Journal of Econometrics 11 (1979) 233–246. © North-Holland Publishing Company

# A RANDOM COEFFICIENT PROBIT MODEL WITH AN APPLICATION TO A STUDY OF MIGRATION*

John S. AKIN and David K. GUILKEY

*University of North Carolina, Chapel Hill, NC 27514, USA*

Robin SICKLES

*University of Pennsylvania, Philadelphia, PA 19174, USA*

Received April 1979

This paper extends the ordered response polytomous probit model to a possibly more realistic framework in which the coefficients are allowed to be random. The new method is then used to analyze family migration behavior. It is seen that the new method allows us to make inferences that were not possible in the fixed coefficient model.

## 1. Introduction

In recent years, the widespread availability of large cross-sectional data sets has helped lead to the development of many new techniques for handling survey data. These new procedures often involve the problem of dealing with some type of limited dependent variable [see, for example, Schmidt and Strauss (1975 and 1976), Heckman (1974), and the NBER special issue (1976)]. The purpose of this paper is to extend one of the more commonly used limited dependent variable techniques, the probit model, to a possibly more realistic framework in which the coefficients of the independent variables are allowed to be random.

Allowing the parameters to be random in the linear regression model has long been argued in the econometrics literature, see, for example, Klein (1953), Nerlove (1965), Swamy (1970), and Zellner (1969). As will be discussed later, this specification seems quite reasonable in an analysis of family migration behavior which is an additional purpose of this paper. This particular application is of interest, not only because we are able to make inferences that are not possible in the standard fixed coefficient probit model,

*The authors would like to thank David Crawford, Peter Schmidt, Robert Strauss, P.A.V.B. Swamy, and Donald Waldman for helpful comments on an earlier draft of this paper. Constructive criticism on an earlier draft was provided by participants at the University of Pennsylvania Econometrics Workshop.

but also because it demonstrates the relative computational ease with which our model can be applied.

The plan of this paper is as follows. Section 2 presents the theoretical model and contrasts it to a related model by Hausman and Wise (1978). The study of migratory behavior using the new method is presented in section 3. Section 4 concludes.

## 2. The model specification

Consider the following model:

$$Y_t^* = X_t\beta_t + \mu_t, \qquad t=1,2,\dots,T, \tag{1}$$

where $X_t$ is a $1\times k$ vector of non-stochastic explanatory variables, $Y_t^*$ is an unobserved dependent variable, and the $\mu_t$ are unknown disturbances which are independent, identically distributed random variables with mean zero and variance $\sigma_\mu^2$. Finally, $\beta_t$ is a $k\times 1$ vector of unobserved random coefficients whose typical element can be written as

$$\beta_{ti} = \bar{\beta}_i + \varepsilon_{ti}, \qquad i=1,2,\dots,k, \tag{2}$$

where $\bar{\beta}_i$ is the unobserved mean of $\beta_{ti}$ and we assume that the $\varepsilon_{ti}$ come from a multivariate normal distribution with the following characteristics:

$$\begin{aligned} E(\varepsilon_{ti}) &= 0 \\ E(\varepsilon_{ti}\varepsilon_{sj}) &= \sigma_{ij} \quad \text{for} \quad t=s, \\ &= 0 \quad \text{for} \quad t\neq s. \end{aligned}$$

For the purpose of this study $\mu_t$ and the $\varepsilon_{ti}$ are assumed to be independent.

This specification for $\beta_t$ was first discussed by Rao (1965) and was also analyzed by Hildreth and Houck (1968) and Swamy and Mehta (1977). It allows individuals in a survey to have heterogeneous responses with respect to how a change in an explanatory variable affects the latent variable $Y_t^*$.

