JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. (2015)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jae.2434

REPLICATION SECTION

SIMULATION ESTIMATION OF TWO-TIERED DYNAMIC PANEL TOBIT MODELS WITH AN APPLICATION TO THE LABOUR SUPPLY OF MARRIED WOMEN: A COMMENT

ZHOU XUN[a,b] AND MICHEL LUBRANO[a,c*]
[a] *AMSE, Aix–Marseille University, France*
[b] *GREQAM, Centre de la Vieille Charité, Marseille, France*
[c] *GREQAM-CNRS, EHESS, Centre de la Vieille Charité, Marseille, France*

SUMMARY
We find that the empirical results reported in Chang (*Journal of Applied Econometrics* 2011; **26**(5): 854–871) are contingent on the specification of the model. The use of Heckman's initial conditions combined with observed and not latent lagged dependent variables leads to a counter-intuitive estimation of the true state dependence. The use of Wooldridge's initial conditions together with the observed lagged dependent variable and a proper modelling of censoring provides a much more natural estimate of the true state dependence parameters together with a clearer interpretation of the decision to participate in the labour market in the two-tiered model.

Received 19 September 2013; Revised 17 November 2014

1. INTRODUCTION

Chang (2011b) proposed a computationally practical simulation estimator for the two-tiered dynamic panel Tobit model. Estimation is undertaken using the Geweke–Hajivassiliou–Keane (GHK) simulator. The one-tiered dynamic Tobit model for panel data and autocorrelated errors is used, first, for modelling the rich dynamic structure of the labour force participation decision of married women and, second, for modelling the number of working hours. This initial model is written as

$$y_{it}^* = y_{i,t-1}\lambda + x_{it}\beta + \bar{x}_{i.}\omega + \epsilon_{it} \tag{1}$$

$$y_{it} = \max\{y_{it}^*, 0\} \tag{2}$$

with the following error structure:

$$\epsilon_{it} = d_i + \nu_{it} \tag{3}$$

* Correspondence to: Michel Lubrano, GREQAM-CNRS, EHESS, Centre de la Vieille Charité, Marseille, France. E-mail: michel.lubrano@univ-amu.fr

$$\nu_{it} = \zeta \nu_{i,t-1} + u_{it} \tag{4}$$

The two-tiered structure implies that the probability of participating ($\Pr(y_{it}^* > 0)$) is computed with a first set of parameters (λ_1, β_1, ω_1), while the number of hours worked (i.e. the conditional expectation of y_{it}), conditioned on the decision of participating, is determined by a second set of parameters ($\lambda_2, \beta_2, \omega_2$). The other parameters (error variances of d_i and u_{it} and ζ) are common to the two decisions.

Unfortunately, we encountered a number of difficulties in reproducing the estimates reported in Chang (2011b). Moreover, the true state dependence parameter λ had a negative value (leading to oscillations) which is hard to interpret. Following Heckman (1981b), individuals having experienced an event in the past (e.g. working) are more likely to experience that same event in the future (e.g. still working), implying a positive λ. True state dependence has to be disentangled from spurious state dependence which corresponds solely to the individual propensity and which is measured by ζ, the residual autocorrelation parameter. The purpose of this paper is to re-estimate Chang's model with different likelihood specifications (different initial conditions) and a possibly more efficient optimization method in order to recover a more satisfactory measure of true state dependence.

2. LIKELIHOOD FUNCTION AND INITIAL CONDITIONS

Let us define the indicator function I_{it}:

$$I_{it} = 1 \text{ if } y_{it}^* > 0, \text{and zero otherwise} \tag{5}$$

The likelihood contribution with fixed initial conditions for individual i is

$$L_i = \frac{1}{R}\sum_{r=1}^{R}\prod_{t=1}^{T}[f^{(r)}(y_{it}|y_{i,t-1}, d_i, x_{it})]^{I_{it}} \times [P^{(r)}(I_{it} = 0|y_{i,t-1}, d_i, x_{it})]^{1-I_{it}} \tag{6}$$

