MD

91-15

Working Paper

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS
SYMONS HALL
UNIVERSITY OF MARYLAND
COLLEGE PARK 20742

[redacted]

Estimation of Endogenous Switching Regression Models with Discrete Dependent Variables: An Application to the Estimation of Farm Women's Farm and Off-Farm Participation Equations

Ayal Kimhi

Department of Agricultural and Resource Economics
University of Maryland
College Park, Maryland 20742

Working Paper No. 91-15

June 1991

Estimation of Endogenous Switching Regression Models with Discrete Dependent Variables: An Application to the Estimation of Farm Women's Farm and Off-Farm Participation Equations

by

Ayal Kimhi*

Department of Agricultural and Resource Economics

University of Maryland

June 1991

Abstract

A two-step procedure for estimating a switching regression model is developed for the case in which all endogenous variables are only observed in binary form. The heteroskedastic residuals of the second stage equations are normalized and the resulting nonlinear equations are estimated by limited information maximum likelihood methods. This procedure is applied to the estimation of participation equations of farm women in farm and off-farm work, and the estimators were in general not significantly different from full information maximum likelihood estimators.

---

* I thank David Stapleton, John Horowitz and Erik Lichtenberg for helpful discussions. The research was supported by the Department of Agricultural and Resource Economics, University of Maryland, by Maryland Agricultural Experiment Station, and by the Center for Agricultural Economic Research, the Hebrew University.

Estimation of Endogenous Switching Regression Models with Discrete Dependent Variables: An Application to the Estimation of Farm Women's Farm and Off-Farm Participation Equations

1. Introduction

This paper suggests an estimation method for a special case of the endogenous switching regression model, in which all dependent variables are discrete. This case is a variation of the "Multivariate Probit Model with Structural Shift" described by Heckman (1978), in which structural shift exists in one equation only (in a two-equation model).

Maddala (1983, p. 223) suggests a two-stage estimation method for the traditional switching regression model: estimating the selection equation in a first stage, calculating selection correction terms and inserting them into the other equations, which are then estimated in the second stage by least squares. When all the endogenous variables are discrete, the second stage of the estimation method has to involve maximum likelihood methods (Probit in our case). This is problematic since by construction, the modified stochastic terms are heteroskedastic. We show that this can be solved by dividing each equation by its standard deviation, which is a function of both first-stage and second-stage parameters. The resulting second-stage model is non-linear in the parameters.

We apply this method to the estimation of participation equations of farm women in farm work and off-farm work, in which the off-farm participation equation parameters depend on farm participation. We also show how to calculate the true covariance matrix of the second-stage estimators, using the Murphy-Topel (1985) method, and compare the estimation results to those obtained by maximum likelihood. The general model and the two-stage estimation procedure are described in section 2, and in section 3 the calculation of the true covariance matrix is presented. Section 4 includes the empirical application and its results, and some concluding remarks are in section 5.

2. The Model and the 2-Stage Estimation Procedure

Heckman (1978) discusses the following model:

(1a) $$y_1^* = X_1 \cdot \alpha_1 + d_1 \cdot \beta_1 + y_2^* \cdot \gamma_1 + U_1$$

(1b) $$y_2^* = X_2 \cdot \alpha_2 + d_1 \cdot \beta_2 + y_1^* \cdot \gamma_2 + U_2$$

(1c) $$d_1 = 1 \quad \text{iff} \quad y_1^* > 0$$

$$d_1 = 0 \quad \text{otherwise}$$

Where the U_i's are $N(0,\sigma_i^2)$ random variables uncorrelated with the X_i's.

