Estimation of the Female Labor Supply Models by Heckman's Two-Step Estimator and the Maximum Likelihood Estimator

K. Nawata

Department of Geosystem Engineering, Graduate School of Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan(nawata@geosys.t.u-tokyo.ac.jp)

Abstract: The female labor supply models have been widely used by various authors as early as Heckman (1974). The models are given by: $h_i = x_{1i}'\alpha + \gamma\ w_i + u_i$, and $w_i = x_{2i}'\beta + v_i$,where h_i is the number of hours worked, x_{1i} and x_{2i} are vectors of explanatory variables which describe the woman's characteristics, and w_i is her wage. h_i and w_i are not observed unless she is working. Unlike other types of econometric models, the maximum likelihood estimator (MLE) is seldom used because of its computational difficulty. The models are usually estimated by Heckman's two-step estimator. However, Heckman's estimator sometimes performs poorly. The problems of the estimator are: (i) the estimator cannot be calculated if x_{1i} contains all variables belonging to x_{2i}, and (ii) the estimator is not efficient even if it can be calculated. Therefore, it is reasonable to use the MLE to estimate the model. This paper considers an estimation of the models by the maximum likelihood method. The likelihood function is presented, and a new algorithm which makes calculation of the MLE possible is proposed. The finite sample properties of Heckman's two-step estimator and the MLE are compared using Monte Carlo experiments. Although it is a widely used method, Heckman's two-step estimator performs quite poorly for this model. Meanwhile, the MLE performs well in all cases. The MLE is much better than Heckman's two-step estimator. Hence, the model used in this study should be estimated by the MLE and all empirical studies which use Heckman's two-step estimator should be revised from this viewpoint.

Keywords: Female labor supply; Sample selection bias; Maximum likelihood estimator; Heckman's two-step estimator

1. INTRODUCTION

The female labor supply models have been widely used by various authors as early as Heckman [1974]. For examples and details of the models, see [Jacoby, 1993; Gangadharan and Rosenbloom, 1996; Averett and Hotchkiss, 1997; and Fernández and Rodríguez-Poo, 1997]. Unlike other types of econometric models, the maximum likelihood estimator (MLE) is seldom used because of its computational difficulty. The models are usually estimated by Heckman's two-step estimator. However, Heckman's estimator sometimes performs poorly [Nawata, 1993 and 1994; and Nawata and Nagese, 1996]. This paper considers an estimation of the models by the maximum likelihood method. The likelihood function is presented, and a new algorithm which makes calculation of the MLE possible is proposed. The finite sample properties of Heckman's two-step estimator and the MLE are compared using Monte Carlo experiments

2. MODEL

The model considered in this paper has been widely used by various authors. The i-th woman's reservation wage (value of time) w_i^* is given by

$$w_i^* = x_{1i}'\alpha^* + \gamma^* h_i + u_i^*, \quad (1)$$

where h_i is the number of hours worked and x_{1i} is a vector of explanatory variables which describe the woman's characteristics. Let w_i be her offered wage. (w_i = log(wage) is often used in empirical studies.) w_i is independently determined from h_i and given by

$$w_i = x_{2i}'\beta + v_i, \quad (2)$$

where x_{2i} is another vector of explanatory variables. x_{1i} and x_{2i} may contain different explanatory variables. x_{1i} and x_{2i} are exogenously determined and satisfy the standard assumptions. u_i^* and v_i are independent and follow normal distributions with means 0, and variances $\sigma_{u^*}^2$ and σ_v^2, respectively.

The i-th woman works if

$$w_i^*(h_i = 0) = x_{1i}'\alpha + u_i^* < w_i \quad (3)$$

When she works, she chooses h_i so that $w_i^* = w_i$. Her labor supply equation is given by

$$h_i = x_{1i}'\alpha + \gamma\ w_i + u_i, \quad (4)$$

where $\alpha = -\alpha^*/\gamma^*, \gamma = 1/\gamma^*,\quad u_i = u_i^*/\gamma^*$ and $V(u_i) = \sigma_\varepsilon^2$.

