DISCUSSION PAPER SERIES

IZA DP No. 3039

Simplified Implementation of the Heckman Estimator of the Dynamic Probit Model and a Comparison with Alternative Estimators

Wiji Arulampalam
Mark B. Stewart

September 2007

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Simplified Implementation of the Heckman Estimator of the Dynamic Probit Model and a Comparison with Alternative Estimators

Wiji Arulampalam
University of Warwick
and IZA

Mark B. Stewart
University of Warwick

Discussion Paper No. 3039
September 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3039
September 2007

ABSTRACT

Simplified Implementation of the Heckman Estimator of the Dynamic Probit Model and a Comparison with Alternative Estimators*

This paper presents a convenient shortcut method for implementing the Heckman estimator of the dynamic random effects probit model using standard software. It then compares the three estimators proposed by Heckman, Orme and Wooldridge based on three alternative approximations, first in an empirical model for the probability of unemployment and then in a set of simple simulation experiments.

JEL Classification: C23, C25, C13, C51

Keywords: dynamic discrete choice models, initial conditions, dynamic probit, panel data

Corresponding author:

Wiji Arulampalam
Department of Economics
University of Warwick
Coventry, CV4 7AL
United Kingdom
E-mail: wiji.arulampalam@warwick.ac.uk

* Wiji Arulampalam is grateful for financial support from the ESRC under Research Grant no. RES-000-22-0651.

1. Introduction

The initial conditions problem is well-recognised in the estimation of dynamic discrete choice models. Its cause is the presence of both the past value of the dependent variable and an unobserved heterogeneity term in the equation and the correlation between them. The strict exogeneity assumption for regressors, standardly used in static discrete choice models in order to marginalise the likelihood function with respect to the unobserved heterogeneity, cannot be used in a dynamic setting due to the presence of the lagged dependent variable.

The standard estimator in this context is that suggested by Heckman (1981a, 1981b), who was the first to explicitly address this problem. His approach involves the specification of an approximation to the reduced form equation for the initial observation and maximum likelihood estimation using the full set of sample observations and allowing cross-correlation between the main and initial period equations. However, use of the estimator has been limited by its requiring separate programming, since standard packages have not included it. (See Arulampalam and Bhalotra (2006) and Stewart (2007) for recent applications). This has led to the suggestion of alternative estimators that have the advantage of requiring only standard software. The estimators suggested by Orme (1997, 2001) and Wooldridge (2005), based on alternative approximations, are commonly used in place of the Heckman estimator for this reason. The main merit claimed by both Orme and Wooldridge for their estimators relative to Heckman's is that theirs can be straightforwardly estimated using standard software.

This paper presents a convenient shortcut method for implementing the Heckman estimator of the dynamic random effects discrete choice model using standard software designed for the estimation of static random effects models with heteroskedastic random effects (such as the gllamm program in Stata) or constrained random coefficient models. The increased ease and availability of the Heckman estimator that this provides removes the need for simpler alternatives. However since the Heckman estimator is itself based on an approximation, this raises the question of the relative finite sample performance of these three approximation-based estimators. This paper therefore also provides an examination of the relative merits of the Heckman, Orme and Wooldridge estimators in the absence of the software issue. It examines differences between the three estimators first in the context of an empirical illustration using a model for the probability of unemployment and then presents a small-scale Monte Carlo experiment of their finite sample performance in circumstances favourable to the Heckman estimator. The Orme and Wooldridge estimators are found to perform as well as, and in some aspects better than, the Heckman estimator.

2. Econometric Model and Estimators

The model for the observed dependent variable y_{it} is specified as

$$\Pr ob\left[y_{it} = 1 \mid y_{it-1}, \mathbf{x}_{it}, \alpha_i \right] = \Phi(\gamma y_{it-1} + \mathbf{x}_{it}\,'\boldsymbol{\beta} + \alpha_i) \qquad (1)$$

where $i=1,\ldots, N$; $t=2,\ldots, T_i$ (the panel may be unbalanced), $\boldsymbol{x}$ is a vector of strictly exogenous observed explanatory variables and $\boldsymbol{\beta}$ is the vector of coefficients associated with $\boldsymbol{x}$. The model includes the *observed* status of the dependent variable in the previous period, y_{it-1}. The model also has a random intercept α_i to account for individual-specific unobserved characteristics. Φ is the cumulative distribution function of a standard normal variate.

