ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

ESTIMATION OF MODELS WITH JOINTLY DEPENDENT QUALITATIVE VARIABLES: A SIMULTANEOUS LOGIT APPROACH

Peter Schmidt and Robert P. Strauss

VOL. 43, No. 4—July 1975

Econometrica, Vol. 43, No. 4 (July, 1975)

ESTIMATION OF MODELS WITH JOINTLY DEPENDENT QUALITATIVE VARIABLES: A SIMULTANEOUS LOGIT APPROACH

BY PETER SCHMIDT AND ROBERT P. STRAUSS[1]

This paper considers the estimation of a simultaneous equations model in which the dependent variables are qualitative. The model is a simultaneous version of the multivariate logit model, and can be estimated by maximum likelihood. An example is presented, dealing with the prediction of occupation and industry of employment of a worker based on certain demographic variables.

1. INTRODUCTION

A WIDE VARIETY of behavior may be more properly interpreted as qualitative than continuous in nature. The probit and logit models are well-known specifications which deal with the former situation when the variable to be explained takes on a finite number of discrete values. Usually there are only two values for the dependent variable representing the occurrence or non-occurrence of an event.

A limitation of such models is that they allow for only one-way analysis of the effects of right-hand side variables on the dependent variable of interest. There are a variety of circumstances when such an interpretation may be untenable; that is, there are circumstances when the discontinuous analogue to simultaneous equation bias may be reasonably thought to exist. We seek in this paper to construct a simultaneous logit model which meets this difficulty and apply it to the problem of predicting occupation and industry of employment. While the estimator developed is applied to an economic problem, it should have wide application in other areas of research, as jointly dependent qualitative behavior would seem quite common. Sections 2 and 3 develop the model, Section 4 contains the application, and Section 5 concludes.

2. A SIMULTANEOUS LOGIT MODEL

Suppose that we have two variables, X and Y, each of which takes on values 0 and 1. Suppose that the logit specification is appropriate, and that the probability of X (or Y) taking on the value of 1 is influenced by a certain number of exogenous variables, *and* by the value of Y (or X). Then a natural specification is

$$\log\left[\frac{P(X=1/Y)}{P(X=0/Y)}\right]_t = R_t\beta + \alpha Y_t,$$

(1)

$$\log\left[\frac{P(Y=1/X)}{P(Y=0/X)}\right]_t = Q_t\gamma + \delta X_t.$$

[1] We wish to thank the Data and Program Library Service, University of Wisconsin, for an efficient version of the 1967 Survey of Economic Opportunity data file. Financial support from the United States Department of Labor is gratefully acknowledged. After the first draft of this paper was written, we became aware of the similar work of Nerlove and Press [3].

Here R_t is a $1 \times K_X$ vector of values of the exogenous variables that affect X, β is a $K_X \times 1$ vector of unknown parameters, α is a parameter, Q_t is a $1 \times K_Y$ vector of values of the exogenous variables that affect Y, γ is a $K_Y \times 1$ vector of parameters, and δ is a parameter. The observation index t runs from 1 to T.

From the specification above, it necessarily follows that $\alpha = \delta$. To see this, note that:

$$\begin{aligned}\alpha &= \log\left[\frac{P(X=1/Y=1)}{P(X=0/Y=1)}\right] - \log\left[\frac{P(X=1/Y=0)}{P(X=0/Y=0)}\right] \\ &= \log\left[\frac{P(X=1, Y=1)}{P(X=0, Y=1)}\right] - \log\left[\frac{P(X=1, Y=0)}{P(X=0, Y=0)}\right] \\ &= \log\left[\frac{P(X=1, Y=1)P(X=0, Y=0)}{P(X=0, Y=1)P(X=1, Y=0)}\right].\end{aligned}$$

On the other hand,

$$\begin{aligned}\delta &= \log\left[\frac{P(Y=1/X=1)}{P(Y=0/X=1)}\right] - \log\left[\frac{P(Y=1/X=0)}{P(Y=0/X=0)}\right] \\ &= \log\left[\frac{P(Y=1, X=1)P(Y=0, X=0)}{P(Y=0, X=1)P(Y=1, X=0)}\right] = \alpha.\end{aligned}$$

This "symmetry" condition enables us to write the model as:

$$\log\left[\frac{P(X=1/Y)}{P(X=0/Y)}\right]_t = R_t\beta + \alpha Y_t,$$

(2)

$$\log\left[\frac{P(Y=1/X)}{P(Y=0/X)}\right]_t = Q_t\gamma + \alpha X_t.$$

Note that it implies that if the occurrence of $Y = 1$ makes $X = 1$ more likely (or less likely), then the converse is also necessarily true.