The reason that $Y_t^*$ is not observed is because it is frequently the case that in surveys we simply ask individuals to categorize their feelings on some subject. What we actually observe is some discrete realization of $Y_t^*$ which gives us an indication of the strength of these feelings. If we let $Y_t$ represent the observed discrete response, and the $A_i$ represent the thresholds through which $Y_t^*$ passes, then we may write the relationship between $Y_t^*$ and $Y_t$ as follows:

$$\begin{aligned} Y_t &= 1 \quad \text{if} \quad Y_t^* < A_1, \\ &= 2 \quad \text{if} \quad A_1 \leqq Y_t^* \leqq A_2, \\ &\;\vdots \\ &= n \quad \text{if} \quad A_{n-1} \leqq Y_t^*. \end{aligned} \tag{3}$$

If there are only two possible values of $Y_t$ $(n=2)$, then the relationship between $Y_t^*$ and $Y_t$ is that of the familiar dichotomous probit model. The polytomous probit model $(n>2)$ with fixed coefficients was developed by McKelvey and Zavonia (1975) and is referred to as an ordered response model by Amemiya (1975).

We can develop an estimator for the random coefficient probit model by rewriting (1) taking (2) into account,

$$Y_t^* = \bar{\beta}_1 X_{t1} + \ldots + \bar{\beta}_k X_{tk} + X_{t1}\varepsilon_{t1} + \ldots + X_{tk}\varepsilon_{tk} + \mu_t,$$

or

$$Y_t^* = X_t \bar{\beta} + \delta_t, \tag{4}$$

where

$$\bar{\beta} = [\bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_k]' \quad \text{and} \quad \delta_t = X_{t1}\varepsilon_{t1} + \ldots + X_{tk}\varepsilon_{tk} + \mu_t.$$

Given our assumptions on $\varepsilon_{ti}$ and $\mu_t$, the distribution of $\delta_t$ is normal with mean zero and variance

$$C_t^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} X_{ti} X_{tj} \sigma_{ij} + \sigma_\mu^2.$$

The covariance between $\delta_t$ and $\delta_s$ $(t \neq s)$ is zero. Thus we can express the probabilities associated with the $Y_t$'s as

$$\begin{aligned} P(Y_t = 1) &= \int_{-\infty}^{(A_1 - X_t\bar{\beta})/C_t} \phi(\lambda)\,\mathrm{d}\lambda, \\ P(Y_t = 2) &= \int_{(A_1 - X_t\bar{\beta})/C_t}^{(A_2 - X_t\bar{\beta})/C_t} \phi(\lambda)\,\mathrm{d}\lambda, \\ &\;\vdots \\ P(Y_t = n) &= 1 - P(Y_t = 1) - \ldots - P(Y_t = n-1), \end{aligned}$$

where $\phi(\lambda)$ is the density function of a standard normal random variable.

The likelihood function can then be written as

$$L=\prod_{t\in\theta_1} P(Y_t=1)\prod_{t\in\theta_2} P(Y_t=2)\dots\prod_{t\in\theta_n} P(Y_t=n), \tag{5}$$

where $t\in\theta_i$ if $Y_t^*$ is in the $i$th category as defined in (3).

Maximum likelihood estimators of the parameters can be obtained by maximizing (5) with respect to $\bar{\beta}$, $\sigma_\mu^2$, $\sigma_{ij}$ $(i,j=1,2,\dots,k)$, and $A_m$ $(m=1,2,\dots,n-1)$. However, it is not possible to identify all of the parameters. Even in the fixed coefficient probit model ($\sigma_{ij}=0$ for all $i$ and $j$) some normalizations must be imposed before estimation can proceed. The usual restrictions are to set $\sigma_\mu^2$ equal to one and $A_1$ equal to zero [see Nelson (1976, p. 496)]. However, since the variance of the constant term has a constant weight in the expression for $C_t$ as does $\sigma_\mu^2$ it would not be identifiable. We thus chose to restrict the constant term (and hence its variance) to equal zero instead of $A_1$.

The analytic first and second derivations of the likelihood function and other computational considerations are discussed in the appendix.

As was mentioned in section 1 the Conditional Probit model of Hausman and Wise (1978) and our Random Coefficient Polytomous Probit model bear some resemblances. With our model it is assumed that a particular choice occurs when the latent variable $Y_t^*$ falls within a certain interval. The measure given to the discrete observable decision (e.g. not to migrate, to migrate within a county, to migrate ouside a county) is a monotonic function of the latent variable. That is, the larger the value of this measure, the larger the value of the latent variable $Y_t^*$. One could think of the polytomous choices as hierarchical in that the result of decision $n$ is more important than that of decision, $n-1$. In contrast, Hausman and Wise specify an underlying continuous variable associated with *each* discrete choice and there is no hierarchical nature to the discrete choices.