The standard uncorrelated random effects model assumes α_i uncorrelated with x_{it}. Alternativel, following Mundlak (1978), correlation between α_i and the observed characteristics can be allowed by assuming a relationship of the form

$$\alpha_i = x_i' a + \varepsilon_i$$

where $x_i = (x_{i1},\ldots,x_{iT})$ and with ε_i independent of x_i. (Averages of the x-variables over t have also been used.) Thus the model can be written as

$$\Pr ob\left[y_{it} = 1 \mid y_{it-1}, \mathbf{x}_{it}, \alpha_i \right] = \Phi(\gamma y_{it-1} + \mathbf{x}_{it}\,'\boldsymbol{\beta} + x_i' a + \varepsilon_i) \qquad (1a)$$

To simplify notation, specification (1) will be used here, with the understanding that when the Mundlak correlated random effects (CRE) model is used, x_{it} in (1) implicitly subsumes a full set of period-specific versions of the (time-varying) x-variables.

The Initial Conditions Problem and Heckman's Estimator

Correlation between α_i and y_{it-1} makes the latter endogenous in equation (1). Heckman suggested the use of an approximation to the process generating the first period observations using the same form of equation as for the rest of the observations but with some restrictions. Specifically he proposed the use of

$$\text{Prob}(y_{i1}=1|\alpha_i) = \Phi[\boldsymbol{z}_i'\lambda + \theta\,\alpha_i] \qquad i=1,\ldots,N \tag{2}$$

where $\boldsymbol{z}_i$ is a vector of exogenous covariates. This would be expected to include $\boldsymbol{x}_{i1}$ and additional variables that can be viewed as "instruments" such as pre-sample variables. Exogeneity corresponds to $\theta = 0$ and can be tested accordingly. The distribution function is assumed to be the same as in (1).

Equations (1) and (2) together specify a complete model for the process. In this model the contribution to the likelihood function for individual i is given by

$$L_i = \int \left(\Phi\left[\left(\mathbf{z}_{\mathbf{i1}}'\boldsymbol{\lambda} + \theta\alpha_i \right)\left(2y_{i1} - 1\right) \right] \prod_{t=2}^{T_i} \Phi\left[\left(\mathbf{x}_{it}'\boldsymbol{\beta} + \gamma y_{it-1} + \alpha_i \right)\left(2y_{it} - 1\right) \right] \right) g(\alpha_i)\,d\alpha_i \tag{3}$$

where $g(\alpha)$ is the probability density function of the unobservable individual-specific heterogeneity. In the standard case considered here, α is taken to be normally distributed and the integral in (3) can be evaluated using Gaussian--Hermite quadrature (Butler and Moffitt, 1982).

Shortcut Setup for Implementing Heckman's Estimator

The simplified implementation procedure proposed here involves the creation of a dummy variable, $D_{it} = 1$ if the observation belongs to period 1, $D_{it} = 0$ otherwise. Equations (1) and (2) can then be combined to give

$$\Pr ob\left[y_{it} = 1 \mid y_{it-1}, \mathbf{x}_{it}, \mathbf{z}_{i1}, \alpha_i\right] = \Phi\left[(\gamma y_{it-1} + \mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i) * (1 - D_{it}) + \left(\mathbf{z}_{i1}'\boldsymbol{\lambda} + \theta\alpha_i\right) * D_{it}\right] \tag{4}$$

This can be rewritten as

$$\Pr ob\left[y_{it} = 1 \mid y_{it-1}, \mathbf{x}_{it}, \mathbf{z}_{i1}, \alpha_i\right] = \Phi\left[\gamma(1 - D_{it}) y_{it-1} + (1 - D_{it})\mathbf{x}_{it}'\boldsymbol{\beta} + D_{it}\mathbf{z}_{i1}'\boldsymbol{\lambda} + \left\{1 + (\theta - 1)D_{it}\right\}\alpha_i\right] \tag{5}$$

This can be viewed as a standard random effects specification, but with a heteroskedastic factor loading for the random effect in period 1. Software that allows this form of heteroskedasticity, such as the gllamm program in Stata, can be used to estimate (5).