The endogenous switching regression model of Maddala (1983, p. 223) is derived from this model by assuming that y_1^* is unobserved, $\gamma_1 = \gamma_2 = \beta_1 = 0$, $\beta_2 = X_3 \cdot \alpha_3$, where $X_3 \supseteq X_2$, and that $U_2 = d_1 \cdot U_{22} + (1-d_1) \cdot U_{21}$, where the assumptions with regard to the U_i's apply for the U_{2i}'s as well. The resulting model is:

$$\text{(2a)} \qquad y_1^* = X_1 \cdot \alpha_1 + U_1$$

$$\text{(2b)} \qquad y_{21}^* = X_{21} \cdot \alpha_{21} + U_{21} \qquad \text{iff} \quad d_1 = 0$$

$$\text{(2c)} \qquad y_{22}^* = X_{22} \cdot \alpha_{22} + U_{22} \qquad \text{iff} \quad d_1 = 1$$

We assume further that the y_{2i}^*'s are unobserved. We observe d_{21} or d_{22} defined as:

$$\text{(2d)} \qquad \left.\begin{array}{lll} d_{21} = 1 & \text{iff} & y_{21}^* > 0 \\ d_{21} = 0 & \text{otherwise} & \end{array}\right\} \quad \text{iff} \quad d_1 = 0$$

$$\left.\begin{array}{lll} d_{22} = 1 & \text{iff} & y_{22}^* > 0 \\ d_{22} = 0 & \text{otherwise} & \end{array}\right\} \quad \text{iff} \quad d_1 = 1$$

The log likelihood function of the model described in (2a)-(2d) is:

$$\mathcal{J} = \Sigma \; d_1 \cdot d_{22} \cdot \ln\Phi(-A_1, -A_{22}, \rho_2) + d_1 \cdot (1-d_{22}) \cdot \ln\Phi(-A_1, A_{22}, -\rho_2) +$$
$$+ (1-d_1) \cdot d_{21} \cdot \ln\Phi(A_1, -A_{21}, \rho_1) + (1-d_1) \cdot (1-d_{21}) \cdot \ln\Phi(A_1, A_{21}, \rho_1)$$

where ρ_i is the correlation coefficient between U_1 and U_{2i} $(i=1,2)$, $A1=-X_1 \cdot \alpha_1/\sigma_1$, $A21=-X_{21} \cdot \alpha_{21}/\sigma_{21}$, and $A22=-X_{22} \cdot \alpha_{22}/\sigma_{22}$.

Maddala (1983) suggested a two-stage estimation method for his model. The first stage of this method is unchanged by the new assumption: estimate (2a) by probit to get estimates of α_1/σ_1. The second stage, however, must also be estimated by probit, in each of the subsamples defined by $d_1=1$ and $d_1=0$, respectively. In order to correct for selectivity, we write, following Johnson & Kotz (1970,

p. 81), and assuming that $\sigma_1 = \sigma_2 = 1$:[1]

(3a) $U_{21} = \rho_1 \cdot U_1 + u_1$

(3b) $U_{22} = \rho_2 \cdot U_1 + u_2$

(4a) $E(U_1|d_1=0) = E(U_1|U_1<-X_1 \cdot \alpha_1) = -\phi(-X_1 \cdot \alpha_1)/\Phi(-X_1 \cdot \alpha_1) \equiv \lambda_1$

(4b) $E(U_1|d_1=1) = E(U_1|U_1>-X_1 \cdot \alpha_1) = \phi(-X_1 \cdot \alpha_1)/(1-\Phi(-X_1 \cdot \alpha_1)) \equiv \lambda_2$

where the u_i's are independent of U_1. Therefore:

(5a) $E(U_{21}|d_1=0) = \rho_1 \cdot \lambda_1$

(5b) $E(U_{22}|d_1=1) = \rho_2 \cdot \lambda_2$

Define:

(6a) $\varepsilon_1 = U_{21} - E(U_{21}|d_1=0)$

(6b) $\varepsilon_2 = U_{22} - E(U_{22}|d_1=1)$

and put into (2b) and (2c) to get:

(7a) $y_{21}^* = X_{21} \cdot \alpha_{21} + \rho_1 \cdot \lambda_1 + \varepsilon_1$

(7b) $y_{22}^* = X_{22} \cdot \alpha_{22} + \rho_2 \cdot \lambda_2 + \varepsilon_2$

where:

---

[1] Two normalizations are necessary for identification since the latent variables y_{2i}^* are not observed.

