Discussion Paper

Central Bureau of Statistics, P.B. 8131 Dep, 0033 Oslo 1, Norway

No. 28 16 December 1987

A MODIFICATION OF HECKMAN'S TWO STAGE ESTIMATION PROCEDURE THAT IS APPLICABLE WHEN THE BUDGET SET IS CONVEX

BY

JOHN K. DAGSVIK

ABSTRACT

The paper develops a simple estimation procedure for a labor supply model with non-linear convex budget set. The procedure is an extension of Heckman's two stage method. The asymptotic properties of the estimators are derived.

1. Introduction

Heckman (1979) has demonstrated that it is possible to estimate a two-equation model of the Tobit-type by a two stage procedure where only the first stage involves a non-linear problem, namely the estimation of the parameters in a Probit model. Heckman's demonstration was motivated by the econometrics related to the standard labor supply model. This model consists of two equations: One equation specifies the (log) market wage as a function of individual characteristics (qualification variables) while the other equation specifies the (log) shadow price of leisure (marginal rate of substitution) as a function of hours worked and variables that affect the preferences. Since the individual's decision problem involves a corner solution (work or not work) the transformed model that gives hours of work as a function of wages and individual characteristics becomes non-linear in the parameters even if the log wage and log shadow price equations are linear. The Heckman procedure relies on the assumption that the reduced form of the market wage and shadow price equations are linear in parameters. This assumption is restrictive for the following reason: Even if the shadow price is structurally linear, e.g., linear in parameters as a function of wage and consumption the reduced form may often be non-linear because the budget constraint is non-linear. One example of non-linear budget constraints is the case with progressive taxes. Even if the budget constraint is linear the reduced form shadow price may become non-linear because labor income is the product of hours and wage.

The procedure presented here modifies Heckman's original method so as to deal with general concave and smooth budget constraints.[1] This modified method avoids the use of instrument variables to accommodate for endogenous consumption.

The method is applied to estimate labor supply functions for West-Germany and France, see Dagsvik et al. (1987a), (1987b). The empirical evidence from these studies demonstrate that the estimators for the structural parameters of the labor supply equation perform very well.

---

[1] Rudiments of the central idea were presented in Dagsvik (1987) but the estimators proposed there are inconsistent.

The paper is organized as follows: In section 2 the individual decision model is presented and in section 3 the corresponding econometric model is developed. Section 4 discusses the estimation procedure and section 5 contains proofs of the large sample properties of the estimators. The final section discusses the efficiency of estimating the wage equation with the modified method instead of using Heckman's orginal method.

2. The individual decision model

We assume that the individual has utility in leisure, L, and consumption, C, that is of the following type

$$(2.1)\quad U(L,C) = \left(\frac{L^{\alpha}-1}{\alpha}\right)A + \left(\frac{C^{\beta}-1}{\beta}\right)B$$

where $\alpha\leq 1$ and $\beta\leq 1$ are parameters and A and B are tasteshifters that depend on individual characteristics. This is a CES type separable utility function that allows for quite flexible patterns of Cournot and Slutsky elasticities. The budget constraints is given by

$$(2.2)\quad C = f(hW+I),\ L+h = M$$

where $f(\cdot)$ is the function that transforms gross income to consumption (composite consumption). This function also depends on the price of C and of actual deductions and tax-free transfers. I is nonlabor income, h is hours of work, M is total number of hours per year and W is the market wage the individual faces.

Provided $f(\cdot)$ is concave and differentiable the first order conditions yield the following decision rule:

Work if and only if

$$(2.3)\quad f'(I)W > \frac{A}{B}\cdot M^{\alpha-1}f(I)^{1-\beta}.$$

If (2.3) holds hours of work, $\bar{h}$, is determined by

$$(2.4)\quad f'(I+\bar{h}W)W = \frac{A}{B}\,(M-\bar{h})^{\alpha-1}f(I+\bar{h}W)^{1-\beta}.$$

We recognize the left hand side of (2.4) as the marginal wage evaluated at $\bar{h}$. The right hand side is the marginal rate of substitution (or shadow price) evaluated at $\bar{h}$. Similarly, (2.3) expresses that working is optimal when the marginal wage at zero hours is larger than the shadow price at zero hours.

3. The econometric model

In order to estimate the model of section 2 it is necessary to specify a wage equation and A/B. We adopt here the standard specifications

$$(3.1) \qquad \log W = X_1\theta + \varepsilon_1$$

and

$$(3.2) \qquad \log(\frac{A}{B}) + (\alpha-1)\log M = X_2\gamma + \varepsilon_2$$

where ε_1 and ε_2 are jointly normally distributed random variables with zero mean and variances σ_1^2 and σ_2^2, respectively. X_1 is a vector that consists of one, length of schooling and experience. X_2 consists of variables such as: one, age, number and age of children. θ and γ are parameter vectors to be estimated. The error terms ε_1 and ε_2 are supposed to account for variables that are known to the individual but unobservable to the econometrician. The errors ε_1 and ε_2 may be correlated.