To estimate the model, maximum likelihood is, at least conceptually, the most straightforward alternative. To develop this, we need to calculate the probabilities of the various possible occurrences. Straightforward (though tedious) algebra reveals the following:

(3)

$$\begin{aligned}P(X_t = 0, Y_t = 0) &= 1/\Delta_t, \\ P(X_t = 0, Y_t = 1) &= \exp(Q_t\gamma)/\Delta_t, \\ P(X_t = 1, Y_t = 0) &= \exp(R_t\beta)/\Delta_t, \text{ and} \\ P(X_t = 1, Y_t = 1) &= \exp(R_t\beta + Q_t\gamma + \alpha)/\Delta_t,\end{aligned}$$

where

(4) $$\Delta_t = 1 + \exp(Q_t\gamma) + \exp(R_t\beta) + \exp(R_t\beta + Q_t\gamma + \alpha).$$

If we define

$$(5) \qquad \Theta_{ij} = \{t | X_t = i, Y_t = j\} \qquad (i, j = 0, 1),$$

then we have the likelihood function

$$(6) \qquad L = \prod_{i=0}^{1} \prod_{j=0}^{1} \prod_{t \in \Theta_{ij}} P(X_t = i, Y_t = j).$$

Maximum likelihood estimates of β, γ, and α can be obtained by the maximization of (6) using a numerical maximization program.

3. THE MORE GENERAL CASE

While the most common case is that in which X and Y each have only two possible values, it is sometimes the case that a variable to be predicted may have three or more possible values. Let us suppose, therefore, that X has m possible values, say $1, 2, \ldots, m$, and Y has n possible values, say $1, 2, \ldots, n$. (This is a slight change in notation from the previous section, where the two values of X and Y were 0 and 1 rather than 1 or 2, but involves no matter of substance.)

In such a case the analogue of equation (2) above is

$$(7) \qquad \begin{aligned} &\log\left[\frac{P(X = i/Y)}{P(X = 1/Y)}\right]_t = R_t\beta_i + \sum_{k=2}^{n} \alpha_{ik} Y_{kt}^* && (i = 2, 3, \ldots, m); \\ &\log\left[\frac{P(Y = i/X)}{P(Y = 1/X)}\right]_t = Q_t\gamma_i + \sum_{k=2}^{m} \alpha_{ki} X_{kt}^* && (i = 2, 3, \ldots, n). \end{aligned}$$

Here the Y_k^* and X_k^* are dummy variables defined by

$$(8) \qquad \begin{aligned} &Y_{kt}^* = \begin{cases} 1 \text{ if } Y_t = k, \\ 0 \text{ otherwise} \end{cases} && (k = 2, \ldots, n); \\ &X_{kt}^* = \begin{cases} 1 \text{ if } X_t = k, \\ 0 \text{ otherwise} \end{cases} && (k = 2, \ldots, m). \end{aligned}$$

The β's and γ's are, as before, vectors of coefficients corresponding to the exogenous variables in the Z_t and Q_t. The α's are (scalar) coefficients of the dummy variables. Note that the number of α's is $(m-1)(n-1)$ and that the analogue of the "symmetry" condition of the previous section has already been imposed.

To state the likelihood function in this case, we again need to calculate the joint probabilities. They turn out to be

$$(9) \qquad \begin{aligned} &P(X_t = 1, Y_t = 1) = 1/\Delta_t; \\ &P(X_t = 1, Y_t = j) = \exp(Q_t\gamma_j)/\Delta_t && (j = 2, \ldots, n); \\ &P(X_t = i, Y_t = 1) = \exp(R_t\beta_i)/\Delta_t && (i = 2, \ldots, m); \\ &P(X_t = i, Y_t = j) = \exp(R_t\beta_i + Q_t\gamma_j + \alpha_{ij})/\Delta_t && (i = 2, \ldots, m; j = 2, \ldots, n). \end{aligned}$$

In the above expressions,

$$(10)\qquad \Delta_t = 1 + \sum_{i=2}^{m} \exp(R_t\beta_i) + \sum_{j=2}^{n} \exp(Q_t\gamma_j)$$

$$+ \sum_{i=2}^{m} \sum_{j=2}^{n} \exp(R_t\beta_i + Q_t\gamma_j + \alpha_{ij}).$$

This leads to the likelihood function

$$(11)\qquad L = \prod_{i=1}^{m} \prod_{j=1}^{n} \prod_{t\in\Theta_{ij}} P(X_t = i, Y_t = j)$$

where

$$(12)\qquad \Theta_{ij} = \{t | X_t = i, Y_t = j\}.$$

Again, resort to numerical maximization is necessary.

