

MIXED LOGIT AND FULLY RECURSIVE LOGIT MODELS

by

Lung-Fei Lee

Discussion Paper No. 79-120, October 1979

Center for Economic Research  
Department of Economics  
University of Minnesota  
Minneapolis, Minnesota 55455

### Abstract

Mixed logit model has been proposed for modelling simultaneous effects of qualitative and continuous responses in the econometrics literatures. However, we point out that the concept of simultaneity does not apply in that model. It is an equivalent representation of a fully recursive logit model.

# MIXED LOGIT AND FULLY RECURSIVE LOGIT MODELS\*

Lung-Fei Lee\*\*

## 1. Introduction

Econometricians have developed multiple equation models to describe the simultaneous effects among endogenous variables in a system. The formulations have been fully developed for continuous endogenous variables (see, e.g. Theil [1971]). Nerlove and Press [1973] and Schmidt and Strauss [1975] develop multiple equations models for qualitative endogenous variables. A simple representative of their formulation is the two equation model:

$$\ln\left[\frac{P(X=1|Y)}{P(X=0|Y)}\right] = Z\beta + \alpha_1 Y$$
$$\ln\left[\frac{P(Y=1|X)}{P(Y=0|X)}\right] = Z\gamma + \alpha_2 X$$
(1)

where  $X$  and  $Y$  are dichotomous dependent variables  $P(X|Y)$  is the probability of  $X$  given  $Y$ , and  $Z$  is a vector of explanatory variables. Nerlove and Press refer to (1) as multivariate logit model. Schmidt and Strauss refer to (1) as a simultaneous logit model. The authors point out that the equations imply  $\alpha_1 = \alpha_2$  and there is no identification problem analogous to the classical simultaneous equations model for continuous endogenous variables. The constraint,  $\alpha_1 = \alpha_2$ , in this model does not

---

\* I appreciate comments by R. Hoffbeck.

\*\* Department of Economics, University of Minnesota, 1035 Business Administration, 271 19th Avenue South, Minneapolis, Minnesota 55455.

allow for separate effects of the endogenous variables in the system. Nerlove and Press have explicitly derived their model from the log linear models developed by statisticians for analyzing contingency tables (see e.g. Bishop et al [1975]). Brier [1979] strongly emphasizes that the previous analysis is comparable to the use of a correlation coefficient to determine the simultaneous effects of two endogenous variables in a system; more exactly, the concept of simultaneity does not apply here.

In a more recent article, Schmidt and Strauss [1976] further generalize their formulations to incorporate continuous dependent variables. When  $Y$  is a continuous variables, they formulate the so called mixed logit model:

$$\ln \left[ \frac{P(X=1|Y)}{P(X=0|Y)} \right] = Z\beta + \alpha Y \quad (2)$$
$$Y|X \sim N(Z\gamma + \delta X, \sigma^2)$$

where  $Y$  conditional on  $X$  is a normal regression model. Schmidt and Strauss [1976] apply their specification (2) to study the simultaneous effects of union membership and earnings. Olsen [1978] has shown that this model implicitly imposes a constraint on the two effects  $\alpha$  and  $\delta$  for the model (2); specifically,

$$\alpha\sigma^2 = \delta \quad (3)$$

With the constraint (3), Olsen points out that the two effects must both be either positive, negative or zero; consequently the relative size of the two effects is a function of whether  $\sigma^2$  is above or below unity and why the magnitude of the variance should play such a central role in determining the relative magnitude is far from clear. Olsen concludes that while the two effects appear to be distinct they cannot be separated, leaving unanswered the very difficult question of the extent of the two

directions of causation. The difficulty in understanding this model is due to the conditional specifications in (2). Contrary to model (1) there is no contingency table analysis formulation from which the mixed logit model can be derived. Recognizing the constraint (3) imposed on the mixed logit model, Warren and Strauss [1979], in their use of model (2) to study the simultaneous determination of unionization and right-to-work legislation, comment "The presence of this constraint makes the mixed logit model appropriate only when prior theoretical considerations suggest that  $\alpha$  and  $\delta$  have the same sign....Note that the imposition of this constraint amounts to a prior information that X and Y are jointly dependent." Are Warren and Strauss correct in their way to interpret the constraint in the mixed logit model? Can the model detect any possible simultaneous effect of X and Y at all? These questions as well as the ones raised by Olsen are the problems that we will consider.