(8a) $$E(\varepsilon_1|d_1=0) = 0$$

(8b) $$E(\varepsilon_2|d_1=1) = 0$$

In the model discussed by Maddala (1983), this is sufficient to get consistent estimators of α_{21} and α_{22}, after substituting the first stage estimator $\hat{\alpha}_1$ for α_1. In our case, since probit is used in the second stage, we can only identify $\alpha_{2i}/Var(\varepsilon_i)$ if $Var(\varepsilon_i)$ is identical across observations. However, by construction, $Var(\varepsilon_i)$ depends on X_1 (via λ_i), which varies across observations. ε_i is therefore heteroskedastic, and the second-stage probit is not valid. The solution is to calculate $Var(\varepsilon_i)$ explicitly, and use the normalized random variables $\varepsilon_i/[Var(\varepsilon_i)]^{1/2}$, which are distributed $N(0,1)$, instead of ε_i.

Using Johnson & Kotz (1970, p. 83) we get:

(9a) $$Var(U_1|d_1=0) = 1 + \lambda_1 \cdot (-X_1 \cdot \alpha_1 - \lambda_1)$$

(9b) $$Var(U_1|d_1=1) = 1 + \lambda_2 \cdot (-X_1 \cdot \alpha_1 - \lambda_2)$$

Using (3) and (6), and the fact that $Var(u_i) = 1 - \rho^2$:

(10a) $$Var(\varepsilon_1|d_1=0) = Var(U_{21}|d_1=0) = 1 + \rho_1{}^2 \cdot \lambda_1 \cdot (-X_1 \cdot \alpha_1 - \lambda_1) \equiv s_1{}^2$$

(10b) $$Var(\varepsilon_2|d_1=1) = Var(U_{22}|d_1=1) = 1 + \rho_2{}^2 \cdot \lambda_2 \cdot (-X_1 \cdot \alpha_1 - \lambda_2) \equiv s_2{}^2$$

Dividing (7a) and (7b) by s_1 and s_2, and substituting $\hat{\lambda}_1$ and $\hat{\lambda}_2$, which are calculated using $\hat{\alpha}_1$, for λ_1 and λ_2, respectively, we get the second stage probit equations, which are nonlinear in the parameters α_{2i} and ρ_i. Identification is supported by the following intuitive argument: conditional on ρ_i, α_{2i} is identified. Then, it is possible to estimate α_{2i} given different values of ρ_i, and choose the one that results in the highest likelihood value. This depends of course on the familiar condition that $X_{2i} \not\supset X_1$; $i=1,2$. A formal identification proof involves verifying the assumptions of Amemiya (1985, p. 115).[2]

3. Estimating the Correct Covariance Matrix

As Murphy & Topel (1985) pointed out, the second stage of a two-stage estimation procedure does not automatically produce the true asymptotic covariance matrix of the estimated parameters. The reason is that in the second stage, the first stage estimators are treated as if they were the true parameters, when in fact they are measured with error.

Murphy & Topel (1985) show that when the maximum likelihood method is used in both stages, the second stage estimators are asymptotically normal with covariance matrix equal to:

$$(11)\quad R_2^{-1} + R_2^{-1}\cdot[R_3'\cdot R_1^{-1}\cdot R_3 - R_4'\cdot R_1^{-1}\cdot R_3 - R_3'\cdot R_1^{-1}\cdot R_4]\cdot R_2^{-1}$$

where:

---

[2] It could be claimed that ε_i is not independent across observations, because of its dependence on $\hat{\lambda}_i$ which is calculated using all observations. However, if $\hat{\lambda}_i$ is consistent, ε_i can be said to be asymptotically independent, using an argument similar to that of Lee (1979, p. 984).

(12a) $R_1(\theta_1) = E\ (\partial L_1/\partial\theta_1)\cdot(\partial L_1/\partial\theta_1)'$

(12b) $R_2(\theta_2) = E\ (\partial L_2/\partial\theta_2)\cdot(\partial L_2/\partial\theta_2)'$

(12c) $R_3(\theta_1,\theta_2) = E\ (\partial L_2/\partial\theta_1)\cdot(\partial L_2/\partial\theta_2)'$

(12d) $R_4(\theta_1,\theta_2) = E\ (\partial L_1/\partial\theta_1)\cdot(\partial L_2/\partial\theta_2)'$.