Thus the essential difference is that the Hausman and Wise model deals with an unordered response, while our model deals with an *ordered* response. An explicit discussion of the differences between ordered and unordered response models can be found in Amemiya (1975) where he discusses the fixed coefficient counterparts of the two models.[1]

As will be demonstrated in the next section, an attractive feature of our model is the ease with which it can be implemented. This is in contrast to the Hausman and Wise method which, as they point out in their paper, is computationally very burdensome. Of course, the choice between the two models must be made on theoretical grounds. In choosing to model the

[1]It should be noted that when there are only two responses, in which case the order of the responses is irrelevant, our model is a special case of the Hausman and Wise model. However, when the number of responses exceeds two, the statistical similarity vanishes.

migration decision within our framework, we feel that we can categorize a decision to move outside of a county to be a more drastic decision than a decision to move within a county. Another possible application of our model might be a situation in which we had categorized information on inflationary expectations of individuals with respect to prices. In this case, it is clear that if the possible responses were that the inflation rate will go down, stay the same, or go up that a distinct ordering exists between these responses.

With a preliminary discussion of the statistical model behind us we turn to the study of migration in section 3.

## 3. A study of migration

The migration decision is of great interest to policy makers, and, as a result, the number of empirical papers by economists on the topic has been quite large.[2] The most notable characteristic of the literature is that with very few exceptions the empirical work is done using data aggregated to *at least* the SMSA or county level.[3] Migration, however, is clearly a microeconomic decision. Therefore, in this paper we use survey data since it provides the most appropriate level of aggregation for estimation of a model attempting to explain moving decisions.

An immediate problem that we encounter, caused by the micro-unit decision that we are attempting to analyze, is the specification of the dependent variable. A conceptually satisfying approach would be to specify all possible destinations as a function of the characteristics of the destinations. This approach is computationally unmanageable unless the number of destinations is small. A less than perfect proxy for the inclusion of all possible locations is to suggest that moves can reasonably be classified as non-existent, within the same local political jurisdiction, or beyond the local political jurisdiction.

Therefore, in the empirical work that follows, the discrete choices ($Y_t$) that we considered were: (a) do not move, (b) move within a county, and (c) move outside of a county.[4] The underlying latent variable ($Y_t^*$) is sentiment towards moving. We feel that the migration decision specified in this way fits into the framework of the ordered response model for a variety of reasons. First of all, any moving decision is clearly a more drastic action than a decision not to move at all. It is also reasonable to suggest that a move outside the county of present residence is more drastic than a decision to stay within the county, since the out of county move usually involves a job

[2]See Greenwood (1975) for a recent survey of the economics migration literature.

[3]The exceptions of which we are aware are Akin and Polachek (1975), Polachek and Horvath (1976) and Kaluzny (1975).

[4]We observe our units over two cross-sections of the Michigan Income Dynamics Panel and define moves by examining the change or lack of change in location over the one year period.

change. In addition, the change in local governmental area involves changes in tax and expenditure patterns, and forces the mover to acquire new knowledge to function in the public aspects of life.

The next practical problem that we are confronted with is how to capture the multiple decision-maker nature of the family migration model.[5] An obvious difficulty is that we would expect that for a family in which the head of the household makes decisions dictatorially to maximize his or her own expected utility, such factors as spouse's earnings at the origin will affect the decision process differently than in situations where the household is more or less a two vote adult democracy. Characteristics of children will also affect decisions differently according to the decision-making method in the household given the tastes and preferences of the family members. Unfortunately, in order to disentangle the effects of various decision methods on family migration we would need to somehow model the decision-making process of the family unit and build this specification into a larger migration model. Clearly, even if this could be accomplished, the data problems would be insurmountable since there does not exist a data set with sufficient information on individual family members.

The random coefficient probit model, however, allows us to model this multiple person decision with currently available family data. By allowing the coefficients of the explanatory variables in the cross-sectional survey to be random, we allow for differing effects of these variables across families due to different tastes and preferences. As a result, we can get information as to whether various spouse and child related independent variables, and the other variables as well, have heterogeneous effects on the migration decision across families.