It is also possible to consider cases in which there are more than two dependent variables of interest. The approach in such cases would be exactly the same; however, the likelihood function becomes increasingly complex as the number of variables increases. The interested reader may obtain a copy of the likelihood function from the authors.

4. AN APPLICATION OF THE SIMULTANEOUS LOGIT MODEL

In this section we apply the simultaneous logit model to prediction of occupation of employment when occupation and industry are considered to be jointly determined. A previous paper [5] analyzed occupational attainment in a single equation context; however, there is reason to believe that a simultaneous approach may be more appropriate. First, occupation and industry are often dependently defined. For example, one rarely finds physicians outside the health industry. Second, independent of the historical aspects of particular industry and occupation definitions, particular labor markets or occupations may be so specialized that they provide services to only one industry.

We analyze the occupation and industry of employment of a sample of 936 persons surveyed in 1967. The parent, representative portion of the 1967 Survey of Economic Opportunity was stratified to include only those over 14 years of age, those who were full-time employees in 1967, and those who had non-zero earnings in that year. The 936 observations were then randomly drawn from this stratified sample.

Four exogenous variables were used: extent of training as measured by years of schooling, labor market experience measured as age minus years of schooling minus five; race with black = 0 and white = 1; and sex, with female = 0 and male = 1.

To ensure that the analysis and computations remain tractable, we have limited the number of occupations to three: "professional," "skilled," and "menial," and the number of industries to two: "service and government" and "production and distribution." Table I displays the underlying two-digit Census Bureau occupation and industry groupings for this three-way grouping of occupation and two-way grouping of industry.

We estimate, then, functions of the form

$$
\begin{aligned}
\log\left[\frac{P_A}{P_B}\right]_t &= \beta_{11} + \beta_{12}\text{Education}_t + \beta_{13}\text{Experience}_t + \beta_{14}\text{Race}_t \\
&\quad + \beta_{15}\text{Sex}_t + \beta_{16}\text{“Skilled”}_t + \beta_{17}\text{“Professional”}_t, \\
\log\left[\frac{P_2}{P_1}\right]_t &= \beta_{21} + \beta_{22}\text{Education}_t + \beta_{23}\text{Experience}_t + \beta_{24}\text{Race}_t \\
&\quad + \beta_{25}\text{Sex}_t + \beta_{16}\text{“Service and Government”}_t, \\
\log\left[\frac{P_3}{P_1}\right]_t &= \beta_{31} + \beta_{32}\text{Education}_t + \beta_{33}\text{Experience}_t + \beta_{34}\text{Race}_t \\
&\quad + \beta_{35}\text{Sex}_t + \beta_{17}\text{“Service and Government”}_t,
\end{aligned}
\tag{13}
$$

TABLE I

UNDERLYING TWO-DIGIT OCCUPATIONS AND INDUSTRIES

Constructed Grouping	Component Two-Digit Titles
"Professional"	Professional, technical, and kindred workers
	Managers, officials, and proprietors except farm
"Skilled"	Clerical and kindred workers
	Sales workers
	Craftsmen, foremen, and kindred workers
	Farm managers and farmers
"Menial"	Operatives and kindred workers
	Private household workers
	Service workers except private household
	Farm laborers and foremen
	Laborers, except farm and mine
"Service and Government"	Business and repair services
	Personal services
	Entertainment and recreation services
	Professional and related services
	Public administration
"Production and Distribution"	Agriculture, forestry, and fisheries
	Mining
	Construction
	Manufacturing—durables
	Manufacturing—non-durables
	Transportation
	Communications
	Utilities and sanitary services
	Wholesale trade
	Retail trade
	Finance, insurance, and real estate

where A is "service and government" industry, B is "production and distribution" industry, 1 is the "menial" occupation, 2 is the "skilled" occupation, and 3 is the "professional" occupation. The variables in quotation marks are the appropriate dummy variables. We may also derive the equation:

$$\text{(14)} \qquad \log \left[\frac{P_3}{P_2}\right]_t = \beta_{41} + \beta_{42}\text{Education}_t + \beta_{43}\text{Experience}_t + \beta_{44}\text{Race}_t + \beta_{45}\text{Sex}_t + \beta_{46}\text{“Service and Government”}_t,$$

where $\beta_{4i} = \beta_{3i} - \beta_{2i}$, $i = 1, \ldots, 5$, and where $\beta_{46} = \beta_{17} - \beta_{16}$.