Alternatively the model can be viewed as a constrained random coefficients model, by rewriting it again as

$$\Pr ob\left[y_{it} = 1 \mid y_{it-1}, \mathbf{x}_{it}, \mathbf{z}_{i1}, \alpha_i\right] = \Phi\left[\alpha_i + \gamma(1 - D_{it}) y_{it-1} + (1 - D_{it})\mathbf{x}_{it}'\boldsymbol{\beta} + D_{it}\mathbf{z}_{i1}'\boldsymbol{\lambda} + (\theta - 1)\alpha_i D_{it}\right] \tag{6}$$

Unobserved heterogeneity α_i can be thought of as a random intercept term in the model and in this formulation the coefficient on D can be viewed as a random coefficient, with a unit correlation with the random intercept, but a different variance. Software for estimating random coefficient models that allows this form of restriction can therefore also be used.

Orme's two-step estimator

This is in the spirit of Heckman's two-step procedure for addressing the issue of endogenous sample selection. Since the cause of the initial conditions problem is the correlation between the regressor y_{it-1} and the unobservable α_i, Orme (1997) uses an approximation to substitute α_i with another unobservable component that is uncorrelated with y_{it-1}. First write

$$\alpha_i = \delta\eta_i + w_i \tag{7}$$

where, η_i, and w_i are orthogonal by construction. Substitution for α_i in (1) gives

$$\mathrm{Pr}ob[y_{it}=1|y_{it-1},\ldots, y_{i1}, \boldsymbol{x_i}, w_i] = \Phi(\boldsymbol{x_{ij}}'\beta + \gamma y_{ij-1} + \delta\, \eta_i + w_i) \qquad t=2,\ldots, T_i \tag{8}$$

which has two unobserved components, η_i and w_i. Since $E(w_i|y_{i1})=0$ by construction, there is no initial conditions problem if we can take care of η_i in (8). Orme notes that under the assumption that $(\alpha_i\ ,\ \eta_i)$ are distributed as bivariate normal, $E(\eta_i|y_{i1}) = e_i$, where $e_i = (2y_{i1} - 1)\varphi(\lambda' z_{i1}) / \Phi(\{2y_{i1} - 1\}\lambda' z_{i1})$ is the generalised error (inverse Mill's ratio) from the first period probit equation similar to (2), analogous to that used in Heckman's sample selection procedure, and ϕ and Φ are the Normal density and distribution functions respectively. Hence we can estimate (8) as a RE probit model using standard software with η_i replaced with an estimate of e_i after the estimation of (2). A potential problem is that, although $E(w_i|y_{i1})=0$, the conditional variance of w_i is not constant, but depends upon the correlation between α_i and η_i. However, Orme shows that the approximation works reasonably well even when this correlation is fairly different from zero.

Wooldridge's Conditional ML estimator

The Heckman estimator approximates the joint probability of the full observed y sequence. Wooldridge (2005) has proposed an alternative Conditional Maximum Likelihood (CML) estimator that considers the distribution conditional on the initial period value (and exogenous variables). Instead of specifying a distribution for $\mathrm{Pr}ob(y_1|\alpha)$, Wooldridge specifies an approximation for Pr*ob*($\alpha|y_1$). Thus a specification such as the following is assumed in the case of the random effect probit,

$$a_i|y_{i1}, \boldsymbol{z_i} \sim \text{Normal}(\zeta_0 + \zeta_1 y_{i1} + \boldsymbol{z_i}'\zeta, \sigma^2_a) \qquad (9)$$

$$where\ \alpha_i = \zeta_0 + \zeta_1 y_{i1} + \boldsymbol{z_i}'\zeta + a_i \qquad (10)$$

$\boldsymbol{z_i}$ includes variables that are correlated with the unobservable α. The appropriate $\boldsymbol{z}$ may differ from that in the Heckman specification. The idea here is that the correlation between y_{i1} and α is handled by the use of (10) giving another unobservable individual-specific heterogeneity term a that is uncorrelated with the initial observation y_1. Wooldridge in fact specifies z_i to be x_i as in (1a) above (although only for periods 2 to T), but alternative specifications of it would also be possible.