The latent variables are simulated according to

$$y_{it}^{*(r)} = x_{it}'\beta + \lambda y_{i,t-1} + \bar{x}_{i.}'\omega + A_{t.}(\psi)\eta_{i.}^{(r)}(\psi) \tag{7}$$

where ψ represents the set of all the parameters. The GHK simulator is used to simulate the R replications of η_{it} recursively. The initial conditions can be modelled in different ways. The simplest solution is to suppose that they are fixed and observed, allowing the likelihood function above to be straightforwardly maximized. However, this will tend to overstate the true degree of state dependence λ at the expense of the autocorrelation coefficient ζ as noted, for instance, in Stewart (2006) for the dynamic probit model. The other solution is to suppose that the initial conditions are random and correlated with the individual effects d_i. Heckman (1981a) proposed modelling the random initial conditions using an approximation to the reduced form of the model, which can be written

$$y_{i0}^* = z_{i0}'\pi + \theta\bar{x}_{i.}'\omega + \theta d_i + u_{i0} \tag{8}$$

where z_{i0} includes all the exogenous variables plus at least one instrumental variable. Then we can write and maximize the completed likelihood contribution:

$$\begin{aligned} L_i = \frac{1}{R}\sum_{r=1}^{R}\prod_{t=1}^{T}&\left[f^{(r)}(y_{it}|y_{i,t-1}, d_i, x_{it})\right]^{I_{it}} \times \left[P^{(r)}(I_{it} = 0|y_{i,t-1}, d_i, x_{it})\right]^{1-I_{it}} \\ &\times \left[f^{(r)}(y_{i0}|d_i, z_{i0})\right]^{I_{i0}} \times \left[P^{(r)}(I_{i0} = 0|d_i, z_{i0})\right]^{1-I_{i0}} \end{aligned} \tag{9}$$

DOI: 10.1002/jae

Chang (2011b) reports results using Heckman's approach for the initial conditions, but without detailing their precise specification. In particular, he does not mention either the scaling parameter θ or which instrumental variable has been used. Moreover, when using Heckman's initial conditions, it is more natural to use the lagged latent variable and not the observed one because the reduced form concerns the former.

Wooldridge (2005) has proposed a much simpler solution to the treatment of initial conditions in dynamic Tobit models that is much more parsimonious in term of extra parameters and which needs no instrumental variables. Instead of completing the conditional likelihood function by using $f(y_{i0}|c_i, \bar{x}_i)$, Wooldridge (2005) proposes to specify $f(c_i|y_{i0}, \bar{x}_i)$. The method is simple for probit models as the lagged binary outcomes are of the same nature, whatever the form of censoring. For the Tobit model, things are different, as underlined by Wooldridge (2005, example 2, pp. 41–42), as the lagged outcomes have a different scaling depending on censoring. To cope with that, Wooldridge introduces a $\mathbf{g}(y_{t-1})$ function, which gives the following equivalent specification of our model:

$$c_i = \bar{x}_i'\omega + d_i + y_{i0}\delta_1 \mathbb{I}(y_{i0} > 0) + \delta_2 \mathbb{I}(y_{i0} = 0) \tag{10}$$

As the lagged endogenous variable is treated as observed, the formulation of Wooldridge (2005, p. 49) leads to the adoption of two different λs when censored or not censored:

$$y_{it}^* = x_{it}'\beta + y_{i,t-1}\lambda_1 \mathbb{I}(y_{i,t-1} > 0) + \lambda_2 \mathbb{I}(y_{i,t-1} = 0) + c_i + u_{it} \tag{11}$$

For the two-tiered model, the c_i are extended as follows. For the participation equation, we have

$$c_{i1} = \bar{x}_i'\omega_1 + d_i + y_{i0}\delta_{11} \mathbb{I}(y_{i0} > 0) + \delta_{12} \mathbb{I}(y_{i0} = 0) \tag{12}$$

while for the working hours equation

$$c_{i2} = \bar{x}_i'\omega_2 + d_i + y_{i0}\delta_{21} \mathbb{I}(y_{i0} > 0) + \delta_{22} \mathbb{I}(y_{i0} = 0) \tag{13}$$

With four different δs, we allow for a better modelling of the influence of the initial conditions.

3. CHECKING THE EMPIRICAL RESULTS

We have reprogrammed (with the free software R) the one-tiered and two-tiered dynamic panel Tobit models using the indications in Chang (2011b) completed by Chang (2011a) and Stewart (2006). We first focus on the one-tiered version, which corresponds to the second column of Table III in Chang (2011b, p. 866). Following the author's indications, we used 10 draws of the uniform random variable ξ_{it}, which is the basic ingredient in the GHK simulator used to dynamically generate the truncated error term of the model. We have to keep the same numbers for ξ_{it} in order to make results comparable between the different models. We have tried three different options for the initial conditions. First, we implemented Heckman's random initial conditions based on a reduced form. That means using an instrumental variable for the initial state. We used the years of education of the husband. Then, we used fixed initial conditions as a point of comparison. We finally implemented Wooldridge's initial conditions with censoring for the lags both for the one-tiered and the two-tiered models.