In (12), L_1 (L_2) is the likelihood function of the first (second) stage with associated parameters θ_1 (θ_2). It is apparent that the asymptotic covariance matrix is different from R_2^{-1} because the dependence of L_2 on θ_1 is not taken into account in the estimation.

A consistent estimator of this covariance matrix can be obtained by replacing θ_1 and θ_2 with their consistent estimators, and replacing expectations with sample means. R_1 and R_2 are in fact the default covariance matrices of the first and second stages, respectively, and R_4 can be easily calculated using the estimation results. The calculation of R_3 is a little more time consuming.

4. An Empirical Example

As an example to the application of the econometric procedures described above, we use a model of off-farm labor force participation of farmers, in which the latent variable describing the tendency to participate, depends on farm participation, which is also determined by a latent variable crossing a threshold.

The theoretical model assumes utility maximization over consumption and leisure subject to time and budget constraints, where time can be productively used on or off the farm (Kimhi, 1991a). Formally, the optimization problem is:

$$\max_{Th,C,Tf,Tm} \quad U(Th\ ,\ C)$$

$$\text{s.t.} \quad \begin{aligned} &1.\quad C \le \pi(Tf) + W\cdot Tm + I \\ &2.\quad Th + Tf + Tm \le T \\ &3.\quad Tf \ge 0 \\ &4.\quad Tm \ge 0 \end{aligned}$$

where *Th*, *Tf* and *Tm* are time spent on home activities, farm work and off-farm work, respectively, *C* is consumption, *I* is non-earned income, *W* is the off-farm wage and π is the conditional variable profit function described by Lopez (1982).

Two of the Kuhn-Tucker necessary conditions for maximization are:

$$\pi_1 + \delta/U_2 = U_1/U_2 \tag{12a}$$

$$W + \phi/U_2 = U_1/U_2 \tag{12b}$$

where δ and ϕ are positive if and only if farm work and off farm work, respectively, are zero, and subscripts denote partial derivatives. If we proceed to solve (12a) and (12b) simultaneously, we end up with a usual simultaneous equations model. In order to get the switching regression structure, we assume that for some reason (long-run considerations, etc.) the farm participation problem is solved prior to the off-farm participation problem.[3]

[3] This is absolutely arbitrary, since it could have been the other way around, and there are arguments in favor of, as well as against, each of the two possibilities.

This results in the conditioning described in (2), and therefore farm participation is determined solely by (12a).[4]

For those who work on farm, $\delta=0$, $\pi_1 = U_1/U_2$, and off-farm participation occurs if:

$$(13a) \qquad W > \pi_1(Tf^*),$$

assuming all sufficient conditions are met, where Tf^* denotes optimal farm labor supply given no off-farm work. For those who don't work on the farm, participation occurs if:

$$(13b) \qquad W > U_1(I,T)/U_2(I,T).$$

This leads to the following off-farm participation index function:

$$(14) \qquad Y^* = W - RHS$$

$$Y = \begin{cases} 1 & \text{if } Y^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

where *RHS* is the right hand side of equations (13a) and (13b) for farm participants and non-participants, respectively. It is clear that when specifying Y^* as a function of observable variables, this function will depend on farm participation.

In terms of our general model (2), we get:

---

[4] This does not contradict the possibility that the stochastic terms U_{2i} $i=1,2$ are drawn before the farm participation decision is made, which is necessary for the validity of the switching regression model (Poirier & Ruud, 1981).

(2a)' $y_1^* = \pi_1(0) - U_1(I,T)/U_2(I,T)$

(2b)' $y_{21}^* = W - U_1(I,T)/U_2(I,T)$

(2c)' $y_{22}^* = W - \pi_1(Tf^*)$,

and (2d). We specify these unobserved latent variables as linear combinations of explanatory variables, including personal, family and farm characteristics, and use data on farm women from the 1981 Census of Agriculture in Israel in order to estimate the model.[5] Descriptive statistics of the data set are reported in table 1.

