Dynamic Probit models for panel data: A comparison of three methods of estimation

Alfonso Miranda

Keele University and IZA
(A.Miranda@econ.keele.ac.uk)

2007 UK Stata Users Group meeting

September 10.

Motivation

- In a number of contexts researchers have to model a dummy variable y_{it} that is function of $y_{i,t-1}$ (unemployment, migration, health).
- A dynamic probit/logit model is needed.
- In the dynamic setup y_{i0} is likely to be correlated with unobserved heterogeneity u_i affecting y_{it}.
- If y_{i0} is taken as exogenous inconsistent estimators are obtained. This is know as the initial conditions problem.

Motivation

- Three methods of estimation have been suggested: Heckman (1981), Orme (1996), and Wooldridge (2002).
- Heckman's method is computer expensive – not anymore really – while the other two methods are computer inexpensive and easy to implement in conventional econometric software.
- No study has compared the relative performance of such methods with small and large samples, and with low and high correlation between unobservables affecting initial conditions and dynamic equations.

Heckman (1981) method

Heckman suggests to approximate the reduced form of the marginal probability of y_{i0} given u_i with a Probit model and to allow free correlation ρ between y_{i0} and y_{it}.

$$
\begin{aligned}
y_{it}^* &= \mathbf{z_{it}}\boldsymbol{\beta} + \gamma y_{i,t-1} + u_i + \varepsilon_{it} && (1)\\
y_{i0}^* &= \mathbf{x_{it}}\boldsymbol{\theta} + \delta u_i + \eta_{it} && (2)
\end{aligned}
$$

with $y_{it} = 1$ if $y_{it}^* > 0$ and zero otherwise. u_i, η_{it} and ε_{it} are all iid $N(0,1)$. Neither ε_{it} nor η_{it} are serially correlated.

- equations (1) and (2) are estimated as a system.
- Need to integrate out u_i against the density $\phi(u_i)$.
- May use ML + Gauss-Hermite quadrature or Maximum Simulated Likelihood.
- $\rho = \frac{\delta}{\sqrt{(2(\delta^2+1))}}$

Orme (1996) method

Orme suggests a two-step bias corrected procedure that is locally valid when ρ approximates to zero. Define,

$$y^*_{it} = \mathbf{z_{it}}\boldsymbol{\beta} + \gamma y_{i,t-1} + u_i + \varepsilon_{it} \quad (3)$$

$$y^*_{i0} = \mathbf{x_{it}}\boldsymbol{\theta} + \delta u_i + \eta_{it} \quad (4)$$

- Notice that in eq. (3) $E[u_i] = 0$ but $E[u_i|y_{i0}] \neq 0$ when $\delta \neq 0$ (that is, when $\rho \neq 0$).
- Correlation between u_i and y_{i0} can be removed by writing:

$$u_i = E[u_i|y_{i0}] + u^*_i$$

so that $E[u^*_i|y_{i0}] = 0$ by construction.

Orme (1996) method

- Can use, in a first step, a simple probit model for y_{i0} to estimate,

$$E[u|y_{i0}] = E\left[u_i|\delta u_i + \eta_{it} \geq -\mathbf{x_{it}}\boldsymbol{\theta}\right] = \frac{\phi\left(\mathbf{x_{it}}\boldsymbol{\theta}\right)}{\Phi\left(\mathbf{x_{it}}\boldsymbol{\theta}\right)}$$

- And in a second step estimate the dynamic equation using a standard RE probit that includes $E[u_i^*|y_{i0}]$ as a regressor,

$$y_{it}^* = \mathbf{x_{it}}\boldsymbol{\beta} + \gamma y_{i,t-1} + \sigma E[u_i|y_{i0}] + u_i^* + \varepsilon_{it} \qquad (5)$$

- Orme shows that this two-step procedure is locally valid if ρ approximates to zero and argues that the method can perform well even if ρ is 'high'.

Wooldridge (2002) method

$$y_{it}^{*} = \mathbf{x_{it}}\boldsymbol{\beta} + \gamma y_{i,t-1} + u_i + \varepsilon_{it} \quad (6)$$

$$y_{i0}^{*} = \mathbf{z_{it}}\boldsymbol{\theta} + \delta u_i + \eta_{it} \quad (7)$$

- Heckman does the following:

$$f\left(y_{i0}, \cdots, y_{iT}\right) = \int f\left(y_{i1}, \cdots, y_{iT} | y_{i0}, \mathbf{w_{it}}, u_i\right) h\left(y_{i0} | \mathbf{w_{it}}, u_i\right) g(u_i | \mathbf{w_{it}}) du_i$$

  with $\mathbf{w_{it}} = (\mathbf{x_{it}}, \mathbf{z_{it}})$ and use ML.