Let y_i be a dummy variable such that y_i=1 if the i-th woman works and 0 otherwise. Substituting (1) and (2) into (3), we get the labor participation equation given by

$$y_i = 1(y_i^* > 0), \quad (5)$$

$$y_i^* = x_{1i}'\alpha + \gamma(x_{2i}'\beta) + u_i + \gamma v_i$$
$$= x_{1i}'\alpha + \gamma(x_{2i}'\beta) + \varepsilon_i,$$

where 1(•) is the indicator function such that 1(•) = 1 if • is true and 0 otherwise and $\varepsilon_i = u_i + \gamma\ v_i$. Let x_i be the vector of all explanatory variables contained in x_{1i} and x_{2i}. Since the multiplication of a positive constant does not change the sign of y_i^*, the second equation of (5) can be rewritten as

$$y_i^* = x_i'\delta + \varepsilon_i^*, \quad (6)$$

where $\varepsilon_i^* = \varepsilon_i/\sigma_\varepsilon$, $\sigma_\varepsilon^2 = V(\varepsilon_i) = \sigma_u + \gamma^2\sigma_v^2$, and $V(\varepsilon_i^*) = 1$.

3. ESTIMATION OF THE MODEL

3.1 Heckman's Two-Step Estimator

The model given in Section 2 can be estimated by the Heckman's two-step estimator. Let Φ and ϕ be the distribution and density functions of the standard normal distribution. $E(w_i \mid y_i = 1)$ and $E(h_i \mid y_i = 1)$ are given by:

$$E(w_i \mid y_i = 1) = x_{2i}'\beta + \sigma_2\lambda(x_i'\delta), \quad (7)$$
$$E(h_i \mid y_i = 1) = x_{1i}'\alpha + \gamma\ E(w_i \mid y_i = 1) + \sigma_1\lambda(x_i'\delta),$$

where $\sigma_1 = Cov(u_i, \varepsilon_i^*)$, $\sigma_2 = Cov(v_i, \varepsilon_i^*)$, and $\lambda(z) = \phi(z)/\Phi(z)$.

Heckman [1976 and 1978] proposed the two-step estimator obtained by the following steps.

- Estimate δ with the probit MLE. Let $\hat{\delta}$ be the probit MLE.
- Replacing δ by $\hat{\delta}$, estimate the first equation of (7) by the least squares method using only the $y_i = 1$ observations. Let $\hat{w}_i$ be the least squares predictor of w_i.
- Replace δ and w_i with $\hat{\delta}$ and $\hat{w}_i$, and estimate the second equation of (7) by the least squares method using only the $y_i = 1$ observations.

3.2 Maximum Likelihood Estimator

Although Heckman's two-step estimator is widely used, the problems of the estimator are:

- the estimator cannot be calculated if x_{1i} contains all variables belonging to x_{2i}, and
- the estimator is not efficient even if it can be calculated.

Therefore, it is reasonable to use the MLE to estimate the model. Let $\vartheta' = (\alpha', \beta', \gamma, \sigma_u, \sigma_v)$. Since $\varepsilon_i = u_i + \gamma\ v_i$, we can get the likelihood function given by

$$L(\vartheta) =$$

$$\prod_{y_i=1}\left[\frac{1}{\sigma_u}\phi\{\frac{h - (x_{1i}'\alpha + \gamma\ w_i)}{\sigma_u}\}\frac{1}{\sigma_v}\phi(\frac{w_i - x_{2i}'\beta}{\sigma_v})\right]$$

$$\times\prod_{y_i=0}\left[1 - \Phi\{\frac{x_{1i}'\alpha + \gamma(x_{2i}'\beta)}{\sqrt{\sigma_u^2 + \gamma^2\sigma_v^2}}\}\right], \quad (8)$$

from the modifications of the standard type III tobit model [for details, see Amemiya 1985, p.390]. It is easy to show that $\hat{\vartheta}'=(\hat{\alpha}',\hat{\beta}',\hat{\gamma},\hat{\sigma}_v)$, which maximizes (8), is consistent and asymptotically normal by the standard arguments of the MLE.