The estimated coefficients and their asymptotic standard errors are given in Table II. "t ratios" formed as the ratio of the estimated coefficient to the asymptotic standard error (obtained from the information matrix), are asymptotically distributed as $N(0, 1)$ under the null hypothesis that the associated coefficient is zero. With almost 1,000 observations, we can feel reasonably confident in using the $N(0, 1)$ distribution for such tests.

Looking at Table II, let us consider first the effects of education. The positive and highly significant coefficients indicate that, other things held constant, more education increases the probability of being in the "professional" occupation, relative to the "skilled" or "menial" occupations. It also makes being in the "skilled" occupation relative to the "menial" occupation more likely. Finally, more education also makes being in the "service and government" industry relative to the "production and distribution" industry more likely.

The effects of labor market experience are less clear since several of the estimated coefficients are statistically insignificant (at reasonable confidence levels). Those that are significant say that more experience increases one's chances of being in the "professional" occupation.

The effects of race are fairly clear-cut and interesting. Other things held constant, to be white increases the chance of being in the "professional" occupation relative to the "skilled" and "menial" occupation and also increases the chance of being in the "skilled" occupation relative to the "menial." In other words, among blacks and whites of equal education and experience, and of the same sex, the whites are apt to get the "better" jobs. On the other hand, being white makes it less likely to be in the "service and government" industry relative to the "production and distribution" industry.

The coefficients of the sex dummy indicate that women tend to be bunched into the "skilled" occupation. To be male makes it more likely to be in the "professional" occupation relative to the "skilled" or "menial" occupation, and less likely to be in the "skilled" occupation relative to the "menial" occupation. Being male also makes it less likely to be in the "service and government" industry relative to the "production and distribution" industry.

Finally, let us consider the dummy variables corresponding to occupation and industry. Looking at the industry equation, we see that to be in the "skilled" occupation (rather than the "menial" occupation) makes being in the "service and government" industry relative to the "production and distribution" industry

TABLE II

Coefficients and Standard Errors[a]

Dependent Variable	Constant	Education	Experience	Race[b]	Sex[c]	"Skilled"	"Prof."	"Serv. and Govt."
$\log (P_A/P_B)$	−.6597	.1547	.00400	−.8737	−1.359	−.5222	.7338	—
	(.62)	(.034)	(.0067)	(.26)	(.17)	(.20)	(.24)	
$\log (P_2/P_1)$	−3.317	.2719	.00073	.7818	−.2594	—	—	−.5222
	(.55)	(.034)	(.0066)	(.27)	(.17)			(.20)
$\log (P_3/P_1)$	−11.420	.6941	.01991	1.411	.8636	—	—	.7338
	(.89)	(.053)	(.0087)	(.44)	(.25)			(.24)
$\log (P_3/P_2)$	−8.103	.4222	.01918	.6292	1.1230	—	—	1.2560
	(.83)	(.045)	(.0081)	(.44)	(.23)			(.23)

[a] A is "Service and Government," B is "Production and Distribution," 1 is "Menial," 2 is "Skilled," and 3 is "Professional."
[b] White is 1, Black is 0.
[c] Male is 1, Female is 0.

TABLE III

PROBABILITIES OF BEING IN EACH OCCUPATION-INDUSTRY COMBINATION, GIVEN AVERAGE EDUCATION AND EXPERIENCE[a]

Race–Sex Combination	Menial, Serv. and Govt.	Menial, Prod. and Dist.	Skilled, Serv. and Govt.	Skilled, Prod. and Dist.	Professional, Serv. and Govt.	Professional, Prod. and Dist.
Black female	.363	.227	.176	.186	.037	.011
Black male	.170	.415	.064	.262	.041	.048
White female	.144	.216	.153	.386	.060	.043
White male	.051	.300	.042	.414	.051	.141

[a] Average education is 11.400, average age is 24.147.

less likely, while being in the "professional" occupation makes being in the "service and government" industry more likely. Conversely, being in the "service and government" industry (rather than in the "production and distribution" industry) makes being in the "skilled" occupation less likely and being in the "professional" occupation more likely, relative to the "menial" occupation. Incidentally, it might be pointed out that the statistical significance of these dummy variables indicates that industry and occupation do, indeed, affect each other, so that a simultaneous specification is useful in this case.