Now taking the logarithm on both sides of (2.3) and (2.4), and inserting in (3.1) and (3.2) yield

$$(3.3) \qquad \varepsilon_2 - \varepsilon_1 = \sigma Z(\bar{h}) + (1-\alpha)\log(1-\frac{\bar{h}}{M})$$

when

$$(3.4) \qquad \varepsilon_2 - \varepsilon_1 < \sigma Z(0)$$

where

$$Z(h) = \frac{(\beta-1)}{\sigma}\log f(hW+I) + \frac{X_1\theta - X_2\gamma}{\sigma} + \frac{1}{\sigma}\log f'(hW+I)$$

and $\sigma^2 = \mathrm{Var}(\varepsilon_1 - \varepsilon_2)$. From (3.4) it follows that the probability of working , p, is given by

$$(3.5) \quad p = p(X) = \Phi(Z(0))$$

where $X = (X_1, X_2, I)$, and $\Phi(\cdot)$ is the standard cumulative normal distribution.

For later reference it will be convenient to state a number of properties of the model (3.3) and (3.4). Let

$$(3.6) \quad V = Z(\bar{h}) - \frac{(\alpha-1)}{\sigma} \log (1- \frac{\bar{h}}{M}),$$

$$(3.7) \quad a = E(Z(\bar{h}) \mid \bar{h}>0) + E\lambda$$

and

$$(3.8) \quad b = E(\log (1- \frac{\bar{h}}{M}) \mid \bar{h}>0)$$

where

$$(3.9) \quad \lambda = \lambda(X) = \frac{\Phi'(Z(0))}{\Phi(Z(0))}$$

is the inverse of Mill's ratio. Then from (3.3) and (3.4) we get by straight forward calculus that

$$(3.10) \quad E\{V \mid \bar{h}>0, X\} = -\lambda$$

and

$$(3.11) \quad \mathrm{Var}\{V \mid \bar{h}>0, X\} = 1 - Z(0)\lambda - \lambda^2.$$

Since ε_1 and ε_2 are jointly normal a standard result gives

$$(3.12) \quad E\{\varepsilon_1 \mid \varepsilon_1 - \varepsilon_2\} = \frac{\sigma_1 \rho}{\sigma} (\varepsilon_1 - \varepsilon_2)$$

where $\rho = \mathrm{corr}(\varepsilon_1, \varepsilon_1 - \varepsilon_2)$. Also we have

$$(3.13) \quad \mathrm{Var}\{\varepsilon_1 \mid \varepsilon_1 - \varepsilon_2\} = \sigma_1^2 (1-\rho^2).$$

Thus by (3.3) and (3.4)

(3.14) $E\{\varepsilon_1 | V, \bar{h}>0\} = -\sigma_1 \rho V$

and

(3.15) $\mathrm{Var}\ \{\varepsilon_1 | V,\ \bar{h}>0,\ X\} = \sigma_1^2 (1-\rho^2).$

From (3.10) and (3.6) we immediately get

(3.16) $\alpha - 1 = \frac{\sigma a}{b}$.

This equation is very interesting because it suggests a way of estimating α from conditional means as we shall discuss in the next section.
Eq. (3.14) implies that

(3.17) $E(\log W | h>0, V, X) = X_1\theta - \sigma_1 \rho V$

which also suggets an alternative estimation procedure of the (conditional) wage equation as we shall see below.

4. A four stage estimation procedure

The procedure suggested by Heckman to estimate the labor supply equations consists in estimating the participation probability and next estimate the parameters in the conditional labor supply and wage equations given participation. Unfortunately , this method is not directly applicable here because our labor supply equation (3.3) is a form of a pseudo-supply function where $\bar{h}$ is determined implicitly.

Here we shall discuss a four stage estimation procedure that avoids the use of instrument variables to account for endogenous consumption and marginal wage. Let the index i refer to individual i and let Ω denote the subsample of those who work. The size of Ω is n_1.

The first stage consists of estimating a reduced form probability of working and in the second stage a wage equation is estimated by using Heckman's method to correct for selectivity bias. Often this selectivity bias is negligible so that estimates of the wage equation can be obtained directly without using an estimate of the participation probability. In the third stage a structural model for the participation probability is estimated by entering the systematic term of the wage equation as an instrument variable for the wage rate. From stage two and three we get estimates $\hat{\beta}$, $\hat{\sigma}$ and $\hat{\delta}$ for β, σ and δ, respectively, where δ is defined by

$$X_3\delta = \frac{X_1\theta}{\sigma} - \frac{X_2\gamma}{\sigma} .$$

Hence we are able to compute estimates, $\hat{Z}_i$, for the variable $Z_i(h_i)$ by

$$(4.1) \quad \hat{Z}_i = \frac{(\hat{\beta}-1)}{\hat{\sigma}} \log f\,(h_iW_i + I_i) + X_{i3}\hat{\delta} + \frac{1}{\hat{\sigma}} \log f'(h_iW_i + I) .$$