Substituting (10) into (1) gives

$$Prob(y_{it}=1|a_i,y_{i1}) = \Phi[\, \boldsymbol{x_{ij}}'\beta + \gamma y_{ij-1} + \zeta_1 y_{i1} + \boldsymbol{z_i}\,\zeta + a_i\,] \qquad t=2,\ldots, T_i \qquad (11)$$

In this model, the contribution to the likelihood function for individual *i* is given by

$$L_i = \int \left(\prod_{t=2}^{T_i} \Phi\left[\left(\mathbf{x}_{it}'\boldsymbol{\beta} + \gamma y_{it-1} + \varsigma_1 y_{i1} + \mathbf{z_i}'\varsigma + a_i \right)\left(2y_{it} - 1\right) \right] \right) g(\mathrm{a}_i) d\mathrm{a}_i \qquad (12)$$

where g(a) is the normal probability density function of the *new* unobservable individual-specific heterogeneity given in (9). Since this is the standard random effects probit model likelihood contribution for individual *i* one can proceed with the maximisation using standard software. Note that if x_i is used for z_i this means the Wooldridge estimator for the uncorrelated random effects specification and for the Mundlak correlated random effects specification are the same, since x_i is already included in the model to be estimated.

3. Empirical Illustration

The empirical illustration uses data from the first six waves of the British Household Panel Survey (BHPS), covering the period 1991-1996, to examine the unemployment dynamics of

British men.[1] The data used are a sub-sample of those used in Stewart (2007). The sample is restricted to those who were in the labour force (employed or unemployed) at each of the six waves. The ILO/OECD definition of unemployment is used, under which a man is unemployed if he does not have a job, but had looked for work in the past four weeks and is available for work.

Results for different estimators for a model for the probability of unemployment of the form of equation (1) above are given in Table 1. Column [1] gives the pooled probit estimates. Additional education, more labour market experience and being married reduce the probability of unemployment. Being in poor health or living in a travel to work area with a high unemployment-vacancy ratio raise the probability. Being unemployed at *t-1* strongly increases the probability of being unemployed at *t*.

Column [2] gives the equivalent standard random effects probit estimates, treating lagged unemployment as exogenous. The coefficients on all the *x*-variables are increased, while that on y_{t-1} is reduced relative to the pooled probit estimates. However the random effects probit and pooled probit models involve different normalizations. To compare coefficients those from the random effects estimator need to be multiplied by an estimate of $\surd(1- \rho)$, where ρ is the constant cross-period error correlation (see Arulampalam, 1999). The scaled coefficient estimate on unemployment at *t-1* in column [2] is 1.35. Compared with the pooled probit estimator, the estimate of γ is reduced by a quarter in the random effects model, but remains strongly significant.

The corresponding results for the Heckman estimator are given in column [3], with the initial period equation including two exogenous pre-labour market instruments and the full set of period-specific versions of the time-varying *x*-variables. (Only the married, poor health and local unemployment-vacancy ratio variables are treated as time-varying. There are

[1] The BHPS contains a nationally representative sample of households whose members are re-interviewed each year. The sample used here contains only Original Sample Members, is restricted to those aged 18-64 and excludes full-time students.

very few changes in the years of education variable in the sample.) The estimate of θ is 0.88, significantly greater than zero, rejecting the exogeneity of the initial conditions. (In fact θ is insignificantly different from 1.) Compared to the random effects estimator treating the initial conditions as exogenous, the Heckman estimator shows a fall in the estimate of γ of about a third and a near doubling in the estimate of ρ. In terms of scaled coefficient estimates, $\gamma(1-\rho)^{½}$, the standard random effects probit with the initial conditions treated as exogenous gives 1.35, while the Heckman estimator gives 0.79.

The Orme two-step estimates for the same model are given in column [4]. The estimator uses two exogenous pre-labour market instruments in conjunction with x_{it} for all time periods in z_i in the initial period equation as in the Heckman estimator. Relative to the Heckman estimator, the Orme estimator gives a slightly higher estimate of γ: 1.11 compared with 1.05 and a slightly lower estimate of ρ: 0.35 compared with 0.43.

The corresponding Wooldridge CML estimates are given in column [5]. The equation estimated contains x_{it} for all time periods. This gives an estimate of γ of 1.06, between the other two estimates and close to the Heckman estimate, and an estimate of ρ of 0.36, also between the other two estimates and close to the Orme estimate. In terms of scaled coefficient estimates, $\gamma(1-\rho)^{½}$, the Wooldridge estimator gives 0.88, about half way between 0.79 for the Heckman estimator and 0.89 for the Orme estimator. However all three of these estimates are fairly close together. The Wooldridge estimates of the elements of β corresponding to education, experience and the local unemployment/vacancy ratio are fairly similar to those from the other estimators. However this is not the case for the coefficients on married and health limits. The latter is cut by about half, the former by about two-thirds. Their standard errors are also appreciably higher than for the other estimators and both are insignificantly different from zero with this estimator. The reason for this is seen in the next paragraph.