- Wooldridge suggests to model the distribution of $\{y_{i1}, \cdots, y_{iT}\}$ given y_{i0} and to use conditional ML.
- To do so one needs to specify the distribution for u_i given y_{i0} and other exog. variables:

$$f\left(y_{i1}, \cdots, y_{iT} | y_{i0}\right) = \int f\left(y_{i1}, \cdots, y_{iT} | y_{i0}, \mathbf{w_{it}}, u_i\right) g\left(u_i | y_{i0}, \mathbf{w_{it}}\right) du_i$$

- It is suggested the following approximation

$$g\left(u_i | y_{i0}, \mathbf{w_{it}}\right) \sim N\left(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{w}_i, \sigma_v^2\right)$$

In other words, we can write

$$u_i = \alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{w}_i + v_i \quad (8)$$

$$v_i \sim N(0, \sigma_v^2) \text{ and independent of } y_{i0}, w_i \quad (9)$$

- substituting (8) in (6)

$$y_{it}^* = \mathbf{z_{it}}\boldsymbol{\beta} + \gamma y_{i,t-1} + \alpha_1 y_{i0} + \alpha_2 \bar{w}_i + v_i + \varepsilon_{it} \quad (10)$$

and estimate (9) by standard RE probit.

The following model is specified:

$$y_{it}^{*} = 0.5 + 0.5z_{it} - 0.5y_{i,t-1} + u_i + \varepsilon_{it} \tag{11}$$

$$y_{i0}^{*} = 1x_{i0} - 1z_{i0} + \delta u_i + \eta_{it} \tag{12}$$

- Random draws from independent standard normal distributions are taken to generate z_{it} and x_{i0}. These variables remain fixed during all simulations.
- At each replication step random draws from independent standard normal distributions are taken to generate u_i, ε_{it} and η_{it}.
- At each iteration the model is estimated using Heckman (MSL with 400 halton draws), Wooldridge, and Orme methods. Estimates for the dynamic equation are kept.

- 1000 replications are taken.
- Various experiments are done comparing the performance of all these three methods using small, medium, and large samples and low and high ρ.
- At the end simulation statistics are calculated:
  - Average estimator (AE)
  - Percentage bias (PB)
  - Average standard error (ASE)
  - Standard error (SDE)
  - Mean square error (MSE)
  - Nominal coverage of 95% confidence intervals (Ncov).

T=3, n=100, N=300, rho=0


```
Number of panels        =  100
Obs per panel           =    3
Total Number of obs     =  300
Delta                   = 0.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.506	1.21	.14	.136	.019	.958
LDV	-.506	-1.14	.261	.25	.063	.958
_cons	.51	1.93	.221	.22	.048	.948
Wooldridge Method						
z	.5	.015	.168	.171	.029	.956
LDV	-.452	9.59	.36	.369	.138	.93
_cons	.494	1.13	.332	.352	.124	.926
Orme Method						
z	.502	.461	.148	.151	.023	.952
LDV	-.48	4.08	.352	.355	.127	.931
_cons	.488	-2.36	.326	.333	.111	.93

T=3, n=100, N=300, rho=0.5


```
Number of panels         =  100
Obs per panel            =    3
Total Number of obs      =  300
Delta                    = 1.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.505	1.04	.136	.13	.017	.966
LDV	-.505	-.969	.252	.238	.057	.965
_cons	.508	1.64	.214	.213	.045	.954
Wooldridge Method						
z	.417	-16.6	.162	.161	.033	.904
LDV	-.466	6.88	.371	.366	.135	.945
_cons	-.222	-144	.267	.277	.597	.232
Orme Method						
z	.412	-17.6	.118	.121	.023	.835
LDV	.162	132	.276	.334	.549	.362
_cons	-7e-3	-101	.266	.302	.348	.44

T=3, n=100, N=300, rho=0.7


```
Number of panels        =  100
Obs per panel           =    3
Total Number of obs     =  300
Delta                   = 10.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.509	1.78	.131	.123	.015	.962
LDV	-.497	.525	.237	.224	.05	.968
_cons	.508	1.58	.191	.199	.04	.942
Wooldridge Method						
z	.474	-5.16	.159	.157	.025	.943
LDV	-.564	-12.8	.421	.396	.161	.932
_cons	-.327	-165	.182	.189	.719	.022
Orme Method						
z	.389	-22.1	.101	.1	.022	.799
LDV	.558	212	.19	.223	1.17	3e-3
_cons	-.042	-108	.853	1.2	1.72	.849

T=3, n=300, N=900, rho=0

Motivation
3 Methods
Monte Carlo Study
Simulation results
Conclusions