### 3.1. Data

The data analyzed are from the University of Michigan Survey Research Center's Income Dynamics Panel. This survey contains family information beginning with the year 1968. We use the two cross-sections for 1970 and 1971. With this panel we are able to determine when and where families move, plus a vast amount of other socio-economic information concerning the family and its members. After deleting families for which information was incomplete we were left with a total sample of approximately 5400 families for the two-year period in our analysis. We randomly sampled these 5400 families to obtain a working sample of 1000 families. We defined as having moved those families whose residence changed between the interview dates in 1970 and 1971.

[5]Family rather than head of household oriented models have been used in only a few recent migration studies. See Kaluzny (1975), De Vanzo (1972), Akin and Polachek (1975), Polachek and Horvath (1976), Sandell (1975), and Mincer (1977).

### 3.2. Results

Table 1 presents the means and standard deviations of our variables. It is clear that all values are as would be expected. The split in the dependent variable was 719 did not move, 214 moved within the county, and 67 moved outside the county.

Table 1

Means and standard deviations for random sample of 1000 families.

| Variable | Mean | Std. dev. |
|---|---|---|
| 1. Earned income, head of household | 6332.68 | 6081.90 |
| 2. Earned income, spouse | 1109.68 | 2190.37 |
| 3. Employed, head (dummy) | 0.733 | 0.433 |
| 4. Married, head (dummy) | 0.616 | 0.487 |
| 5. No. of states, head | 2.035 | 1.269 |
| 6. Sex, head (male = 1, female = 2) | 1.311 | 0.463 |
| 7. No. of children below age 18 | 1.266 | 1.624 |
| 8. Years of experience, head | 17.845 | 14.503 |
| 9. Years of experience, spouse | 4.874 | 8.343 |
| 10. Years of education, head | 10.620 | 3.954 |
| 11. Race, head (white = 1, other = 0) | 0.599 | 0.490 |
| 12. Spouse in professional or managerial position or self-employed in market sector (dummy) | 0.440 | 0.497 |

Our linear empirical equation is based on the normal investment oriented economic theory of migration in which families tend to move in order to increase future earnings or non-monetary sources of utility.[6] With regard to the explanatory variables, it is hypothesized that adding members to a family increases the cost of moving, and, therefore, tends to reduce mobility. Having a working spouse, especially one who is in a professional or managerial occupation, will also tend to add to the cost of a move, but we expect the strength of this spouse related mobility reduction to vary greatly across families due to different decision structures. For the same reason we expect wide variation in the above mentioned effects of children on the mobility decision.

The earned income variables need additional explanation. It is frequently argued that we can proxy relative earnings to different destinations by comparing the actual earnings of our decision-makers in the two actual locations we observe. This approach suffers from the implication that actual earnings in a location to which a unit moves are the same as the expected earnings when the unit chose to move. For this reason we have chosen to use absolute levels of our explanatory variables at the origin. The reasoning is that low (high) absolute levels indicate low (high) levels relative to other

[6]See Akin and Polachek (1975) for a more detailed discussion of a similar model.

places in the cross-section and that absolute levels are, therefore, sufficient for explaining moves that are made in order to obtain improvements. We are also implicitly assuming that the worse the conditions within your present local governmental jurisdiction are, the more likely you are not only to move but also to move outside the jurisdiction.

The other independent variables have obvious interpretations.

Table 2 presents the results of both fixed and random coefficient model estimation. The results for the estimated coefficients are quite similar. We see that having lived in more states is significantly related to higher propensity to migrate. Being employed and having more education also appear to be positively related to migration, but the significance levels are low (especially for education in the fixed model). Significant negative effects are evidenced for head's income (significant only in the fixed model), female headship, number of children, spouse's experience (not highly significant in the fixed coefficient model), head's experience, and white race (low level of significance in both models).