Estimates for the corresponding correlated random effects model, using the Mundlak specification resulting in equation (1a), are given in Table 2. This results in the full set of period-specific versions of the time-varying x-variables being added to the main equation. Recall that the Wooldridge estimator is the same in both cases. The estimates of γ using the Heckman and Orme estimators both fall slightly when this specification is used. The estimates of the coefficients on education and experience are little changed, but those on the (time-varying) married and health limits variables fall considerably and now match closely those from using the Wooldridge estimator.

As indicated above, other specifications of both the z-vector and the relationship between α and the x-variables have been proposed and can be used as alternatives. However the contenders considered here have little effect on the estimates in Tables 1 and 2. To illustrate, using only x_{i1} rather than the whole of x_i in the initial period equation (in addition to the two exogenous pre-labour market instruments) reduces the Heckman estimate of γ in Table 1 from 1.048 to 1.047 and increases the estimate of ρ from 0.430 to 0.433. Replacing the full x_i by the time means changes the estimate of γ to 1.049 and that of ρ to 0.431. Similar very small differences are found for the elements of β, for the other estimators and for the correlated random effects estimates in Table 2.

4. Simulation Illustration

In this section we present a small-scale simulation experiment, to provide a comparison of the estimators in a situation where the true values of the parameters are known. The analysis is limited in scope and considers a specification that makes assumptions favourable to the Heckman estimator.

The set-up for the experiments is as follows. Data is generated for 3,000 observations (N=500, T=6). The model used to generate the data is

$$y_{it} = \mathbf{1}[(\gamma\, y_{it-1} + \beta_0 + \beta\, x_{it} + \alpha_i + u_{it}) > 0] \qquad \text{for } t=2,\ldots,T$$

$$y_{i1} = \mathbf{1}[(\pi_0 + \pi_1\, x_{i1} + \pi_2\, z_i + \theta\, \alpha_i + u_{i1}) > 0]$$

Thus the approximation used in the Heckman estimator is assumed to be the data generating process. The start of the stochastic process is assumed to coincide with the start of the sample period. The objective of the analysis in this section is to examine the situation where the Heckman specification does not involve an approximation for the model for the initial period, and ask how the finite sample performance of the Orme and Wooldridge estimators compares with that of the Heckman estimator in this particular case.

The covariate x is generated as follows. $x_{i1}^* = \chi_2^2/2$, $x_{it}^* = 0.6\, x_{it-1}^* + 0.8\, N(0,1)$ for $t \geq 2$, $x = 0.5\, x^* + 2.5\, \mathbf{1}[x^* < 0]$. The variable z is generated as a standard uniform. Guided by the indications of the empirical illustration in the previous section, the average of x over the six time periods is included both in the main equation to allow for correlated random effects and in the initial period instrument set. The individual effect α is generated in the experiments as $N(0, \sigma_\alpha^2)$ and u as N(0, 1). The inter-period error correlation is therefore given by $\rho = \sigma_\alpha^2/(1+\sigma_\alpha^2)$. Hence $\sigma_\alpha^2 = \rho/(1-\rho)$. Different experiments are conducted for different values of γ, β, ρ and θ. The following parameters are fixed in all experiments: $\beta_0 = -2$, $\pi_0 = -1$, $\pi_1 = 1.5$, $\pi_2 = 0.5$. Each of the experiments reported in Tables 3 and 4 is based on 100 Monte Carlo replications.

Table 3 gives the relative average bias and relative root mean square error (both in percentage terms and both relative to the true value) for the estimates of γ and β using the Heckman, Wooldridge and Orme estimators in each of the first set of experiments. In the base experiment the main parameter values are set to $\gamma = 1.2$, $\beta = 1.0$, $\rho = 0.4$ and $\theta = 0.8$. This implies an estimate of σ_α of 0.816, similar to the standard deviation of x (and hence of βx in this experiment) of 0.862. The relative bias in this experiment for all three estimators is fairly small. The Heckman and Orme estimators give an absolute relative bias in the estimate of γ of just under 2%. The Wooldridge estimator dominates these two for γ with a relative bias of only 0.2%. The relative bias in the estimate of β is less than 1% for all three estimators, but the Heckman estimator has the smallest relative bias for β. The root mean square errors for both γ and β are fairly similar for all three estimators.