Similarly we are able to compute estimates, $\hat{\lambda}_i$, for λ_i by

$$(4.2) \quad \hat{\lambda}_i = \lambda(\hat{Z}_i(0)).$$

Now $\hat{a}$ defined by

$$(4.3) \quad \hat{a} = \frac{1}{n_1} \sum_{i\varepsilon\Omega} (\lambda_i + Z_i)$$

is obviously a consistent estimate for a. Also

$$(4.4) \quad \hat{b} = \frac{1}{n_1} \sum_{i\varepsilon\Omega} \log\,(1 - \frac{h_i}{M})$$

is a consistent estimate for b and therefore by (3.16)

$$(4.5)\qquad \hat{\alpha}-1 = \frac{\hat{\sigma}\,\hat{a}}{\hat{b}}$$

is a consistent estimate for α-1. Accordingly, the fourth stage consists of estimating α by (4.5).

It is also possible to improve the estimation of the wage equation. Define e_i by

$$(4.6)\qquad \log W_i = X_{1i}\Theta - \sigma_1 \rho V_i + e_i.$$

Then it follows from (3.14) and (3.15) that

$$E\{e_i | X_i,\ V_i,\ h_i>0\} = 0$$

and

$$\mathrm{Var}\{e_i | X_i,\ V_i,\ h_i>0\} = \sigma_1^2 (1-\rho^2)$$

which demonstrates that Θ and $\sigma_1 \rho$ can be estimated by OLS provided V_i is known. Now, by using the results from stage one and two we are able to compute an estimate

$$(4.7)\qquad \hat{V}_i = \hat{Z}_i - \frac{\hat{a}}{\hat{b}} \log \left(1 - \frac{h_i}{M}\right)$$

for V_i.

The estimation of the wage equation proposed here differs from Heckman's (1979) method in that he introduces λ_i as an additional regressor while we use V_i. By (3.10) λ_i is the conditional mean of $-V_i$ given X_i and $h_i>0$. One would therefore expect Heckman's method to be less efficient than the method proposed here. Also Heckman's approach implies heteroscedastic disturbances which is avoided in the present procedure.

5. The large sample properties of the estimators

Considers first the asymptotic properties of $\hat{a}/\hat{b}$. Let

$$T_1 = \sqrt{n_1}(\hat{a} - a)\ ,\quad T_2 = \sqrt{n_1}(\hat{b} - b).$$

Inserting T_j into (3.16) yields

$$(5.1)\quad \frac{\hat{a}}{\hat{b}} \cdot \frac{b}{a} = \frac{1 + T_1/a\sqrt{n_1}}{1 + T_2/b\sqrt{n_1}} = 1 + \frac{T_1 - \frac{a}{b}T_2}{a\sqrt{n_1}} + o(\sqrt{n_1}).$$

Eq. (5.1) implies that

$$\sqrt{n_1}\left(\frac{\hat{a}}{\hat{b}} - \frac{a}{b}\right)$$

has the same asymptotic distribution as

$$\frac{T_1 - \frac{a}{b}T_2}{b}.$$

Under quite general conditions (see for instance Judge at al, 1985)

$$(5.2)\quad T - \frac{a}{b}T_2 \rightarrow N(0,\tau_1)$$

where

$$(5.3)\quad \tau_1^{\,2} = \lim_{n_1 \to \infty} \mathrm{Var}[n_1^{-\frac{1}{2}} \sum_{i\varepsilon\Omega} (\lambda_i + V_i)].$$

This variance formulae can be simplified due to (3.11). We have

$$\mathrm{Var}\,[n_1^{-\frac{1}{2}} \sum_{i\varepsilon\Omega} (\lambda_i + V_i)]$$

$$= E[\frac{1}{n_1} \sum_{i\varepsilon\Omega} \mathrm{Var}(\lambda_i + V_i) | \, h_i > 0, X_i]$$

$$+ \mathrm{Var}\,[n_i^{-\frac{1}{2}} \sum_{i\varepsilon\Omega} E(\lambda_i + V_i) \, | h_i > 0, X_i]$$

which by (3.10) and (3.11) reduces to

$$E\,\mathrm{Var}[n_1^{-\frac{1}{2}}(\sum_{i\varepsilon\Omega} V_i)|h_i>0,X_i]$$

$$= E\,[\frac{1}{n_1}\sum_{i\varepsilon\Omega}\mathrm{Var}(V_i|h_i>0,X_i)]$$

$$= \{E\,\frac{1}{n_1}\sum_{i\varepsilon\Omega}(1-Z_i(0)\lambda_i-\lambda_i^2)\}.$$