The results of the two-stage estimation procedure are compared to maximum likelihood results in table 2. In practice, maximum likelihood estimation was not much slower that the two-stage procedure, but because the correlation coefficient was close to -1 in one of the subsamples, it failed when arbitrary starting values were used. Maximum likelihood was successful only when the consistent estimators of the two-stage procedure were used as starting values, and this demonstrates the importance of that procedure in this particular case.

Let's look at farm participation first. The first column in table 2 presents univariate probit estimators, and the second column presents the ML estimators (joint with off-farm participation).[6] We find that the difference between them is only

---

[5] We chose farm women because they were more equally divided into subsamples according to farm participation: only 10% of farm men did not work on farm, compared to 59% of farm women.

[6] The farm participation equation also included a set of regional dummies and a set of village establishment year dummies. The former should have actually been included in the off-farm

marginal, which is not surprising, since univariate probit is equivalent to Quasi-ML (Avery et al., 1983) in this case. Age profiles of farm participation are concave as expected, with participation probability peaking around the age of 47. Schooling has a positive and significant effect. The number of other family members in all age groups affects farm participation negatively, with adults having a greater effect than children. In dairy farms, farm participation of farm women is much higher, which is expected, since dairy farm work is known to be a good complement for house work. Land has a negative effect on participation; in larger farms, women have a lower tendency to work on farm. It could be that hired labor substitutes for family members in larger farms, and the income effect may play a role here too. In contrast, capital stock has a positive effect on farm participation.[7]

Now we turn to the off-farm participation results. Comparing the 2-stage with the ML results, we see that they are not as close as in the case of farm participation. Still, other than the correlation coefficients, the difference between the two sets of

---

participation equation as well (see Tokle & Huffman, 1990), but we excluded it for the purpose of identification of the second stage parameters, as discussed earlier in section 2.

[7] We only included capital assets at least ten years old, in order to avoid the problem of endogeneity of capital stock in the time allocation decision. Also, the land variable is the original land allotment of the farm, which was institutionally determined at the time of establishment of the village. The dairy farm dummy is also considered exogenous since strict milk quotas and large subsidies have kept the subset of farms that produce milk relatively stable over time (because of the endogeneity problem, the number of milk cows and other farm attributes were not used as explanatory variables).

estimators is pretty small. The correlation coefficient was underestimated (in absolute value) by the 2-stage method in both cases. Comparing the off-farm participation equations for the two subsets of farm women, we first notice that the correlation coefficient between the stochastic terms of the farm and off-farm participation equations is close to -1 in the farm non-workers equation, in contrast to a negative but non-significant correlation coefficient for farm workers. We actually expected to find the contrary, as we found elsewhere for farm men (Kimhi, 1991b). We don't have any convincing explanation for this result.[8]

The coefficients of the personal characteristics are not significantly different between the two subsets. Participation probability as a function of age peaks slightly later for farm workers (at the age of 35 versus 33 for non-workers), and in both cases off-farm participation probability peaks much earlier than farm participation probability (47), and declines much faster afterwards. The schooling coefficient is positive and significant, and is approximately twice as large as the schooling coefficient in the farm participation equation, which means that schooling, at least as measured here, contributes more to off-farm earnings than to farm productivity. These results are very much in line with existing studies (Lass et al., 1991), with the exception that age profiles of off-farm participation probability are more concave and peak earlier than in the other studies (between ages 45 and 55).

---

[8] In fact, it undermines the specification of the decision process.

The number of children decreases off-farm participation probability, and the number of adults increases it, in both subsets. These effects are stronger in the non-workers equation (with the exception of family members over 65 years old).

The major difference between the two subsets lies in the coefficients of farm attributes. Land has a negative coefficient for farm workers and positive for non-workers. The dairy farm dummy has a negative coefficient in both cases, but much larger in absolute value for farm workers. Capital stock has a negative and significant coefficient in the non-workers equation, and a positive and non-significant coefficient in the workers equation. The existence of these differences is expected (Kimhi, 1991b), since for those who work on farm, farm attributes affect farm labor demand and therefore affect off-farm labor supply through the time constraint. For those who don't work on the farm, the effect is only through the budget constraint (i.e., both income and substitution effects exist for farm workers, but only income effect for non-workers). We don't have an explanation for the positive coefficient of land in the non-workers equation, but it is evident that the substitution effect and the income effect work in the same direction in the workers' equation, since in larger farms and dairy farms (where family labor demand is relatively higher), farm women have a lower tendency to work on the farm. This last finding is in line with farm men's participation results reported elsewhere (Lass et al., 1991).