The remaining experiments in Table 3 (experiments 2 – 6) retain the same values of γ, β and θ and examine various values of ρ. These experiments shift the balance between the variance of the individual-specific effect, α, and the exogenous variation in the specification (that in βx). In the base experiment, the ratio of the variance of α to the variance of βx is 0.9. Experiment 2 lowers ρ to 0.3, which reduces this ratio of variances to 0.6. The other four experiments in Table 3 increase ρ and hence this variance ratio. Experiments 3 – 6 use values of ρ of 0.5, 0.6, 0.7 and 0.8, giving values of this variance ratio of 1.3, 2.0, 3.2 and 5.4 respectively.

For all these values of ρ the Heckman estimator gives the smallest relative bias for the estimate of β. The Wooldridge estimator gives the largest relative bias and the Orme estimator lies between them in these terms. It is also clear that in broad terms relative bias rises as ρ does for all three estimators (although this is not quite monotonic). The root mean square error for the estimate of β is fairly similar for the three estimators in each case.

The relative biases for the estimate of γ present a somewhat different picture. For low values of ρ the Wooldridge estimator continues to give the smallest relative bias of the three estimators (as it did in the base experiment). However in broad terms this relative bias rises with ρ, which is not the case for the Heckman and Orme estimators. As a result it is dominated by the Heckman estimator for $\rho = 0.7$ and the Orme estimator for both $\rho = 0.7$ and $\rho = 0.8$. There is however a large rise in the relative bias of the Heckman estimator when ρ is increased from 0.7 to 0.8. For this latter experiment, the Heckman estimator of γ has a relative bias of 5%, compared with 2.5% for the Wooldridge estimator and 0.2% for the Orme estimator.

In broad terms the relative bias of the Orme estimator improves as ρ is increased. This is rather surprising, since the Orme estimator is based on an approximation around $\rho = 0$. One should of course be cautious about reading too much into a single set of simulation experiments.

The experiments reported in Table 4 use modifications to the values of γ, β and θ used in the data generating process. Experiment 7 doubles the value of β used compared to the base experiment: from 1.0 to 2.0. The performance of all three estimators deteriorates for both γ and β. When the value of β is halved in experiment 8, this is not universally the case. Experiments 9 and 10 increase and decrease γ respectively, while experiments 11 and 12 increase and decrease θ respectively. In these seven experiments retaining $\rho = 0.4$ (including the base experiment), the Wooldridge estimator of γ dominates the other two in six cases out of seven in terms of relative bias. The Heckman estimator has the largest relative bias (in absolute terms) in five of these seven experiments. In contrast to this, for the estimation of β, the Heckman estimator has the smallest relative bias (in absolute terms) in four of the seven experiments and the Wooldridge estimator the largest in five of the seven experiments. Clearly no estimator dominates the others overall in this set of experiments.

Of the six experiments based on increasing or decreasing γ, β or θ relative to the base experiment, the increase of β to 2.0 in experiment 7 seems to have the largest impact overall. The final experiment therefore examines the effect of this in conjunction with increasing ρ from 0.4 to 0.8 (i.e. two changes relative to the base experiment). Relative to experiment 7, the increase in ρ worsens the Heckman and Wooldridge estimates of γ and the Wooldridge and Orme estimates of β in relative bias terms. It increases the root mean square error for all three estimators of both parameters. Relative to experiment 6, the increase in β worsens the

Heckman and Orme estimates of γ and all three estimates of β in relative bias terms. It increases the root mean square error for the Wooldridge and Orme estimators of both parameters. In this final experiment the Wooldridge estimator has the largest relative bias and root mean square error for both γ and β.

Judged across the full set of experiments conducted, none of the three estimators dominates the other two in all cases, or even in a majority of cases.

5. Conclusions

This paper presents a convenient shortcut method for implementing the Heckman estimator of the dynamic random effects probit model using standard software. This removes the need for separate programming and puts this estimator on a similar footing to the simpler estimators suggested by Orme and Wooldridge based on alternative approximations. The choice between these estimators can therefore be based on performance rather than availability of ease of use. An empirical illustration has been presented in section 3 and a set of simulation experiments in section 4. The former suggests that it is advantageous to allow for correlated random effects using the approach of Mundlak (1978), but that once this is done, the three estimators provide similar results. The simulation experiments suggest that none of the three estimators dominates the other two in all cases.