3.3 Algorithm

Since the likelihood function given by (8) is a complicated (nonlinear) function of ϑ, the calculation of $\hat{\vartheta}$ may not be easy to perform. Although the likelihood function is a concave function of $\eta'=(\alpha',\beta')$ when the values of γ, σ_u and σ_v are given, it is not a concave function of γ, σ_u and σ_v. Therefore, the standard algorithms may not converge to the maximum value. The following method, which is a modification of the scanning procedure of Nawata [1994 and 1995], is used to calculate the MLE.

- Normalize w_i, dividing by its sample standard deviation. Select a proper value of D and choose equidistant M_1 points from $[-D,D]$. Let δ be the distance between any two points.
- Let $\gamma=0$ and calculate $\hat{\alpha}_0,\hat{\beta}_0$, $\hat{\sigma}_{u_0}$, and $\hat{\sigma}_{v_0}$, which maximize the conditional maximum likelihood function. Note that $\hat{\alpha}_0$ and $\hat{\sigma}_{u_0}$ are the tobit MLE of $h_i=y_i(x_{1i}'\alpha+u_i)$ and that $\hat{\beta}_0$ and $\hat{\sigma}_{v_0}$ are the OLS estimators of (2) using $y_i=1$ observations.
- Let $\hat{\alpha}_j,\hat{\beta}_j$, $\hat{\sigma}_{u_j}$, and $\hat{\sigma}_{v_j}$ be the j-th estimators. Increase γ by δ, choose the initial values of the iteration as $\hat{\alpha}_j,\hat{\beta}_j$, $\hat{\sigma}_{u_j}$, and $\hat{\sigma}_{v_j}$, and calculate the (j+1)-th estimator. Since the likelihood is a continuous function of γ, the previous estimators are in the neighborhood of the maximum value.
- Continue the previous step and calculate estimators up to D, the largest value of γ, determined in the first step.
- In the same way, calculate estimators from 0 to $-D$, the smallest value of γ.
- Choose the value of γ which maximizes the conditional likelihood function. If the conditional maximum likelihood function is maximized at D or $-D$, increase the value of D and repeat the steps.
- Choose M_2 points in the neighborhood of the value determined in the previous step and repeat the procedure.
- Determine the final estimators, $\hat{\alpha},\hat{\beta},\hat{\sigma}_u,\hat{\sigma}_v$ and $\hat{\gamma}$. Note that since the values determined in the previous step are sufficiently close to the maximum value, $\hat{\alpha},\hat{\beta},\hat{\sigma}_u,\hat{\sigma}_v$ and $\hat{\gamma}$ can be calculated by using the Newton-Raphson method employing the values determined in the previous step as the initial values of the iteration.

4. MONTE CARLO EXPERIMENTS

In this section some Monte Carlo results are presented for Heckman's two-step estimator and the MLE. The basic model considered in this paper is:

$$h_i=\alpha_0+\alpha_1x_{1i}+\gamma\, w_i+u_i, \tag{9}$$
$$w_i=\beta_0+\beta_2x_{2i}+v_i,$$
$$y_i=1[\alpha_0+\alpha_1x_{1i}+\gamma(\beta_0+\beta_1x_{2i})+u_i+\gamma v_i>0],$$
$$i=1,2,\ldots,N.$$

If $y_i=1$, w_i is observable and $h_i>0$. u_i and v_i are normal random variables with mean 0 and variance 1. The effect of the correlation of x_{1i} and x_{2i} is considered, and the values of x_{1i} and x_{2i} are determined as follows:

$$x_{1i}=\xi_{1i}, \tag{10}$$
$$x_{2i}=[\pi\xi_{1i}+(1-\pi)\xi_{2i}]/\sqrt{\pi^2+(1-\pi)^2}.$$

ξ_{1i} and ξ_{2i} are normal random variables with mean 0 and variance 4. ξ_{1i} and ξ_{2i} are independently distributed. $\pi/\sqrt{\pi^2+(1-\pi)^2}$ is the correlation coefficient of x_{1i} and x_{2i}, and $\pi=0$, 0.5, 0.8 and 1.0 are considered. (Since Heckman's two-step estimator cannot be calculated, only the MLE is considered for the $\pi=1.0$ cases.)