DOI: 10.1002/jae

Table I. Correlated RE+AR(1) one-tiered and two-tiered dynamic panel Tobit models of married women's labour supply

	One-tiered			Two-tiered Wooldridge	
Initial cond.	Heckman	Fixed	Wooldridge	Particip.	Hours
Const	-91.85	-107.21	-18.91	-221.45	438.47
	(300.33)	(153.22)	(132.28)	(148.55)	(103.24)
Edu	81.84	40.03	15.04	28.57	-3.377
	(12.08)	(5.82)	(4.65)	(4.536)	(3.153)
Race	138.78	100.91	63.51	-11.99	79.55
	(58.98)	(28.55)	(22.450)	(22.30)	(17.37)
Age	48.23	13.14	232.01	22.48	12.60
	(12.63)	(6.69)	(59.46)	(7.01)	(4.832)
Age^2	-0.819	-0.31	-37.74	-0.340	-0.197
	(0.15)	(0.080)	(7.222)	(0.083)	(0.058)
Hinc	-1.595	-1.52	-1.444	-1.210	-0.984
	(0.44)	(0.474)	(0.476)	(0.756)	(0.468)
C12	-193.18	-62.09	-63.33	-64.04	-23.42
	(16.24)	(14.36)	(14.26)	(23.21)	(15.27)
C35	-130.92	-37.33	-30.55	4.245	-20.73
	(18.12)	(14.17)	(13.85)	(22.27)	(12.61)
C613	-39.74	17.76	17.41	45.67	10.57
	(15.23)	(10.68)	(10.47)	(17.31)	(9.990)
$\overline{Hinc}$	-6.851	-2.065	-1.139	-0.438	-1.228
	(0.74)	(0.622)	(0.589)	(0.844)	(0.582)
$\overline{C12}$	-278.61	-223.63	-184.04	-153.52	-202.64
	(178.19)	(78.04)	(60.95)	(61.45)	(46.35)
$\overline{C35}$	-308.85	-235.97	-97.04	-5.171	-15.19
	(173.95)	(78.41)	(59.55)	(57.45)	(43.37)
$\overline{C613}$	-88.54	-37.37	20.02	2.537	-14.45
	(49.63)	(23.95)	(19.95)	(25.01)	(15.57)
λ	-0.094	0.659			
	(0.011)	(0.011)			
λ_1			0.526	0.335	0.536
			(0.012)	(0.018)	(0.012)
λ_2			-501.48	-536.09	18.06
			(16.88)	(21.32)	(19.42)
δ_1			0.162	0.034	0.131
			(0.014)	(0.015)	(0.010)
δ_2			-170.13	-205.60	53.11
			(26.05)	(24.94)	(18.71)
ζ	0.700	-0.021	-0.056	-0.089	
	(0.018)	(0.011)	(0.010)	(0.011)	
σ_u	577.05	589.4	583.51	493.61	
	(2.24)	(2.370)	(2.272)	(1.983)	
σ_d	814.28	376.1	274.96	177.31	
	(25.55)	(12.18)	(9.818)	(6.948)	
θ	0.796				
	(0.024)				
heduc	-13.97				
	(1.61)				
ρ	0.666	0.290	0.181	0.114	
Log-likelihood	−91,684	−91,623	−91,110	−90,272	
Iterations	93	118	168	149+48	

Note:The first column of results corresponds to the replication of Chang's paper for the one-tiered model. It implements the model of Heckman with an approximate reduced form using husband's education as an instrumental variable. The second column gives the estimation with fixed initial conditions. The third column corresponds to the full implementation of Wooldridge (2005). The last two columns correspond to the two-tiered model with Wooldridge's initial conditions. The stationarity condition on ζ is introduced by a parametric transformation. We used BFGS and symmetric numerical derivatives. Standard errors are shown in parentheses and are computed using BHHH. For the last two columns, we used BFGS + BHHH in a second round.

DOI: 10.1002/jae

3.1. Optimization Strategy

The positivity constraint for σ_d and σ_u is directly implemented as σ_d and σ_u appear only as squares in the likelihood function. Only the stationarity assumption $|\zeta| < 1$ appears to be crucial for the success of the optimization process. This constraint can be imposed using either `constrOptim` as in Chang (2011b), or using `BFGS` together with a suitable transformation verifying $|g(.)| < 1$ and $g(\zeta) \simeq \zeta$ over $[-b, b]$ for a given $b < 1$. A good candidate that worked well was $g(z) = 0.9 * \sin(1.164387 \times z)$. A symmetric numerical gradient proved to be more efficient than the default numerical gradient of `optim` (`maxLik` directly implements symmetric numerical derivatives).