## 2. Recursive Logit Model vs. Mixed Logit Model

In Maddala and Lee [1976], we have formulated a set of recursive models in the logit framework. One of the two equation models considered is

$$\begin{aligned} X^* &= Z\theta_1 + \epsilon_1 \\ Y &= Z\theta_2 + \delta X + \epsilon_2 \end{aligned} \tag{4}$$

where  $\epsilon_1$ ,  $\epsilon_2$  are independent disturbances,  $X^*$  is a latent variable which determines the observable dichotomous variable X as

$$\begin{aligned} X &= 1 \quad \text{iff } X^* > 0 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

If  $\epsilon_2$  is assumed to be  $N(0, \sigma^2)$  and  $\epsilon_1$   $\text{Sech}^2$  distribution, one has the following equivalent system

$$\ln\left[\frac{P(X=1)}{P(X=0)}\right] = Z\theta_1 \quad (4)'$$

$$Y = Z\theta_2 + \delta X + \varepsilon_2$$

This two equations model is fully recursive and the formulation is similar to the classical formulation of recursive equations models for continuous dependent variables. It is quite clear that since  $\varepsilon_1$  and  $\varepsilon_2$  are independent,  $X$  and  $Y$  are not jointly dependent. In fact, there is only one direction of causation from  $X$  to  $Y$  but not vice versa. Undoubtedly model (4) will be useful only when variable  $X$  does precede (in time or in some other sense) variable  $Y$ . Thus, one might have the impressions that the above fully recursive model and the mixed logit model are quite different in nature. However, this is not the case.

Let us assume that there are no prior constraints on  $\beta$  or  $\gamma$  in (2) as in the empirical applications of the mixed logit models. Similarly there are no prior restrictions on  $\theta_1, \theta_2$  in (4)'. The exogenous variable vector  $Z$  contains a constant term. Let  $A = \{(\theta_1, \theta_2, \delta, \sigma^2) | \sigma^2 > 0\}$  and  $B = \{(\beta, \gamma, \delta, \alpha, \sigma^2) | \alpha\sigma^2 = \delta, \sigma^2 > 0\}$  be the parameters spaces of models in (4)' and (2) respectively. The following proposition will be proved:

Proposition: The models in (2) and (4)' are the same; specifically, there exists a one to one correspondence  $h: A \rightarrow B$  such that  $a \in A, L_A(a) = L_B(h(a))$  where  $L_A$  and  $L_B$  are the likelihood functions of fully recursive logit models in (4)' and mixed logit model in (2) respectively.

### 3. Proof

To prove the proposition, we will use the following relations:

$$\begin{aligned} P(X=1|Y) &= P(X=1, Y) / P(Y) \\ &= P(X=1, Y) / (P(X=0, Y) + P(X=1, Y)) \\ &= (1 + P(X=0, Y) / P(X=1, Y))^{-1} \\ &= \left(1 + \frac{P(Y|X=0) P(X=0)}{P(Y|X=1) P(X=1)}\right)^{-1} \end{aligned} \quad (5)$$

$$\begin{aligned} \text{and } \frac{P(X=0)}{P(X=1)} &= \left( \frac{1}{P(X=1|Y)} - 1 \right) \frac{P(Y|X=1)}{P(Y|X=0)} \\ &= \frac{P(X=0|Y)}{P(X=1|Y)} \frac{P(Y|X=1)}{P(Y|X=0)} \end{aligned} \quad (6)$$

where  $P(X, Y)$  is the joint density of  $X, Y$  and  $P(X)$  is the marginal probability of  $X$ .

Suppose the equations in (4)' hold, we will prove that they can be written as equations in (2). Since  $\varepsilon_2 \sim N(0, \sigma^2)$  and is independent with  $\varepsilon_1$ , the conditional density of  $Y$  on  $X$  is

$$P(Y|X) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2} (Y - Z\theta_2 - \delta X)^2\right\}$$

and it implies

$$\frac{P(Y|X=0)}{P(Y|X=1)} = \exp\left\{-\frac{\delta}{\sigma^2} (Y - Z\theta_2) + \frac{\delta^2}{2\sigma^2}\right\}. \quad (7)$$

Hence from (5),

$$P(X=1|Y) = \frac{\exp\left(Z\theta_1 - \frac{\delta}{\sigma^2}Z\theta_2 - \frac{\delta^2}{2\sigma^2} + \frac{\delta}{\sigma^2}Y\right)}{\{1 + \exp\left(Z\theta_1 - \frac{\delta}{\sigma^2}Z\theta_2 - \frac{\delta^2}{2\sigma^2} + \frac{\delta}{\sigma^2}Y\right)\}}$$