The true parameter values are:

$$\alpha_0=0.0,\quad \alpha_1=1.0,\quad \gamma=0.5, \tag{11}$$
$$\beta_0=0 \text{ and } \beta_1=1.0.$$

The sample sizes of N = 100, 200 and 400 are considered, and the number of trials is 500 for all cases. The MLE is calculated by the method described in Section 3 so that:

Table 1. Heckman's two-step estimator and the MLE ($\pi = 0.0$).

	Mean	S.D.	25%	Median	75%		Mean	S.D.	25%	Median	75%
Heckman						MLE					
N=100						N=100					
α_0	-1.198	10.200	-3.358	0.390	3.208	α_0	-0.002	0.304	-0.166	0.035	0.186
α_1	1.116	1.980	-0.073	0.903	2.178	α_1	1.014	0.247	0.835	1.026	1.201
γ	0.445	4.659	-0.024	0.383	1.101	γ	0.503	0.156	0.423	0.507	0.598
β_0	-0.089	1.221	-0.707	0.026	0.583	β_0	0.011	0.364	-0.244	0.009	0.253
β_1	0.999	0.371	0.779	0.972	1.230	β_1	1.003	0.285	0.789	1.001	1.223
N=200						N=200					
α_0	-0.164	4.645	-2.516	0.201	2.628	α_0	0.012	0.213	-0.110	-0.007	0.178
α_1	1.022	1.293	0.115	0.937	1.847	α_1	1.010	0.188	0.888	1.011	1.146
γ	0.494	0.751	0.104	0.425	0.841	γ	0.503	0.104	0.447	0.509	0.573
β_0	-0.042	0.725	-0.476	0.003	0.387	β_0	0.011	0.265	-0.125	0.014	0.159
β_1	1.008	0.235	0.866	1.004	1.146	β_1	0.989	0.210	0.849	0.981	1.112
N=400						N=400					
α_0	0.014	2.762	-1.446	0.162	1.776	α_0	0.011	0.150	-0.090	0.019	0.126
α_1	0.974	0.872	0.379	0.972	1.521	α_1	1.014	0.126	0.940	1.003	1.087
γ	0.482	0.453	0.193	0.442	0.730	γ	0.503	0.071	0.454	0.499	0.546
β_0	-0.018	0.439	-0.289	-0.009	0.271	β_0	-0.005	0.183	-0.115	-0.005	0.120
β_1	1.007	0.161	0.909	1.000	1.112	β_1	0.992	0.142	0.893	0.981	1.100

Table 2. Heckman's two-step estimator and the MLE ($\pi = 0.5$).

	Mean	S.D.	25%	Median	75%		Mean	S.D.	25%	Median	75%
Heckman						MLE					
N=100						N=100					
α_0	4.610	121.23	−2.366	0.361	2.445	α_0	0.019	0.302	−0.180	0.044	0.235
α_1	1.033	1.403	0.185	0.881	1.660	α_1	1.008	0.288	0.808	1.005	1.192
γ	−0.852	38.618	0.070	0.411	0.877	γ	0.494	0.164	0.392	0.504	0.605
β_0	0.038	1.742	−0.795	0.105	0.990	β_0	0.004	0.376	−0.254	−0.004	0.274
β_1	0.958	0.693	0.580	0.963	1.317	β_1	0.990	0.307	0.804	0.980	1.194
N=200						N=200					
α_0	−0.528	11.244	−1.812	0.134	1.869	α_0	0.017	0.213	−0.142	0.009	0.161
α_1	1.021	0.919	0.379	0.921	1.605	α_1	1.025	0.196	0.885	1.033	1.148
γ	0.658	4.004	0.157	0.426	0.750	γ	0.488	0.108	0.412	0.487	0.569
β_0	−0.044	1.044	−0.634	−0.005	0.608	β_0	0.014	0.255	−0.145	0.048	0.199
β_1	1.008	0.418	0.752	0.986	1.249	β_1	1.001	0.220	0.855	0.997	1.146
N=400						N=400					
α_0	0.055	1.982	−1.033	0.176	1.361	α_0	0.005	0.146	−0.103	0.014	0.103
α_1	0.973	0.619	0.538	0.929	1.350	α_1	1.013	0.141	0.918	1.010	1.118
γ	0.481	0.396	0.214	0.451	0.658	γ	0.490	0.079	0.438	0.485	0.544
β_0	−0.057	0.6636	−0.427	−0.045	0.356	β_0	0.011	0.199	−0.115	0.010	0.141
β_1	1.0247	0.2953	0.832	1.0267	1.2018	β_1	1.008	0.167	0.906	1.017	1.116