Thus

$$(5.4)\quad \tau_1^2 = \lim_{n_1\to\infty} E\,\{\frac{1}{n_1}\sum_{i\varepsilon\Omega}(1-Z_i(0)\lambda_i-\lambda_i^2)\}.$$

However, since we do not know Z_i we must apply the estimate $\hat{Z}_i$ and consequently τ_1 is not the correct asymptotic standard error of $\hat{a}/\hat{b}$. The correct variance that takes into account the sampling error in the coefficients obtained in stage one is defined by

$$(5.5)\quad \tau_2^2 = \lim_{n_1\to\infty}\mathrm{Var}\,\{n_1^{-\frac{1}{2}}\sum_{i\varepsilon\Omega}(\hat{\lambda}_i+\hat{Z}_i-\frac{a}{b}\log(1-\frac{h_i}{M}))\}.$$

In order to simplify the arguments below we introduce the variables

$$Y_i=\begin{cases}1 & \text{when } h_i>0\\ 0 & \text{otherwise}\end{cases}$$

and $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_n)$, $n_1=\sum_i Y_i$.

Thus we can express (5.5) as

$$\tau_2^2 = \lim_{n\to\infty}\mathrm{Var}\,\{n_1^{-\frac{1}{2}}\sum_i(\hat{\lambda}_i+\hat{Z}_i-\frac{a}{b}\log(1-\frac{h_i}{M}))Y_i\}$$

where n is the sample size.
Now observe that conditional on $\underset{\sim}{Y}$ the parameter estimates obtained in stage three become non-stochastic.

As a consequence $\hat{\lambda}_i+\hat{Z}_i$ and $\hat{\lambda}_j+\hat{Z}_j$ become independent for $i\neq j$ whereas they depend on each other when $\underset{\sim}{Y}$ is not given. Thus

$$(5.6)\qquad \tau_2^2 = \lim_{n\to\infty} [E\{\frac{1}{n_1}\sum_i \mathrm{Var}((\hat{\lambda}_i + \hat{Z}_i - \frac{a}{b}\log(1-\frac{h_i}{M}))\, Y_i \mid \underset{\sim}{Y})\}$$

$$+ \mathrm{Var}\{n_1^{-\frac{1}{2}}\sum_i E(\hat{\lambda}_i + \hat{Z}_i - \frac{a}{b}\log(1-\frac{h_i}{M}))\, Y_i \mid \underset{\sim}{Y})\}].$$

Let

$$\Delta_i = \hat{\lambda}_i + \hat{Z}_i - \lambda_i - Z_i$$

and notice that Δ_i is of order $O(n^{-\frac{1}{2}})$ since the variance of the stage one parameter estimates converges to zero with rate n. Hence

$$\mathrm{Var}[(\hat{\lambda}_i + \hat{Z}_i - \frac{a}{b}\log(1-\frac{h_i}{M}))\, Y_i|\underset{\sim}{Y})) = \mathrm{Var}((V_i + \lambda_i + \Delta_i)Y_i|\underset{\sim}{Y}]$$

$$= Y_i\mathrm{Var}(V_i + \lambda_i|\underset{\sim}{Y}) + Y_i\mathrm{Var}(\Delta_i|Y_i) + 2Y_i\mathrm{cov}(V_i + \lambda_i, \Delta_i|\underset{\sim}{Y})$$

$$= Y_i\mathrm{Var}(V_i + \lambda_i|\underset{\sim}{Y}) + O(n^{-1}).$$

We have thus demonstrated that

$$(5.7)\qquad E\{\frac{1}{n_1}\sum_i \mathrm{Var}((\hat{\lambda}_i + \hat{Z}_i - \frac{a}{b}\log(1-\frac{h_i}{M}))\, Y_i|\underset{\sim}{Y})\} = \tau_1^2 + O(n^{-1}).$$

Consider next the second term of (5.6). We have by (3.10)

$$E\{(\hat{\lambda}_i + \hat{Z}_i - \frac{a}{b}\log(1-\frac{h_i}{M}))\, Y_i|\underset{\sim}{Y}\}$$
$$= Y_iE\{\hat{\lambda}_i - \lambda_i + \hat{Z}_i - Z_i|Y_i\}$$

$$= Y_iE\{(\hat{Z}_i(0) - Z_i(0))\lambda_i' + \hat{Z}_i - Z_i|\underset{\sim}{Y}) + O(n^{-\frac{1}{2}}).$$

Let

$$K_1 = E(\log f(\bar{h}\ W+I)|\bar{h}>0) + E(\log f(I)\lambda'),$$

$$K_2 = E(\log f'(\bar{h}\ W+I)|\bar{h}>0) + E(\log f'(I)\lambda'),$$

$$K_3 = EX_3$$

and

$$g = \frac{\beta - 1}{\sigma}.$$

Now from (4.1) it follows that

$$E\ \{(\hat{Z}_i(0) - Z_i(0))\lambda_i' + \hat{Z}_i - Z_i \mid Y_i = 1\}$$

$$= (\hat{g}-g)K_1 + (\hat{\sigma}^{-1}-\sigma^{-1})K_2 + K_3(\hat{\delta}-\delta) + O(n^{-\frac{1}{2}}).$$