In order to illustrate these results in a different way, we have also evaluated the probabilities of being in each occupation-industry combination, using equation (9). This was done for each sex-race combination, at the sample means of education and experience. These are given in Table III.

Given average educational attainment and experience, we find rather striking differences between blacks and whites in their odds of entering certain industry-occupation combinations. For example, a black male has three times the chance of being in the "menial" occupation—"services and government" industry than a white male, while the white male has a three-fold better chance than the black male of being in the "professional" occupation—"production and distribution" industry. Similar disparities occur for a black-white comparison of females. For example, white females are four times as likely as black females to be in the "professional" occupation—"production and distribution" industry.

A final check on the specification used here was to check whether the results varied by region or not. The above model was rerun using random samples of size 1,000 from each of the four regions (Northeast, North Central, South, and West) in the Survey of Economic Opportunity data. The coefficients and standard errors are given in Table IV, and the implied probabilities in Table V. Without discussing these in any detail, we will comment that the inter-regional differences are less substantial than might have been expected.

5. CONCLUSION

This paper developed a simultaneous equations version of the logit model. It is designed to handle the joint prediction of two qualitative variables, where each

TABLE II

Coefficients and Standard Errors[a]

Dependent Variable	Constant	Education	Experience	Race[b]	Sex[c]	"Skilled"	"Prof."	"Serv. and Govt."
$\log (P_A/P_B)$	−.6597	.1547	.00400	−.8737	−1.359	−.5222	.7338	—
	(.62)	(.034)	(.0067)	(.26)	(.17)	(.20)	(.24)	
$\log (P_2/P_1)$	−3.317	.2719	.00073	.7818	−.2594	—	—	−.5222
	(.55)	(.034)	(.0066)	(.27)	(.17)			(.20)
$\log (P_3/P_1)$	−11.420	.6941	.01991	1.411	.8636	—	—	.7338
	(.89)	(.053)	(.0087)	(.44)	(.25)			(.24)
$\log (P_3/P_2)$	−8.103	.4222	.01918	.6292	1.1230	—	—	1.2560
	(.83)	(.045)	(.0081)	(.44)	(.23)			(.23)

[a] A is "Service and Government," B is "Production and Distribution," 1 is "Menial," 2 is "Skilled," and 3 is "Professional."
[b] White is 1, Black is 0.
[c] Male is 1, Female is 0.

TABLE III
PROBABILITIES OF BEING IN EACH OCCUPATION-INDUSTRY COMBINATION, GIVEN AVERAGE EDUCATION AND EXPERIENCE[a]

Race–Sex Combination	Menial, Serv. and Govt.	Menial, Prod. and Dist.	Skilled, Serv. and Govt.	Skilled, Prod. and Dist.	Professional, Serv. and Govt.	Professional, Prod. and Dist.
Black female	.363	.227	.176	.186	.037	.011
Black male	.170	.415	.064	.262	.041	.048
White female	.144	.216	.153	.386	.060	.043
White male	.051	.300	.042	.414	.051	.141

[a] Average education is 11.400, average age is 24.147.

less likely, while being in the "professional" occupation makes being in the "service and government" industry more likely. Conversely, being in the "service and government" industry (rather than in the "production and distribution" industry) makes being in the "skilled" occupation less likely and being in the "professional" occupation more likely, relative to the "menial" occupation. Incidentally, it might be pointed out that the statistical significance of these dummy variables indicates that industry and occupation do, indeed, affect each other, so that a simultaneous specification is useful in this case.

In order to illustrate these results in a different way, we have also evaluated the probabilities of being in each occupation-industry combination, using equation (9). This was done for each sex-race combination, at the sample means of education and experience. These are given in Table III.

Given average educational attainment and experience, we find rather striking differences between blacks and whites in their odds of entering certain industry-occupation combinations. For example, a black male has three times the chance of being in the "menial" occupation—"services and government" industry than a white male, while the white male has a three-fold better chance than the black male of being in the "professional" occupation—"production and distribution" industry. Similar disparities occur for a black-white comparison of females. For example, white females are four times as likely as black females to be in the "professional" occupation—"production and distribution" industry.

A final check on the specification used here was to check whether the results varied by region or not. The above model was rerun using random samples of size 1,000 from each of the four regions (Northeast, North Central, South, and West) in the Survey of Economic Opportunity data. The coefficients and standard errors are given in Table IV, and the implied probabilities in Table V. Without discussing these in any detail, we will comment that the inter-regional differences are less substantial than might have been expected.