```
Number of panels        =  300
Obs per panel           =    3
Total Number of obs     =  900
Delta                   = 0.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.505	.941	.077	.078	6e-03	.948
LDV	-.492	1.54	.147	.142	.02	.962
_cons	.497	-.587	.126	.12	.014	.961
Wooldridge Method						
z	.488	-2.49	.09	.088	8e-3	.947
LDV	-.399	20.3	.197	.205	.052	.904
_cons	.46	-7.9	.185	.193	.039	.928
Orme Method						
z	.491	-1.83	.081	.082	7e-3	.931
LDV	-.436	12.8	.195	.198	.043	.922
_cons	.452	-9.67	.179	.18	.035	.928

T=3, n=300, N=900, rho=0.5


```
Number of panels        =  300
Obs per panel           =    3
Total Number of obs     =  900
Delta                   = 1.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.504	.88	.075	.076	6e-3	.948
LDV	-.493	1.34	.142	.135	.018	.964
_cons	.497	-.637	.122	.116	.014	.964
Wooldridge Method						
z	.421	-15.9	.088	.089	.014	.823
LDV	-.442	11.6	.21	.231	.057	.938
_cons	-.225	-145	.153	.153	.549	7e-3
Orme Method						
z	.401	-19.9	.068	.07	.015	.62
LDV	.209	142	.167	.207	.545	.112
_cons	-.048	-110	.155	.177	.332	.174

T=3, n=300, N=900, rho=0.7


```
Number of panels       =  300
Obs per panel          =    3
Total Number of obs    =  900
Delta                  = 10.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.506	1.22	.071	.07	5e-3	.957
LDV	-.49	2	.133	.127	.016	.955
_cons	.497	-.53	.109	.108	.012	.949
Wooldridge Method						
z	.472	-5.58	.088	.086	8e-3	.928
LDV	-.517	-3.46	.267	.245	.06	.924
_cons	-.33	-166	.103	.1	.699	0
Orme Method						
z	.399	-20.1	.058	.06	.014	.567
LDV	.575	215	.109	.126	1.17	0
_cons	-.27	-154	.555	.796	1.23	.58

T=3, n=3000, N=9000, rho=0


```
Number of panels       = 3000
Obs per panel          =    3
Total Number of obs    = 9000
Delta                  = 0.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.5	-.024	.023	.022	5e-4	.962
LDV	-.501	-.176	.046	.046	2e-3	.951
_cons	.501	.134	.039	.039	1e-3	.948
Wooldridge Method						
z	.493	-1.38	.028	.026	7e-4	.959
LDV	-.464	7.23	.069	.063	5e-3	.939
_cons	.483	-3.3	.061	.06	4e-3	.944
Orme Method						
z	.493	-1.48	.025	.024	6e-4	.95
LDV	-.469	6.12	.065	.059	4e-3	.944
_cons	.477	-4.64	.06	.055	3e-3	.946

T=3, n=3000, N=9000, rho=0.5


```
Number of panels       = 3000
Obs per panel          =    3
Total Number of obs    = 9000
Delta                  = 1.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.5	.059	.023	.021	4e-4	.968
LDV	-.5	-.063	.045	.044	2e-3	.955
_cons	.5	.049	.038	.038	1e-3	.951
Wooldridge Method						
z	.419	-16.3	.026	.03	7e-3	.163
LDV	-.415	16.9	.062	.101	.017	.486
_cons	-.218	-144	.047	.047	.517	0
Orme Method						
z	.397	-20.7	.022	.025	.011	.02
LDV	.245	149	.06	.085	.562	0
_cons	-.081	-116	.054	.07	.342	0

T=3, n=3000, N=9000, rho=0.7


```
Number of panels        = 3000
Obs per panel           =    3
Total Number of obs     = 9000
Delta                   = 10.00
```

	AE	PB	ASE	SDE	MSE	Ncov
Heckman Method						
z	.501	.156	.022	.02	4e-4	.966
LDV	-.499	.294	.042	.041	2e-3	.951
_cons	.499	-.234	.034	.034	1e-3	.945
Wooldridge Method						
z	.472	-5.52	.027	.026	1e-3	.84
LDV	-.545	-8.93	.095	.084	9e-3	.938
_cons	-.328	-166	.033	.033	.687	0
Orme Method						
z	.403	-19.4	.018	.017	8e-3	0
LDV	.571	214	.034	.04	1.15	0
_cons	-.353	-171	.149	.157	.753	0

Conclusions

- Heckman's method delivers estimators that are hardly subject to bias and that are estimated with high precision.
- The methods suggested by Wooldridge and Orme (W&O) deliver estimators that can be subject to substantial bias and low precision.
- W&O: The bias does not seem to decrease as sample size (number of panels n) increases.
- W&O: The bias increases when ρ gets higher.
- Nominal coverage of confidence intervals is satisfactory in Heckman's method but can be extremely bad in the case of W&O when ρ is high.

Conclusions

- Evidence suggest that Heckman's method offers substantial advantages.
- Today Heckman's method is not really computer expensive anymore (can use MSL and BHHH algorithm to speed the process).