Accordingly, the second term on the right hand side of (5.6) can be written as

$$(5.8)\qquad \text{Var}\ \{n_1^{-\frac{1}{2}} \sum_i E(\hat{\lambda}_i + \hat{Z}_i - \frac{a}{b}\log(1 - \frac{h_i}{M}))\ Y_i \mid \underset{\sim}{Y})$$

$$= \text{Var}\ \{n_1^{\frac{1}{2}}[\hat{g}-g)K_1 + (\hat{\sigma}^{-1}-\sigma^{-1})K_2 + K_3(\hat{\delta}-\delta)]\} + O(\frac{1}{n})$$

$$= K\ \Sigma\ K'd + O(\frac{1}{n})$$

where

$K = (K_1, K_2, K_3)$,

Σ is the asymptotic covariance matrix of

$n^{\frac{1}{2}}(\hat{g}-g,\ \hat{\sigma}^{-1} - \sigma^{-1},\ \hat{\delta}-\delta)$

and

$$d = \underset{\substack{n_1 \to \infty \\ n \to \infty}}{\text{plim}} \frac{n_1}{n},\quad 0 < d < 1.$$

Combining (5.6), (5.7) and (5.8) yields

$$(5.9)\qquad \tau_2^{\ 2} = \tau_1^{\ 2} + d\ K\ \Sigma\ K'.$$

Accordingly, we have demonstrated that when the stochastic parameter estimates from stage three are accounted for we have

$$(5.10) \qquad n_1^{\frac{1}{2}}\left(\frac{\hat{a}}{\hat{b}} - \frac{a}{b}\right) \underset{n \to \infty}{\to} N\left(0, \frac{\tau_2}{b}\right).$$

The asymptotic variance, τ_2^2, can be estimated in a straight forward manner as follows. Let

$$(5.11) \quad \hat{\tau}_1^2 = \frac{1}{n_1} \sum_{i \in \Omega} (1 - \hat{Z}_i(0)\hat{\lambda}_i - \hat{\lambda}_i^2)$$

$$(5.12) \quad \hat{K}_1 = \frac{1}{n_1} \sum_{i \in \Omega} \log f\,(h_i W_i + I_i) - \frac{1}{n} \sum_i (\hat{\lambda}_i^2 + \hat{Z}_i(0)\hat{\lambda}_i) \log f\,(I_i)$$

$$(5.13) \quad \hat{K}_2 = \frac{1}{n_1} \sum_{i \in \Omega} \log f'(h_i W_i + I_i) - \frac{1}{n} \sum_i (\hat{\lambda}_i^2 + \hat{Z}_i(0)\hat{\lambda}_i) \log f'(I_i)$$

and

$$(5.14) \quad \hat{K}_3 = \frac{1}{n} \sum_i X_{i3}.$$

Notice that in (5.12) and (5.13) we have exploited the fact that $\lambda'(x) = -\lambda(x) - x\lambda(x)^2$. From (5.11) - (5.14) it is clear that τ_2^2 can be consistently estimated by

$$(5.15) \quad \hat{\tau}_2^2 = \hat{\tau}_1^2 + \frac{n_1}{n} \hat{K} \hat{\Sigma} \hat{K}'$$

where $\hat{K} = (\hat{K}_1, \hat{K}_2, \hat{K}_3)$ and $\hat{\Sigma}$ is the estimated asymptotic covariance matrix of the stage three parameter estimates.

The asymptotic variance of

$$\hat{\alpha} - 1 = \frac{\hat{\sigma}\hat{a}}{\hat{b}}$$

can be derived in a completely analogous manner. Let $\hat{\tau}_3^2$ denote the estimator of the asymptotic variance

$$\tau_3^2 = b^2 \lim_{n \to \infty} \text{Var} \left\{ n_1^{\frac{1}{2}} \left(\frac{\hat{\sigma}\hat{a}}{\hat{b}} - \frac{\sigma a}{b} \right) \right\}.$$

Then

$$(5.16) \quad \hat{\tau}_3^2 = \hat{\sigma}^2 \hat{\tau}_1^2 + \frac{n_1}{n} \tilde{K} \tilde{\Sigma} \tilde{K}'$$

where

$$\tilde{K} = (\hat{K}_1, \hat{K}_3)$$

and $\tilde{\Sigma}$ is the asymptotic covariance matrix of $n^{\frac{1}{2}}(\hat{\beta}-\beta, \hat{\sigma}\hat{\delta} - \sigma\delta)$.