Table 2 also presents estimates of the standard deviations of the coefficients of the random coefficient model [with reference to (2), what we are reporting are $\sqrt{\hat{\sigma}_{ij}}$ $(i=1,2,\ldots,k)$]. Note that in this case we set $\sigma_{ij}=0$ for $i \neq j$. In other words, we followed Hildreth and Houck and did not allow for covariation between the random coefficients. To do so would have added an additional sixty-six parameters [12(12-1)/2].

We see that income of wife, number of children, experience of head and education of head all have statistically significant variance in the coefficients across families. The set of significantly varying coefficients includes every important spouse and children characteristic except the professional and managerial occupation dummy. This clearly seems to be supportive of our contention that a more general model which allows for a differing structure of decision making across households represents an improvement over earlier models. The $\chi^2$ test statistic for the null hypothesis that all the variances are jointly zero is 47.722. The critical value for the $\chi^2_{0.005}$ (df=12) is 28.300, a result that is clearly in support of our specification of randomness in the coefficients.

An additional result which is quite interesting is that even though the fixed coefficient estimation suggests a zero effect for spouse's income, the random coefficient model indicates that there is significant variation in the effect of this independent variable across families at the 5% level of significance for a one-sided test of the null hypothesis of zero variation. In other words, even though 'on average' there is no effect, the probability of getting a positive *or* negative effect for spouse's income in some families is quite high.

This last result is clearly an example of an inference that was not possible in the fixed coefficient model. Based on the results of the fixed coefficient model, one would not consider spouse's income an important variable and it

Table 2

Parameter estimates for fixed and random coefficient probit models.

| Variable | Fixed | | Random | | | |
|---|---|---|---|---|---|---|
| | Coefficient | $t$-statistic | Coefficient means | $t$-statistic | Coefficient std. dev. | $t$-statistic |
| 1. Income, head | −0.0985 | −1.56 | −0.1291 | −1.01 | 0.0000 | 0.0000 |
| 2. Income, spouse | −0.0155 | −0.29 | −0.1066 | −0.57 | 0.9022 | 1.72 |
| 3. Employed | 0.0584 | 1.08 | 0.1121 | 1.04 | 0.0115 | 0.01 |
| 4. Married | −0.1004 | −0.85 | −0.0425 | −0.01 | 0.0323 | 0.03 |
| 5. States | 0.0745 | 1.78 | 0.1681 | 1.96 | 0.1726 | 0.46 |
| 6. Sex | −0.2419 | −4.23 | −0.6095 | −2.52 | 0.0115 | 0.01 |
| 7. Children | −0.2028 | −4.05 | −0.3576 | −1.90 | 0.4862 | 1.53 |
| 8. Experience, spouse | −0.0630 | −1.02 | 1.0592 | −2.35 | 1.3031 | 2.30 |
| 9. Esperience, head | −0.4886 | −8.79 | −1.1692 | −2.56 | 1.0669 | 2.09 |
| 10. Education | 0.0244 | 0.49 | 1.5901 | 1.35 | 0.9463 | 2.11 |
| 11. Race | −0.0658 | −1.41 | −0.1160 | −1.17 | 0.0046 | 0.01 |
| 12. Professional | −0.0984 | −0.94 | −0.0023 | 0.01 | 0.2356 | 0.18 |
| Thresholds | | | | | | |
| $A_1$ | 0.6700 | 14.45 | 1.7162 | 2.63 | | |
| $A_2$ | 1.0469 | 16.20 | 3.8682 | 2.69 | | |

might possibly be eliminated from the model. The random coefficient model clearly suggests that it is important, even though we cannot say anything systematic about its influence on the migration decision.

As a final comparison, we present the effect on the probability of moving of changes in the categorized explanatory variables. Table 3 presents the six different configurations we chose for the five dummy variables. Table 4 then shows the probability of (a) not moving, (b) moving within the county, (c) moving outside the county, for each of the six configurations when the continuous variables are held fixed at their averages both for the fixed and random coefficient model.

Table 3

Configurations of dummy variables used to generate probabilities.[a]

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1. Employed | No | Yes | Yes | Yes | Yes | Yes |
| 2. Married | No | No | Yes | Yes | Yes | Yes |
| 3. White | No | No | No | Yes | Yes | Yes |
| 4. Spouse, professional | No | No | No | No | Yes | Yes |
| 5. Female | No | No | No | No | No | Yes |

[a]All continuous variables set at sample average.