TABLE 1: Unemployment probability model: Alternative estimators

	[1] Probit	[2] RE probit	[3] Heckman	[4] Orme	[5] Wooldridge
Unem(t-1)	1.837	1.536	1.048	1.107	1.062
	[0.074]	[0.122]	[0.130]	[0.115]	[0.115]
Education	-0.043	-0.050	-0.058	-0.054	-0.055
	[0.011]	[0.014]	[0.017]	[0.016]	[0.017]
Experience	-0.048	-0.068	-0.072	-0.064	-0.066
	[0.030]	[0.037]	[0.045]	[0.043]	[0.045]
Married	-0.186	-0.236	-0.309	-0.280	-0.092
	[0.066]	[0.082]	[0.100]	[0.093]	[0.227]
Health limits	0.429	0.503	0.585	0.569	0.289
	[0.093]	[0.114]	[0.133]	[0.126]	[0.185]
Local u/v	0.654	0.849	0.941	0.919	0.880
	[0.229]	[0.268]	[0.306]	[0.292]	[0.396]
λ				0.459	
				[0.076]	
Unem(1)					1.016
					[0.161]
ρ		0.225	0.430	0.354	0.357
		[0.065]	[0.063]	[0.044]	[0.043]
θ			0.882		
			[0.189]		
LogL	-1052.00	-1044.81	-1341.14	-1024.24	-1014.01

Estimators:
[1] Pooled Probit
[2] Standard Random Effects Probit (takes initial condition to be exogenous)
[3] Heckman estimator, with x in all periods and 2 exog instruments in initial period equation
[4] Orme estimator, with x in all periods and 2 exog instruments in initial period equation
[5] Wooldridge estimator, with x in all periods included in z

Notes:
1. Sample size = 10,092.
2. LogL in [3] is for joint model for all periods. Those in other columns are for 2-T only.

TABLE 2: Unemployment probability model: Alternative estimators with Mundlak correction for correlated individual effects

	[1] Probit	[2] RE probit	[3] Heckman	[4] Orme	[5] Wooldridge
Unem(t-1)	1.811	1.500	1.009	1.074	1.062
	[0.075]	[0.124]	[0.130]	[0.115]	[0.115]
Education	-0.044	-0.052	-0.060	-0.056	-0.055
	[0.012]	[0.015]	[0.018]	[0.017]	[0.017]
Experience	-0.050	-0.072	-0.077	-0.070	-0.066
	[0.031]	[0.040]	[0.048]	[0.045]	[0.045]
Married	-0.041	-0.063	-0.095	-0.090	-0.092
	[0.194]	[0.212]	[0.231]	[0.226]	[0.227]
Health limits	0.211	0.254	0.299	0.287	0.289
	[0.158]	[0.174]	[0.189]	[0.185]	[0.185]
Local u/v	0.633	0.900	0.896	0.873	0.880
	[0.338]	[0.378]	[0.406]	[0.396]	[0.396]
λ				0.469	
				[0.076]	
Unem(1)					1.016
					[0.161]
ρ		0.232	0.439	0.357	0.357
		[0.066]	[0.063]	[0.044]	[0.043]
θ			0.885		
			[0.189]		
LogL	-1044.03	-1044.81	-1332.14	-1015.40	-1014.01

Notes:
1. Estimators as in Table 1 with x in all periods added to main equation.
2. Sample size = 10,092.
3. LogL in [3] is for joint model for all periods. Those in other columns are for 2-T only.

TABLE 3: Simulation Results: Base Experiment and Variations in ρ

Experiment	Estimator	γ Relative Bias (%)	γ Relative RMSE (%)	β Relative Bias (%)	β Relative RMSE (%)
1) Base	Heckman	-1.850	7.664	0.073	5.668
	Wooldridge	0.218	7.483	0.752	5.705
	Orme	-1.733	7.634	0.152	5.669
2) $\rho = 0.3$	Heckman	-1.563	7.426	0.129	5.165
	Wooldridge	0.255	7.389	0.581	5.200
	Orme	-1.504	7.404	0.167	5.165
3) $\rho = 0.5$	Heckman	-1.700	7.620	0.475	5.962
	Wooldridge	0.575	7.521	1.330	6.094
	Orme	-1.522	7.598	0.610	5.967
4) $\rho = 0.6$	Heckman	-1.259	8.396	0.565	6.279
	Wooldridge	1.026	8.353	1.700	6.562
	Orme	-1.127	8.346	0.870	6.379
5) $\rho = 0.7$	Heckman	-0.012	10.083	0.331	6.809
	Wooldridge	1.614	10.245	2.083	7.147
	Orme	-0.483	10.010	1.098	6.906
6) $\rho = 0.8$	Heckman	4.881	10.873	-1.141	7.363
	Wooldridge	2.453	10.551	2.740	7.973
	Orme	0.188	10.011	1.934	7.921