Finally, we tried to evaluate the unconditional marginal effects of the explanatory variables on the latent off-farm participation patterns of farm women (Huang et al., 1991), rather than the partial effect represented by the estimated coefficients. The results are qualitatively unchanged with respect to personal and family characteristics. Land size has a small, probably not statistically significant, positive effect on the marginal off-farm participation tendency.[9]

5. Summary and Conclusions

The usual two-step procedure for the estimation of an endogenous switching regression model is inappropriate when only qualitative realizations of the dependent variables are observed. This is because after correcting for selectivity, the residuals are conditionally heteroskedastic. In this paper an alternative method is proposed, according to which the residual is normalized by its calculated (observation specific) standard deviation, which is a function of both first and second stage parameters. This allows the second stage regressions, which are nonlinear in the parameters, to be estimated by ML probit. We also show how to calculate the true covariance matrix of the second stage estimators.

We demonstrated this method by applying it to the estimation of participation equations of farm women in farm and off-farm work,

---

[9] The term "probably" means that standard errors of the unconditional marginal effects were not derived, and the assertion is based on the standard errors of the relevant coefficients in the switching regression. The calculations were performed using sample means of the latent variables determining farm and off-farm participation, of $\phi(-X_1 \cdot \alpha_1)$ and of $\Phi(-X_1 \cdot \alpha_1)$.

where the farm participation equation serves as the selection equation. The results are compared to ML results and found to be quite close. In fact, in this particular problem ML convergence could not have been achieved without using the initial consistent estimators of the two-stage procedure, and this indicates the usefulness of the method.

Topics left for future research include a formal identification proof, and testing against alternative possible specifications of the decision process.

Table 1

Descriptive Statistics

a. Explanatory Variables

Variable	Mean by Farm Participation All	Workers	Non-Workers	Range	Units
Age	43.6	43.0	44.0	14-80	years
$(Age)^2/50$	41.3	40.0	42.3	4-128	$(years)^2/50$
Schooling	8.6	9.4	8.1	0-20	years
Family 0-14[a]	1.6	1.6	1.7	0-11	heads
Family 15-21	.88	.78	.95	0-8	heads
Family 22-65	2.4	2.2	2.4	0-10	heads
Family 66+	.23	.19	.26	0-3	heads
Total Land[b]	3.1	3.0	3.2	0-8	ln(dunams)[c]
Dairy Farm	.08	.09	.07	0-1	dummy
Old Capital[d]	1.0	1.1	.97	0-7.3	ln($81)[e]

b. Participation

	Working Off-Farm	Not Working	Total
Working on Farm	1075 (6%)	5866 (35%)	6941 (41%)
Not Working	2580 (15%)	7494 (44%)	10074 (59%)
Total	3655 (21%)	13360 (79%)	17015 (100%)

[a] Number of family members in each age group.
[b] Original land allotment.
[c] 1 dunam = 0.23 acre.
[d] Normative value of capital assets at least ten years old.
[e] In 1981 prices.