3.2. Empirical Results for the One-Tiered Model

We have tried to reproduce in column one of Table I the results of Chang's one-tiered correlated RE+AR(1) model. Using Chang's results as starting values along with 1 for θ, 0 for the instrument parameter and (600, 800) for two error components, we obtained the maximum of the likelihood function using Heckman's initial conditions in 93 iterations with `BFGS` and symmetric numerical derivatives. Overall, our parameter estimates are rather close to those reported by Chang (within one standard deviation), but there are notable exceptions for the impact of the number of children. The standard errors are generally slightly smaller than those reported by Chang. The state dependence parameter is negative. If we now use fixed initial conditions, the algorithm converges in 118 iterations, with slightly lower standard deviations and now a positive and significant true state dependence parameter, a slightly lower value for ζ and a higher likelihood value.

We next implement the Wooldridge specification for the initial conditions, allowing for different parameter values for λ and δ when the lagged and initial dependent variables are censored or not, as recommended in Wooldridge (2005, p. 49). In the third column, the log-likelihood value is now much higher, and the structure of the dynamics is completely altered. There is a much more important positive true state dependence measured by λ_1 for uncensored past working hours, while the value of ζ becomes negative while remaining significant. The intra-class correlation is now much smaller. For a censored lagged endogenous variable, the coefficient is negative and significant. The two fixed effects coefficients have the same sign as the corresponding λ.

3.3. Empirical Results for the Two-Tiered Model

Since we now have two equations, and as we distinguish censored and uncensored events, we have four different values for the λs and the δs when using Wooldridge's specification for the initial condition, given in equations (11)–(13). The four individual effect parameters are significant and asymmetric. We obtained positive and significant effects for the uncensored λ_1 in the two equations. The censored λ_2 is negative and significant for the participation equation and is positive and insignificant in the working hours equation as reported in the last two columns of Table I. Similarly to the results reported in the final estimation of the one-tiered model, ζ has a significant negative effect which refers to unobserved oscillations. The log-likelihood function has a much higher value than that of the corresponding one-tiered model. Clearly, the two-tiered model is an improvement over the one-tiered model not only because the log-likelihood value is higher but also because this model relaxes many constraints allowing the asymmetric effects between the two equations to be captured. With our specification, the number of children has a clear impact on the decision to participate. Mothers with young babies do not participate, with mid-age children the effect is not significant, while with children between 6 and 13 mothers are significantly more likely to participate.

DOI: 10.1002/jae

4. CONCLUSION

What are our main conclusions? First, we could not reproduce Chang's 2011b results in a satisfactory way as, among other things, it was not clear how to empirically implement Heckman's initial conditions with Chang's dataset. Second, the choices made for the specification of the initial conditions are of prime importance for accurately measuring the true state dependence. Heckman's approach should ideally be combined with a latent lagged endogenous variable. Wooldridge's specification is much simpler as it allows naturally for an observed lagged endogenous variable, even if a distinction has to be made when that variable is censored or not. It seems to us to be the most satisfactory approach. The estimate of true state dependence it provides is in accordance with intuition and it reduces the importance of individual random effects, the latter having a much lower variance. Finally, it provides a clear interpretation of the influence of the number of children on the decision of mothers to participate in the labour market.

ACKNOWLEDGEMENTS

We would like to thank Stephen Bazen for useful discussions and Christelle Lecourt for giving us the idea to write this paper. The comments by two anonymous referees and the Editor helped to improve greatly the contents of this paper. The usual disclaimers apply. This work has been carried out thanks to the support of the A*MIDEX project (ANR-11-IDEX-0001-02), funded by the 'Investissements d'Avenir' French Government program, managed by the French National Research Agency (ANR).

REFERENCES

Chang S-K. 2011a. A computationally practical simulation estimation for dynamic panel Tobit models. *Academia Economic Papers* **39**(1): 1–32.

Chang S-K. 2011b. Simulation estimation of two-tiered dynamic panel Tobit models with an application to the labor supply of married women. *Journal of Applied Econometrics* **26**(5): 854–871.

Heckman J. 1981a. The incidental parameter problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In *Structural Analysis of Discrete Panel Data with Econometric Applications*, Manski C, McFadden D (eds). MIT Press: Cambridge, MA; 179–195.

Heckman J. 1981b. Statistical models for discrete panel data. In *Structural Analysis of Discrete Panel Data with Econometric Applications*, Manski C, McFadden D (eds). MIT press: Cambridge, MA; 114–178.

Stewart M. 2006. Maximum simulated likelihood estimation of random- effects dynamic probit models with autocorrelated errors. *Stata Journal* **6**(2): 256–272.

Wooldridge JM. 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics* **20**(1): 39–54.

DOI: 10.1002/jae