Table 3. Heckman's two-step estimator and the MLE ($\pi = 0.8$).

	Mean	S.D.	25%	Median	75%		Mean	S.D.	25%	Median	75%
Heckman						MLE					
N=100						N=100					
α_0	−1.522	38.021	−2.139	0.656	3.012	α_0	0.052	0.339	−0.173	0.068	0.274
α_1	0.982	2.223	−0.114	0.806	1.924	α_1	1.021	0.341	0.810	1.014	1.239
γ	1.483	23.117	−0.677	0.350	1.322	γ	0.475	0.183	0.374	0.500	0.597
β_0	0.326	3.328	−1.046	0.448	2.014	β_0	0.011	0.433	−0.246	0.033	0.313
β_1	0.828	1.386	0.133	0.770	1.433	β_1	0.989	0.350	0.745	0.971	1.200
N=200						N=200					
α_0	0.6535	18.219	−1.405	0.4736	2.1302	α_0	0.007	0.241	−0.154	0.007	0.187
α_1	1.0361	1.3675	0.1577	0.833	1.676	α_1	1.001	0.246	0.837	1.008	1.151
γ	0.2077	8.246	−0.326	0.2842	0.9305	γ	0.499	0.127	0.432	0.497	0.579
β_0	0.053	2.009	−1.107	0.235	1.219	β_0	−0.003	0.295	−0.193	−0.008	0.200
β_1	0.964	0.885	0.446	0.870	1.462	β_1	0.998	0.248	0.814	0.982	1.176
N=400						N=400					
α_0	0.202	12.492	−1.166	0.339	1.518	α_0	0.013	0.173	−0.120	0.016	0.137
α_1	0.947	0.791	0.396	0.850	1.433	α_1	1.010	0.172	0.893	1.018	1.115
γ	0.487	5.365	−0.066	0.437	0.891	γ	0.498	0.084	0.439	0.495	0.561
β_0	0.108	1.445	−0.712	0.156	1.027	β_0	0.003	0.228	−0.142	0.015	0.148
β_1	0.946	0.652	0.495	0.919	1.332	β_1	0.988	0.187	0.857	0.974	1.118

Table 4 . Heckman's two-step estimator and the MLE ($\pi = 1.0$).

	Mean	S.D.	25%	Median	75%		Mean	S.D.	25%	Median	75%
Heckman						MLE					
N=100						N=100					
α_0	–	–	–	–	–	α_0	0.033	0.309	−0.180	0.042	0.244
α_1	–	–	–	–	–	α_1	0.995	0.348	0.775	1.001	1.230
γ	–	–	–	–	–	γ	0.488	0.185	0.387	0.500	0.613
β_0	–	–	–	–	–	β_0	0.010	0.463	−0.279	0.017	0.334
β_1	–	–	–	–	–	β_1	0.994	0.375	0.769	0.990	1.237
N=200						N=200					
α_0	–	–	–	–	–	α_0	0.020	0.229	−0.160	0.025	0.170
α_1	–	–	–	–	–	α_1	0.985	0.226	0.824	0.995	1.152
γ	–	–	–	–	–	γ	0.500	0.119	0.426	0.505	0.572
β_0	–	–	–	–	–	β_0	−0.003	0.305	−0.228	0.014	0.182
β_1	–	–	–	–	–	β_1	0.994	0.260	0.826	0.973	1.142
N=400						N=400					
α_0	–	–	–	–	–	α_0	0.013	0.159	−0.100	0.005	0.122
α_1	–	–	–	–	–	α_1	1.006	0.171	0.881	1.005	1.128
γ	–	–	–	–	–	γ	0.500	0.088	0.439	0.507	0.561
β_0	–	–	–	–	–	β_0	−0.002	0.228	−0.162	0.003	0.165
β_1	–	–	–	–	–	β_1	0.995	0.167	0.883	1.000	1.104