Finally we consider the sampling properties of the estimator $\widehat{\sigma_1 \rho}$ and $\hat{\theta}$ for $\sigma_1 \rho$ and Θ in (4.6). We demonstrated above that the conditional variance of e_i given X_{1i} and V_i does not depend on these variables. However since we use $\hat{V}_i$ instead of V_i the real error term is

$$e_i^* = e_i + c(\hat{V}_i - V_i)$$

where $c = \sigma_1 \rho$. The problem is therefore to obtain the limiting distribution of

$$n_1^{\frac{1}{2}} \begin{bmatrix} \hat{\theta}-\theta \\ \hat{c}-c \end{bmatrix} = B_n \, n_1^{-\frac{1}{2}} \begin{bmatrix} \sum_{i\varepsilon\Omega} X_{1i}' e_i^* \\ \sum_{i\varepsilon\Omega} \hat{V}_i e_i^* \end{bmatrix}$$

where

$$B_{n_1} = n_1 \begin{bmatrix} \sum_{i\varepsilon\Omega} X_{1i}' X_{1i}, & \sum_{i\varepsilon\Omega} X_{1i}' \hat{V}_i \\ \sum_{i\varepsilon\Omega} X_{1i} \hat{V}_i, & \sum_{i\varepsilon\Omega} \hat{V}_i^2 \end{bmatrix}^{-1}.$$

Under general conditions of the regressors (Amemiya, 1973)) we have

$$(5.17) \quad B = \operatorname*{plim}_{n\to\infty} B_{n_1} = \operatorname*{plim}_{n\to\infty} n_1 \begin{bmatrix} \sum_{i\varepsilon\Omega} X_{1i}' X_{1i}, & -\sum_{i\varepsilon\Omega} X_{1i}' \lambda_i \\ \sum_{i\varepsilon\Omega} X_{1i} \lambda_i, & \sum_{i\varepsilon\Omega} (1 - Z_i(0)\lambda_i) \end{bmatrix}^{-1}$$

since by (3.10) and (3.11)

$$E(V_i^{\,2}|X_i) = \lambda_i^{\,2} + Var\ (V_i|X_i) = 1-Z_i(0)\lambda_i .$$

Note that we can write

$$\hat{V}_i - V_i = D_i(\hat{\kappa}-\kappa)$$

where $D_i = (D_{i1}, D_{i2}, D_{i3}, D_{i4})$ is defined by

$$D_{i1} = \log f(h_i W_i + I_i),$$

$$D_{i2} = \log f'(h_i W_i + I_i),$$

$$D_{i3} = -\log(1-\frac{h_i}{M}),$$

$$D_{i4} = X_{i3}$$

and

$$\kappa = (\frac{\beta-1}{\sigma}, \frac{1}{\sigma}, \frac{a}{b}, \delta)'.$$

Hence

$$n_1^{-\frac{1}{2}} \sum_{i\varepsilon\Omega} X_{1i}^{'}(\hat{V}_i-V_i) = (\frac{\sum_i Y_i X_{1i}^{'} D_i}{n})\, n^{\frac{1}{2}}\, (\hat{\kappa}-\kappa) \cdot (\frac{n}{n_1})^{\frac{1}{2}}.$$

By the strong law of large numbers

$$\frac{\sum_i Y_i X_{1i}^{'} D_i}{n} \underset{n\to\infty}{\overset{a.s.}{\to}} E(Y_i X_{1i}^{'} D_i) = E(YX_1^{'}D).$$

Accordingly

$$(5.18)\quad n_1^{-\frac{1}{2}} \sum_{i\varepsilon\Omega} X_{1i}^{'} e_i^{*} \sim d^{-\frac{1}{2}} cE(YX_1^{'}D)\, n^{\frac{1}{2}}\, (\hat{\kappa}-\kappa) + \eta_1$$

where

$$(5.19)\quad \eta_1 \sim N(0;\gamma_1),\ \gamma_1 = \sigma_1^{\,2}(1-\rho^2)\ \text{plim} \sum_{i\varepsilon\Omega} \frac{X_{1i}^{'} X_{1i}}{n_1} = d^{-1} E(YX_1^{'}X_1)\sigma_1^{\,2}(1-\rho^2).$$

Similarly

$$(5.20)\quad n_1^{-\frac{1}{2}} \sum_{i\varepsilon\Omega} V_i e_i^* \sim d^{-\frac{1}{2}} cE(YVD)\, n^{\frac{1}{2}}(\hat{\kappa}-\kappa) + \eta_2$$

where

$$(5.21)\quad \eta_2 \sim N(0,\gamma_2^2),\ \gamma_2^2 = \sigma_1^2(1-\rho^2)\ \text{plim} \sum_{i\varepsilon\Omega} \frac{V_i^2}{n_1} = \sigma_1^2(1-\rho^2)\ (\tau_1^2$$

$$+\ \text{plim} \sum_{i\varepsilon\Omega} \lambda_i^2/n_1) = \sigma_1^2(1-\rho^2)\ \text{plim} \sum_{i\varepsilon\Omega} (1-Z_i(0)\lambda_i)/N_1.$$