Table 4

Example of changes in probabilities between fixed and random coefficient models.

| Configuration | | Fixed coefficient | Random coefficient |
|---|---|---|---|
| 1 | a | 0.6336 | 0.7488 |
| | b | 0.2839 | 0.2398 |
| | c | 0.0825 | 0.0114 |
| 2 | a | 0.7151 | 0.7615 |
| | b | 0.2318 | 0.2283 |
| | c | 0.0531 | 0.0102 |
| 3 | a | 0.6610 | 0.7159 |
| | b | 0.2671 | 0.2697 |
| | c | 0.0719 | 0.0143 |
| 4 | a | 0.7311 | 0.7183 |
| | b | 0.2207 | 0.2681 |
| | c | 0.0481 | 0.0136 |
| 5 | a | 0.7311 | 0.7183 |
| | b | 0.2207 | 0.2681 |
| | c | 0.0481 | 0.0136 |
| 6 | a | 0.8541 | 0.7954 |
| | b | 0.1281 | 0.1943 |
| | c | 0.0178 | 0.0103 |

## 4. Conclusion

This paper has presented a generalization of the polytomous probit model that allows individuals to have randomly different responses to a change in an explanatory variable. A clear advantage over the standard probit model was that in our analysis of family migratory behavior we were able to model this multiple person decision given current data limitations. As is discussed in the appendix, this generalization was achieved with only a moderate increase in computational burden.

## Appendix

The purpose of this appendix is to discuss in greater detail computational considerations associated with the polytomous probit model with random coefficients. First of all, since many computer maximization routines require first and second derivatives of the likelihood function, we present them for the model of section 3 where $\sigma_{ij}=0$ for $i\neq j$. To keep the notation simple, we write $\sigma_{ii}$ as $\sigma_i^2$ for this case.

We can proceed by taking the logarithm of (5),

$$\mathscr{L}=\log L=\sum_{i=1}^{n}\sum_{t_i}\log\omega_{it},$$

where

$$\omega_{it}=\Phi(Z_{it})-\Phi(Z_{i-1,t}),$$

$$\Phi(Z_{it})=\int_{-\infty}^{Z_{it}}\phi(\lambda)\,\mathrm{d}\lambda,$$

$$\Phi(Z_{0t})=0,$$

$$\Phi(Z_{nt})=1-\Phi(Z_{n-1,t}),$$

$$Z_{it}=(A_i-X_t\beta)/C_t,$$

$$C_t=(\sigma_1^2X_{t1}^2+\ldots+\sigma_k^2X_{tk}^2+1)^{\frac{1}{2}},$$

$$t_i=t\in\theta_i.$$

The first derivatives can now be written as

$$\frac{\partial\mathscr{L}}{\partial A_j}=\sum_{t_j}\frac{\omega_{jt}\phi(Z_{jt})}{C_t}-\sum_{t_{j+1}}\frac{\omega_{j+1,t}\phi(Z_{jt})}{C_t},\qquad j=1,2,\ldots,n-1,$$

$$\frac{\partial \mathscr{L}}{\partial \beta_m}=\sum_{i=1}^{n}\sum_{t_i}\frac{\omega_{it}X_{tm}}{C_t}[-\phi(Z_{it})+\phi(Z_{i-1,t})], \qquad m=1,2,\ldots,k,$$

$$\frac{\partial \mathscr{L}}{\partial \sigma_p}=\sum_{i=1}^{n}\sum_{t_i}\frac{\omega_{it}X_{tp}^2\sigma_p}{C_t^2}[-Z_{it}\phi(Z_{it})+Z_{i-1,t}\phi(Z_{i-1,t})], \qquad p=1,2,\ldots,k.$$