- D =3 and $\delta = 0.1$ are chosen in the first step.
- Steps are repeated three times. $\delta = 0.01$ and $\delta = 0.001$ are used in the second and third repetitions.

The results of estimates of α's, β's and γ are presented in Tables 1-4. Note that the following notations are used in the tables.

S.D.: Standard Deviation, 25%: 25% Percentile, and 75%: 75% Percentile.

The MLE performs well. The biases are quite small and the standard errors are reasonably small in all cases. However, Heckman's two-step estimator performs quite poorly especially when N is small and π is large. The standard errors are much larger than those of MLE, and the biases are significantly large in some cases.

5. CONCLUSION

This paper considers estimations of the female labor supply model by Heckman's two-step estimator and the MLE. A new algorithm, which makes calculation of the MLE possible, is considered. The finite sample performance of the two estimators is examined by Monte Carlo experiments. Although it is a widely used method, Heckman's two-step estimator performs quite poorly for this model. Meanwhile, the MLE performs well in all cases. The MLE is much better than Heckman's two-step estimator. Hence, the model used in this study should be estimated by the MLE and all empirical studies which use Heckman's two-step estimator should be revised from this viewpoint.

REFERENCES

Amemiya, T., *Advanced Econometrics*, Harvard University Press, Cambridge, Massachusetts, 1985.

Averett, S. L., and J. L. Hotchkiss, Female Labor Supply with a Discontinuous, Non-Convex Budget Constraint: Incorporation of a Part-time/Full-time Wage Differential, *Review of Economics and Statistics*, 79, 461-470, 1997.

Blundell, R., and R. J. Smith, Coherency and Estimation in Simultaneous Models with Censored or Qualitative Dependent Variables, *Journal of Econometrics*, 64, 355-373, 1994.

Fernández, A. I. and J. M. Rodríguez-Poo, Estimation and Specification Testing in Female Labor Participation Models: Parametric and Semiparametric Methods, *Econometric Reviews*, 16, 229-247, 1997.

Gangadhan, J., and J. L. Rosenbloom, The Effects of Child-Bearing on Married Woman's Labor Supply and Earnings: Using Twin Births as a Natural Experiment, Working Paper 5647, Bureau of Economic Research, Inc., 1996.

Heckman, J., Shadow Prices, Market Wages, and Labor Supply, *Econometrica*, 42, 679-693, 1974.

Heckman, J., The Common Structure of Statistical Models for Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *Annals of Economic and Social Measurement* 5, 475-492, 1976.

Heckman J., Dummy Endogenous Variables in a Simultaneous Equation System, *Econometrica*, 46, 931-960, 1978.

Heckman J., Sample Selection Bias as Specific Error, *Econometrica*, 47, 152-161, 1979.

Jacoby, H. G., Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra, *Review of Economic Studies*, 60, 903-921, 1993.

Lemieux, T., P. Fortin, and P. Frechette, The Effect of Taxes on Labor Supply in the Underground Economy, *American Economic Reviews*, 84, 231-254, 1994.

Nawata, K., A Note on Estimation of Models with Sample-Selection Biases, *Economic Letters*, 42, 15-24, 1993.

Nawata K., Estimation of the Sample-Selection Models by the Maximum Likelihood Estimator and Heckman's Two-Step Estimator, *Economic Letters*, 45, 33-40, 1994.

Nawata K., Estimation of Sample-Selection Models by the Maximum Likelihood Method, *Mathematics and Computers in Simulation*, 39, 299-303, 1995.

Nawata, K., and N. Nagase, Estimation of Sample Selection Models, *Econometric Reviews*, 15, 387-400, 1996.