Let  $\alpha = \frac{\delta}{\sigma^2}$ ,  $\beta = \theta_1 - \theta_2 \frac{\delta}{\sigma^2} - \frac{\delta^2}{2\sigma^2} e_1$  where  $e_1' = (1, 0, \dots, 0)$ . It follows

$$\ln \frac{P(X=1|Y)}{P(X=0|Y)} = Z\beta + \alpha Y.$$

Conversely, suppose the equations in (2) hold. From (6), (7), and (3),

$$\begin{aligned} \frac{P(X=0)}{P(X=1)} &= \exp(-Z\beta - \alpha Y) \cdot \exp\left(\alpha(Y - Z\gamma) - \frac{\delta}{2} \alpha\right) \\ &= \exp(-Z\beta - Z\alpha\gamma - \frac{\delta}{2}\alpha) \end{aligned}$$

and hence

$$\ln \left[ \frac{P(X=1)}{P(X=0)} \right] = Z\theta_1$$

where  $\theta_1 = \beta + \alpha\gamma + \frac{\delta\alpha}{2} e_1$ .

It is obvious that the mapping  $h$

$$h'(\theta_1, \theta_2, \delta, \sigma^2) = \left( \theta_1' - \theta_2' \frac{\delta}{\sigma^2} - \frac{\delta^2}{2\sigma^2} e_1', \theta_2', \delta, \sigma^2 \right)$$

defines a one to one correspondence between  $A$  and  $B$  and  $a \in A$ ,

$h(a) \in B$  define the same density function.

Q.E.D.

As an immediate corollary, we have

Corollary 1: The least squares single equation estimates of  $\gamma$ ,  $\delta$  and  $\sigma^2$  from the regression of  $Y$  on  $X$  in (2) are the full information maximum likelihood estimates.

#### 4. Conclusions

As for the cases we considered, the mixed logit model in (2) is only an equivalent representation of the fully recursive equations system in (4). There is no doubt that specification (2) may be quite misleading for modelling simultaneous determinations of qualitative and continuous response variables. The answers to Olsen's questions are clear. From the recursive model formulation in (4), it is clear that  $\delta = \sigma_{XY} / \sigma_X^2$ , where  $\sigma_X^2$  is the variance of  $X$  and  $\sigma_{XY}$  the covariance of  $X$  and  $Y$ , and  $\alpha$  in (2) is  $\sigma_{XY} / \sigma^2 \sigma_X^2$ . Thus  $\alpha$  and  $\delta$  are measures of correlation between  $X$  and  $Y$  expressed in different units and are not distinct simultaneous effects. Similar to the arguments in Brier [1979], we can conclude that the concept of simultaneity does not apply to the mixed logit model. It is not true that the imposition of constraint amounts to a prior information that  $X$  and  $Y$  are jointly dependent.

References

- Bishop, Y.M.M., S.E. Fienberg and P.W. Holland, 1975, Discrete Multivariate Analysis: Theory and Practice. Cambridge, MA, M.I.T. Press.
- Brier, S.S., 1979, The Utility of Systems of Simultaneous Logistic Response Equations, Sociological Methodology, Ch. 5, San Francisco, CA, Jossey-Bass, Inc., Publishers.
- Nerlove, M., and S.J. Press, 1973, Univariate and Multivariate Log-Linear and Logistic Models, Santa Monica: Rand R1306-NIH/EDA.
- Maddala, G.S. and L.F. Lee, 1976, Recursive Models with Qualitative Endogenous Variables, Annals of Economic and Social Measurement, 5, 525-545.
- Olsen, R.J., 1978, Comment on "The Effect of Unions on Earnings and Earnings on Unions: A Mixed Logit Approach", International Economic Review 19, 259-61.
- Schmidt, P. and R. P. Strauss, 1975, Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach, Econometrics 43, 745-755.
- Schmidt, P. and R. P. Strauss, 1976, The Effect of Unions on Earnings and Earnings on Unions: A Mixed Logit Approach, International Economic Review XVII, 204-212.
- Warren, R.S. and R.P. Strauss, 1979, A Mixed Logit Model of the Relationship Between Unionization and Right-to-Work Legislation, Journal of Political Economy, 87, 648-655.
- Theil, H., 1971, Principles of Econometrics, John Wiley & Sons, Inc., New York.