Derivation of the second derivatives is more tedious but yields the following:

$$\frac{\partial^2 \mathscr{L}}{\partial A_j^2}=\sum_{t_j}\frac{\phi(Z_{jt})}{C_t^2}[\phi(Z_{jt})-\omega_{jt}Z_{jt}] + \sum_{t_{j+1}}\frac{\phi(Z_{jt})}{C_t^2}[\phi(Z_{jt})-\omega_{j+1,t}Z_{jt}],$$

$$\frac{\partial^2 \mathscr{L}}{\partial A_j \partial A_{j+1}}=-\sum_{t_{j+1}}\frac{\phi(Z_{j+1,t})\phi(Z_{jt})}{C_t^2},$$

$$\frac{\partial^2 \mathscr{L}}{\partial A_j \partial A_{j-1}}=-\sum_{t_j}\frac{\phi(Z_{j-1,t})\phi(Z_{jt})}{C_t^2},$$

$$\frac{\partial^2 \mathscr{L}}{\partial A_j \partial A_{j-1}}=0 \quad \text{for all } j, \quad p\neq j-1,\, j,\, j+1,$$

$$\frac{\partial^2 \mathscr{L}}{\partial A_j \partial \beta_m}=\sum_{t_j}\frac{\phi(Z_{jt})X_{tk}}{C_t^2}[\omega_{jt}Z_{jt}-\phi(Z_{jt})+\phi(Z_{j-1,t})] + \sum_{t_{j+1}}\frac{\phi(Z_{jt})X_{tk}}{C_t^2}[\omega_{j+1,t}Z_{jt}+\phi(Z_{j+1,t})-\phi(Z_{j-1,t})],$$

$$\frac{\partial^2 \mathscr{L}}{\partial A_j \partial \sigma_p}=\sum_{t_j}\frac{\phi(Z_{jt})X_{tp}^2\sigma_p}{C_t^3}[\omega_{jt}Z_{jt}^2-\phi(Z_{jt})Z_{jt}+\phi(Z_{j-1,t})Z_{t-1,t}] - \sum_{t_{j+1}}\frac{\phi(Z_{jt})X_{tp}^2\sigma_p}{C_t^3}[\omega_{j+1,t}Z_{jt}^2-\phi(Z_{j+1,t})Z_{j+1,t}+\phi(Z_{jt})Z_{jt}],$$

$$\frac{\partial^2 \mathscr{L}}{\partial \beta_m \partial \beta_p}=\sum_{i=1}^{n}\sum_{t_i}\frac{X_{tp}X_{tm}}{C_t^2}[-Z_{it}\phi(Z_{it})\omega_{it} + Z_{i-1,t}\phi(Z_{i-1,t}\omega_{it}+(-\phi(Z_{it})+\phi(Z_{i-1,t}))^2],$$

$$\frac{\partial^2 \mathscr{L}}{\partial \beta_m \partial \sigma_p}=\sum_{i=1}^{n}\sum_{t_i}\frac{X_{tm}X_{tp}^2\sigma_p}{C_t^3}\{\phi(Z_{it})\omega_{it}(1-Z_{it}^2) - \phi(Z_{i-1,t})\omega_{it}(1-Z_{i-1,t}^2) - [\phi(Z_{it})-\phi(Z_{i-1,t})][Z_{it}\phi(Z_{it})-Z_{i-1,t}\phi(Z_{i-1,t})]\},$$

$$\frac{\partial^2 \mathscr{L}}{\partial\sigma_p \partial\sigma_q} = \sum_{i=1}^{n} \sum_{t_i} g_{it} f_t \{ g_{it} - 3\omega_{it} + \omega_{it}[-Z_{it}^3 \phi(Z_{it}) + Z_{i-1,t}^3 \phi(Z_{i-1,t})]/g_{it} \},$$

where

$$p \neq q,$$

$$g_{it} = -Z_{it}\phi(Z_{it}) + Z_{i-1,t}\phi(Z_{i-1,t}),$$

$$f_t = \frac{X_{tp}^2 \sigma_p X_{tq}^2 \sigma_q}{C_t^4},$$

and

$$\frac{\partial^2 \mathscr{L}}{\partial\sigma_p^2} = \sum_{i=1}^{n} \sum_{t_i} \frac{g_{it} X_{tp}^4 \sigma_p^2}{C_t^4} \{ \omega_{it} C_t^2 / \sigma_p^2 X_{tp}^2 + g_{it} - 3\omega_{it} + \omega_{it}[-Z_{it}^3 \phi(Z_{it}) - Z_{i-1,t}^3 \phi(Z_{i-1,t})]/g_{it} \}.$$

The actual computer maximization routine was the DFP algorithm which used the analytic first derivatives presented above. The $t$-statistics that we report are obtained from the numerical second derivatives of the likelihood function. Since our sample size is large, we can be reasonably confident that the test statistics are normal with mean zero and variance one.