Notes:

1. Data generating process for base experiment (experiment 1) has γ=1.2, β=1.0, ρ=0.4, θ=0.8 in addition to the specification given in the text.
2. Data generating processes for experiments 2 – 6 as for base experiment except for specified value of ρ given in Column 1.
3. 100 Monte Carlo replications used in each experiment.

TABLE 4: Simulation Results: Additional Experiments

Experiment	Estimator	γ Relative Bias (%)	γ Relative RMSE (%)	β Relative Bias (%)	β Relative RMSE (%)
1) Base	Heckman	-1.850	7.664	0.073	5.668
	Wooldridge	0.218	7.483	0.752	5.705
	Orme	-1.733	7.634	0.152	5.669
7) β = 2.0	Heckman	-2.557	7.988	2.702	6.276
	Wooldridge	-0.309	7.648	2.415	6.037
	Orme	-2.569	8.017	2.561	6.186
8) β = 0.5	Heckman	-1.837	8.265	0.897	9.541
	Wooldridge	0.040	8.246	2.542	9.942
	Orme	-1.815	8.230	0.974	9.545
9) γ = 2.4	Heckman	-2.033	4.911	0.874	6.565
	Wooldridge	-0.133	4.586	0.836	6.654
	Orme	-1.979	4.899	0.926	6.585
10) γ = 0.6	Heckman	-1.755	13.569	-0.369	5.250
	Wooldridge	1.504	13.523	0.498	5.312
	Orme	-1.550	13.534	-0.288	5.253
11) θ = 1.2	Heckman	-1.102	7.713	0.265	5.651
	Wooldridge	1.285	7.803	1.171	5.764
	Orme	-0.877	7.700	0.436	5.669
12) θ = 0.4	Heckman	-1.753	7.419	0.047	5.514
	Wooldridge	-0.117	7.313	0.285	5.511
	Orme	-1.691	7.416	0.061	5.511
13) β = 2.0	Heckman	3.787	10.647	2.647	7.084
and ρ = 0.8	Wooldridge	4.459	11.433	5.259	8.866
	Orme	1.558	10.521	4.854	8.532

Notes:
1. Data generating process for base experiment (experiment 1) has γ=1.2, β=1.0, ρ=0.4, θ=0.8 in addition to the specification given in the text.
2. Data generating processes for experiments 7 – 11 as for base experiment except for specified parameter values given in Column 1.
3. 100 Monte Carlo replications used in each experiment.

References

Arulampalam, W. (1999) A Note on estimated effects in random effect probit models", *Oxford Bulletin of Economics and Statistics*, 61(4), 597-602.

Arulampalam, W. and Bhalotra, S. (2006) Sibling death clustering in India: genuine scarring vs unobserved heterogeneity, *Journal of the Royal Statistical Society Series A*, 169, 829-848.

Butler, J. S. and Moffitt, R. (1982) A computationally efficient quadrature procedure for the one-factor multinomial probit model, *Econometrica*, 50, 761-4.

Heckman, J. J. (1981a) Heterogeneity and state dependence, in S. Rose (ed.), *Studies in Labor Markets*, Chicago Press, Chicago, IL.

Heckman, J. J. (1981b). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process, in C. F. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, MA, 114-178.

Mundlak, Y (1978). On the pooling of time series and cross section data. *Econometrica*, 46, 69-85.

Orme, C. D. (1997) The initial conditions problem and two-step estimation in discrete panel data models, mimeo, University of Manchester.

Orme, C. D. (2001) Two-step inference in dynamic non-linear panel data models, mimeo, University of Manchester.

Stewart, M. B. (2007) Inter-related dynamics of unemployment and low-wage employment, *Journal of Applied Econometrics*, 22, 511-531.

Wooldridge, J. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity, *Journal of Applied Econometrics*, 20, 39-54.