The variables $\hat{\kappa}-\kappa$ and η_j are uncorrelated because

$$E((\hat{\kappa}-\kappa)e_i) = EE((\hat{\kappa}-\kappa)e_i \,|\, \underset{\sim}{Y}) = E((\hat{\kappa}-\kappa)E(e_i \,|\, \underset{\sim}{Y}) = 0.$$

The covariance matrix, Q, of

$$\begin{bmatrix} cd^{-\frac{1}{2}}E(YX_1'D)n^{\frac{1}{2}}(\hat{\kappa}-\kappa) + \eta_1 \\ \\ cd^{-\frac{1}{2}}E(YVD)n^{\frac{1}{2}}(\hat{\kappa}-\kappa) + \eta_2 \end{bmatrix}$$

has the form

$$(5.22)\quad Q = \sigma_1^2(1-\rho^2)B^{-1} + c^2d^{-1} \begin{bmatrix} E(YX_1'D)\Sigma E(YD'X_1), & E(YX_1'D)\Sigma E(YVD') \\ E(YVD)\Sigma E(YD'X_1), & E(YVD)\Sigma E(YVD') \end{bmatrix}$$

Consequently

$$(5.23)\quad \sqrt{n_1} \begin{bmatrix} \dot{\hat{\theta}}-\theta \\ \hat{c}-c \end{bmatrix} \sim N(0,\ BQB').$$

6. Comparison with the Heckman procedure for estimating for wage equation

Above we mentioned that the procedure for estimating the wage equation proposed here intuitively appears to be more efficient than Heckman's method. This is true when the effect of the inserted stochastic parameter estimates from stage one is small. If the influence of these parameters estimates is large the conclusion is not clear.

In the present section we shall only consider the relationship to the Heckman procedure in the particular case where we ignore the effect of stochastic parameter estimates.

Also we shall ignore the heteroschedasticity problem in Heckman's method by replacing η_i in Heckman's formulae for ϕ by $1-\rho^2$.

Let Γ and $\tilde{\Gamma}$ denote the asymptotic covariance matrices for our and Heckman's estimators, respectively.

Then

$$(6.1) \qquad \Gamma - \tilde{\Gamma} = \sigma_1{}^2\,(1-\rho^2)\,(B-\tilde{B})$$

where B is defined by (5.17) and $\tilde{B}$ is denoted B in Heckman's paper. By inspection we realize that

$$(6.2) \qquad B^{-1} = \tilde{B}^{-1} + tR$$

where

$$R = \begin{bmatrix} 0\ldots0 \\ \vdots \\ 0\ldots1 \end{bmatrix}$$

and

$$t = d^{-1}\,\mathrm{Var}(YV^2) = d^{-1}\tau_1{}^2 .$$

Standard manipulation of (6.2) yields

$$(6.3) \qquad B = (\tilde{B}^{-1} + tR)^{-1} = \tilde{B}\,(I-(I+tR\tilde{B})^{-1}tR\tilde{B}).$$

Furthermore

$$(6.4)\qquad (I+t\tilde{R}\tilde{B})^{-1}\, t\tilde{R}\tilde{B} = \frac{t}{1+t\tilde{B}_{mm}} \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ \tilde{B}_{m1} & \tilde{B}_{m2},\ldots & \tilde{B}_{mm} \end{bmatrix}$$

where m is the dimension of $\tilde{B}$ and B. Let $\hat{\theta}$ and $\hat{c}$ denote the estimates obtained by Heckman's procedure. Then (5.22), (5.23), (6.2) and (6.4) imply that

$$\text{Var}\ \hat{c} = \frac{\sigma_1^2(1-\rho^2)\tilde{B}_{mm}}{1 + t\tilde{B}_{mm}}$$

and

$$\text{Var}\ \hat{\theta}_j = \sigma_1^2(1-\rho^2)\left(\tilde{B}_{jj} - \frac{t\tilde{B}_{mj}^2}{1 + t\tilde{B}_{mm}}\right).$$

Let

$$t^* = \frac{t}{\sigma_1^2(1-\rho^2)}.$$

Then the variance expressions above can be expressed as

$$(6.5)\qquad \text{Var}\ \hat{c} = \frac{\text{Var}\tilde{c}}{1+t^*\text{Var}\tilde{c}}$$

$$(6.6)\qquad \text{Var}\ \hat{\theta}_j = \text{Var}\ \tilde{\theta}_j - \frac{\text{cov}(\tilde{c}, \tilde{\theta}_j)^2\, t^*}{1+t^*\ \text{Var}\ \tilde{c}} = \text{Var}\ \tilde{\theta}_j(1-\text{corr}(\tilde{c}, \tilde{\theta}_j)\ t^*\text{Var}\ \hat{c}).$$

Formulae (6.6) tells us that the variance reduction increases when the correlation between the estimators for c and θ increases or if the variance of c increases (provided corr $(\tilde{c}, \tilde{\theta}_j)$ is fixed).

