table 1. From these results, it is easy to see that once the correction is made to the RE probit coefficient estimates, they do not look very different to those from a simple pooled probit model. This is also true with regard to the marginal effects. It is also seen that when the corrections are not made to the coefficients and marginal effects the results can be very misleading.

Table 1 - Coefficient Estimates and Marginal Effects (std. errors)

	Pooled Probit		Random Effects Probit			
	Coefficient	Marginal Effect	Unadjusted		Adjusted	
			Coefficient	Marginal Effect	Coefficient	Marginal Effect
Intercept	0.705 (0.045)	0.216 (0.007)	1.092 (0.112)	0.230 (0.006)	0.706 (0.052)	0.216 (0.008)
x	-0.402 (0.048)	-0.123 (0.014)	-0.651 (0.083)	-0.137 (0.017)	-0.421 (0.045)	-0.129 (0.013)
ρ			0.581 (0.065)			
Log Likelihood	-523.6514		-497.4611			

Notes: The marginal effects are calculated at the means (1, -0.058).

V CONCLUSION

Not all commercially available software provide calculations of marginal effects in very routinely estimated models such as the random effects probit model. These are of practical importance to applied researchers. Comparison between different random effects probit model coefficient estimates (marginal effects), and between these and the pooled probit coefficient estimates (marginal effects) can be very misleading for the very simple reason of the normalisation that is implemented in software to facilitate easy estimation. This note discusses this point and provides the necessary adjustments to enable applied researchers to make valid comparisons.

References

- Arulampalam, W. and Booth, A. L. (1998) - 'Training and labour market flexibility: Is there a trade-off?', forthcoming, British Journal of Industrial Relations, Dec.
- Booth, A.L., Garcia-Serrano, C. and Jenkins, S. P. (1996) - 'New men and new women: Is there convergence in patterns of labour market transition?', Institute of Labour Research, Discussion Paper 96/01, University of Essex.
- Greene, W. H. (1997) - Econometric Analysis, Prentice-Hall International, Inc.
- Greene, W. H. (1998) - LIMDEP, Version 7.0: User's Manual, Bellport, NY: Econometric Software.
- Guilkey, D. K. and Murphy, J. L. (1993) - 'Estimation and testing in the random effects probit model', Journal of Econometrics, 59, 301-317.
- Heckman, J. J. (1981) - 'Statistical models for discrete panel data', in Structural Analysis of Discrete Data with Econometric Applications, ed. C. F. Manski and D. McFadden, 114-178, Cambridge: MIT Press.
- Maddala, G. S. (1987) - 'Limited dependent variable models using panel data', Journal of Human Resources, XX11,3, 307-337.
- Robinson, P. M. (1982) - 'On the asymptotic properties of estimators of models containing limited dependent variables', Econometrica, 50(1), 27-42.
- STATA (1997) - Statistical Software, Release 5, Stata Corporation.