Table 2: A Comparison of Two-Stage and ML Estimators

Variable	Farm[a]		Off-Farm (Farm Workers)		Off-Farm (Non-Workers)	
	2-Stage	ML	2-Stage	ML	2-Stage	ML
Intercept	-2.07 (13.)	-2.07 (14.)	-2.72 (6.9)	-2.52 (6.1)	-2.98 (13.)	-3.14 (17.)
Family 0-14	-.025 (3.3)	-.022 (2.9)	-.066 (4.1)	-.065 (4.1)	-.087 (8.4)	-.090 (9.0)
Family 15-21	-.045 (4.5)	-.045 (4.5)	.012 (.54)	.014 (.66)	-.014 (1.1)	-.019 (1.3)
Family 22-65	-.105 (9.7)	-.104 (9.8)	.021 (.90)	.024 (1.0)	.045 (3.1)	.043 (2.9)
Family 65+	-.099 (3.9)	-.095 (3.9)	.179 (3.6)	.181 (3.6)	.154 (4.2)	.157 (4.7)
Land[b]	-.036 (2.7)	-.039 (3.1)	-.108 (4.2)	-.098 (3.8)	.079 (5.2)	.091 (6.2)
Dairy Farm	.354 (9.2)	.340 (8.9)	-.438 (4.9)	-.449 (5.1)	-.160 (3.0)	-.177 (3.3)
Old Capital	.025 (4.5)	.024 (4.2)	.006 (.52)	.005 (.42)	-.034 (4.5)	-.035 (4.9)
Age	.086 (13.)	.085 (13.)	.091 (5.5)	.085 (5.0)	.073 (6.9)	.078 (8.6)
Schooling	.034 (12.)	.036 (13.)	.082 (13.)	.078 (11.)	.082 (14.)	.087 (23.)
$(Age)^2/50$	-.045 (12.)	-.045 (13.)	-.064 (6.7)	-.061 (6.2)	-.055 (8.7)	-.059 (11.)
Correlation	-----	-----	-.187 (1.6)	-.270 (2.1)	-.860 (17.)	-.925 (3.7)

Notes: t-ratios in parenthesis. ML: asymptotic; 2-Stage: actual.

[a] Farm equation also included sets of regional dummies and village establishment year dummies.

[b] Land and capital stock were measured in natural logarithms to minimize the effects of outliers. A normalization was used such that a zero remained a zero.

References

Amemiya, T. (1985), *Advanced Econometrics*, Cambridge: Harvard University Press.

Avery, R.B., Hansen L.P. & Hotz, V.J. (1983), "Multiperiod Probit Models and Orthogonality Condition Estimation," *International Economic Review 24(1)*: 21-35.

Heckman, J.J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica 46(6)*: 931-59.

Huang, C.L., Raunikar, R. & Misra, S. (1991), "The Application and Economic Interpretation of Selectivity Models", *American Journal of Agricultural Economics 73(2)*: 496-501.

Johnson, N.L. & Kotz, S. (1970), *Distributions in Statistics: Continuous Univariate Distributions-2*, Boston: Houghton Mifflin.

Kimhi, A. (1991a), *Occupational Choice in Israeli Cooperative Villages*, Ph.D. Dissertation, The University of Chicago, Chicago, Illinois.

Kimhi, A. (1991b), *The Relevance of the Extent of Farm Work to the Analysis of Off-Farm Labor Supply of Farmers*, Working Paper No. 91-11, Department of Agricultural and Resource Economics, University of Maryland.

Lass, D.A., Findeis, J.L. & Hallberg, M.C. (1991, forthcoming), "Factors Affecting the Supply of Off-Farm Labor: A Review of Empirical Evidence," in *Multiple Job-Holding Among Farm Families*, ed. Hallberg, M.C., Findeis, J.L. & Lass, D.A., Ames: Iowa State University Press.

Lee, L.F. (1979), "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables," *Econometrica 47(4)*: 977-95.

Lopez, R.E. (1982), "Applications of Duality Theory to Agriculture," *Western Journal of Agricultural Economics 7(2)*: 353-65.

Maddala, G.S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

Murphy, K.M. & Topel, R.H. (1985), "Estimation and Inference in Two-Step Econometric Models," *Journal of Business & Economic Statistics 3(4)*: 370-9.

Poirier, D.J. & Ruud P.A. (1981), "On the Appropriateness of Endogenous Switching," *Journal of Econometrics 16(2)*: 249-56.

Tokle, J.G. & Huffman, W.E. (1990), *Local Labor Market Conditions: effects on Labor Demand and Wage Labor Supply Decisions of Farm and Rural Nonfarm Couples, 1978-82*, Staff Paper No. 215, Iowa State University, Ames, Iowa.