References

Amemiya, T. (1973): "Regression Analysis when the Dependent Variable is Truncated Normal". Econometrica, 41, 997-1017.

Dagsvik, J.K. (1987): "A Simplified Version of Heckman's two Stage Estimation Procedure that is Applicable when the Selectivity Effect is Strong." Discussion Paper, nr. 113. University of Heidelberg.

Dagsvik, J.K., K. Holst, S. Strøm, G. Wagenhals and J. Østervold (1987a): "Female Labor Supply in West-Germany." Memorandum, Dept. Economics, University of Oslo.

Dagsvik, J.K., F. Laisney and S. Strøm (1987b): "Female Labor Supply and the Tax Benefit system in France 1979". Memorandum, Dept. Economics, University of Oslo.

Heckman, J.J. (1979): "Sample Selection Bias as a Specification Error". Econometrica, nr. 47, 153-162.

Judge, G.G., Griffiths, W.E., Hill, R.C., Lütkepohl, H., and Lee, T.C. (1980): "The theory and practice of econometrics," Wiley, New York.

ISSUED IN THE SERIES DISCUSSION PAPER

No. 1 I. Aslaksen and O. Bjerkholt: Certainty Equivalence Procedures in the Macroeconomic Planning of an Oil Economy.

No. 3 E. Biørn: On the Prediction of Population Totals from Sample surveys Based on Rotating Panels.

No. 4 P. Frenger: A Short Run Dynamic Equilibrium Model of the Norwegian Prduction Sectors.

No. 5 I. Aslaksen and O. Bjerkholt: Certainty Equivalence Procedures in Decision-Making under Uncertainty: an Empirical Application.

No. 6 E. Biørn: Depreciation Profiles and the User Cost of Capital.

No. 7 P. Frenger: A Directional Shadow Elasticity of Substitution.

No. 8 S. Longva, L. Lorentsen, and Ø. Olsen: The Multi-Sectoral Model MSG-4, Formal Structure and Empirical Characteristics.

No. 9 J. Fagerberg and G. Sollie: The Method of Constant Market Shares Revisited.

No.10 E. Biørn: Specification of Consumer Demand Models with Stocahstic Elements in the Utility Function and the first Order Conditions.

No.11 E. Biørn, E. Holmøy, and Ø. Olsen: Gross and Net Capital, Productivity and the form of the Survival Function . Some Norwegian Evidence.

No.12 J. K. Dagsvik: Markov Chains Generated by Maximizing Components of Multidimensional Extremal Processes.

No.13 E. Biørn, M. Jensen, and M. Reymert: KVARTS - A Quarterly Model of the Norwegian Economy.

No.14 R. Aaberge: On the Problem of Measuring Inequality.

No.15 A-M. Jensen and T. Schweder: The Engine of Fertility - Influenced by Interbirth Employment.

No.16 E. Biørn: Energy Price Changes, and Induced Scrapping and Revaluation of Capital - A Putty-Clay Approach.

No.17 E. Biørn and P. Frenger: Expectations, Substitution, and Scrapping in a Putty-Clay Model.

No.18 R. Bergan, Å. Cappelen, S. Longva, and N. M. Stølen: MODAG A - A Medium Term Annual Macroeconomic Model of the Norwegian Economy.

No.19 E. Biørn and H. Olsen: A Generalized Single Equation Error Correction Model and its Application to Quarterly Data.

No.20 K. H. Alfsen, D. A. Hanson, and S. Glomsrød: Direct and Indirect Effects of reducing SO_2 Emissions: Experimental Calculations of the MSG-4E Model.

No.21 J. K. Dagsvik: Econometric Analysis of Labor Supply in a Life Cycle Context with Uncertainty.

No.22 K. A. Brekke, E. Gjelsvik, B. H. Vatne: A Dynamic Supply Side Game Applied to the European Gas Market.

No.23 S. Bartlett, J. K. Dagsvik, Ø. Olsen and S. Strøm: Fuel Choice and the Demand for Natural Gas in Western European Households.

No.24 J. K. Dagsvik and R. Aaberge: Stochastic Properties and Functional Forms in Life Cycle Models for Transitions into and out of Employment.

No.25 T. J. Klette: Taxing or Subsidising an Exporting Industry.

No.26 K. J. Berger, O. Bjerkholt and Ø. Olsen: What are the Options for non-OPEC Producing Countries.

No.27 A. Aaheim: Depletion of Large Gas Fields with Thin Oil Layers and Uncertain Stocks.

No.28 J. K. Dagsvik: A Modification of Heckman's Two Stage Estimation Procedure that is Applicable when the Budget Set is Convex.