Although differences in computer software and hardware make it difficult, if not impossible, to make general statements about the feasibility of the application of our specification to large scale models with many observations, our limited experience is encouraging. The model of section 3 which involved a search over 26 parameters with 1000 observations converged after 117 iterations. The convergence criterion used was that the change in the likelihood function was less than $1.0 \times 10^{-15}$. This is probably a much stricter convergence criterion than is really needed to obtain reasonable results.

## References

Akin, J. and S. Polachek, 1975, A family model of the decision to migrate, Paper presented at the Western Economics Society Meetings (San Diego, CA).

Albright, R., R. Lerman and C. Manski, forthcoming, Report on the development of an estimation program for the multinomial probit model, Econometrica.

Amemiya, T., 1975, Qualitative response models, Annals of Economic and Social Measurement, Summer, 363–372.

Bowles, S., 1970, Migration as investment: Empirical tests of the human investment approach to geographic mobility, Review of Economics and Statistics, Nov., 356–362.

Da Vanzo, J., 1972, An analytical framework for studying the potential effects of an income maintenance program on U.S. intergenerational migration, Rand Report R-1081-EAD, Dec.

Greenwood, M., 1975, Research on internal migration in the United States: A survey, Journal of Economic Literature 13, June, 307–433.

Hausman, J. and D. Wise, 1978, A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences, Econometrica, March, 403–426.

Heckman, J., 1974, Shadow prices, market wages, and labor supply, Econometrica, July, 679–694.

Hildreth, C. and J. Houck, 1968, Some estimators for a linear model with random coefficients, Journal of the American Statistical Association, 584–595.

Kaluzny, R., 1975, Determinants of household migration: A comparative study by race and poverty level, Review of Economics and Statistics 52, Aug., 269–274.

Klein, L., 1953, A textbook of econometrics (Row, Peterson and Co., Evanston, IL).

Mincer, J., 1976, Family migration decisions, Mimeo., Jan.

NBER special issue on discrete, qualitative and limited dependent variables, 1976, Annals of Economic and Social Measurement, Fall.

Nelson, F., 1976, On a general computer algorithm for the analysis of models with limited dependent variables, Annals of Economic and Social Measurement, Fall, 493–510.

Nerlove, M., 1965, Estimation and identification of Cobb–Douglas production functions (Rand McNally, Chicago, IL; North-Holland, Amsterdam).

Polachek, S. and F. Horvath, 1976, A life cycle approach to migration: Analysis of the perspicacious peregrinator, Mimeo., Aug.

Rao, C.R., 1965, The theory of least squares when the parameters are stochastic, Biometrika, 447–458.

Sandell, S., 1975, The economics of family migration, in: Dual careers: A longitudinal analysis of the labor market experience of women 4, 141–160.

Schmidt, P. and R. Strauss, 1975, Estimation of models with jointly dependent qualitative variables: A simultaneous equations approach, Econometrica, July, 745–756.

Schmidt, P. and R. Strauss, 1976, The effect of unions on earnings and earnings on unions: A mixed logit approach, International Economic Review, Feb., 204–212.

Sjaastad, L., 1962, The costs and returns of human migration, Journal of Political Economy, Suppl., Oct., 80–93.

Swamy, P., 1970, Efficient inference in a random coefficient regression model, Econometrica, March, 311–323.

Swamy, P. and J. Mehta, 1977, Estimation of linear models with time and cross-sectionally varying coefficients, Journal of the American Statistical Association, Dec., 890–898.

Zellner, A., 1969, On the aggregation problem: A new approach to a troublesome problem, in: K.A. Fox et al., eds., Economic models, estimation and risk programming: Essays in honor of Gerhard Tintner (Springer-Verlag, New York).