5. CONCLUSION

This paper developed a simultaneous equations version of the logit model. It is designed to handle the joint prediction of two qualitative variables, where each

TABLE IV

Coefficients and Standard Errors by Region

Region	Dependent Variable	Constant	Education	Experience	Race	Sex	"Skilled"	"Prof."	"Serv. and Govt."
Northeast	$\log(P_A/P_B)$	−1.028	.1587	.00644	−.7910	−1.052	−.0910	.9980	
		(.70)	(.033)	(.0060)	(.26)	(.16)	(.20)	(.23)	
	$\log(P_2/P_1)$	−3.006	.2482	.00214	.4690	−1.367	—	—	−.0910
		(.52)	(.035)	(.0059)	(.26)	(.16)			(.20)
	$\log(P_3/P_1)$	−10.280	.6319	.01595	1.566	.3258	—	—	.9980
		(.81)	(.047)	(.0078)	(.47)	(.22)			(.23)
	$\log(P_3/P_2)$	−7.274	.3835	.01381	1.097	.4625	—	—	1.089
		(.74)	(.040)	(.0072)	(.46)	(.20)			(.20)
North Central	$\log(P_A/P_B)$	−.6818	.1252	.01570	−.7969	−1.208	−.4479	.8711	—
		(.63)	(.035)	(.0064)	(.28)	(.17)	(.20)	(.23)	
	$\log(P_2/P_1)$	−3.363	.2502	.00892	.6346	−.0738	—	—	−.4479
		(.56)	(.038)	(.0062)	(.28)	(.16)			(.20)
	$\log(P_3/P_1)$	−11.210	.6725	.02673	1.300	.8625	—	—	.8711
		(.88)	(.051)	(.0822)	(.51)	(.23)			(.23)
	$\log(P_3/P_2)$	−7.847	.4223	.01781	.6654	.9363	—	—	1.319
		(.83)	(.044)	(.0077)	(.52)	(.22)			(.22)
South	$\log(P_A/P_B)$	−.6009	.1644	.02768	−1.383	−1.613	−.500	1.076	—
		(.56)	(.031)	(.0067)	(.22)	(.17)	(.22)	(.25)	
	$\log(P_2/P_1)$	−4.202	.2304	.01231	1.5871	.3810	—	—	−.5000
		(.48)	(.033)	(.0066)	(.24)	(.17)			(.22)
	$\log(P_3/P_1)$	−10.990	.6173	.03696	1.802	.7857	—	—	1.076
		(.76)	(.047)	(.0087)	(.36)	(.23)			(.25)
	$\log(P_3/P_2)$	−6.788	.3869	.02465	.2149	.4047	—	—	1.576
		(.72)	(.041)	(.0081)	(.38)	(.22)			(.23)
West	$\log(P_A/P_B)$	−.9529	.1770	.00711	−.9450	−1.298	−.2369	.4838	—
		(.63)	(.032)	(.0059)	(.34)	(.16)	(.19)	(.22)	
	$\log(P_2/P_1)$	−2.634	.1663	−.00413	1.7663	−.5468	—	—	−.2369
		(.63)	(.035)	(.0066)	(.40)	(.18)			(.19)
	$\log(P_3/P_1)$	−9.769	.5890	.01989	1.837	.1138	—	—	.4838
		(.84)	(.046)	(.0074)	(.48)	(.22)			(.22)
	$\log(P_3/P_2)$	−7.135	.4227	.02402	.07071	.6606	—	—	.7207
		(.79)	(.039)	(.0069)	(.54)	(.19)			(.19)

TABLE V

PROBABILITIES OF BEING IN EACH OCCUPATION-INDUSTRY COMBINATION, BY REGION, GIVEN AVERAGE EDUCATION AND EXPERIENCE

Region	Race-Sex Combination	Menial, Serv. and Govt.	Menial, Prod. and Dist.	Skilled, Serv. and Govt.	Skilled, Prod. and Dist.	Professional, Serv. and Govt.	Professional, Prod. and Dist.
Northeast[a]	Black female	.244	.252	.204	.232	.050	.019
	Black male	.131	.386	.096	.310	.037	.040
	White female	.102	.233	.137	.343	.100	.084
	White male	.046	.301	.054	.386	.063	.150
North Central[b]	Black female	.332	.255	.160	.193	.045	.015
	Black male	.154	.396	.069	.278	.050	.054
	White female	.145	.247	.132	.352	.073	.052
	White male	.052	.298	.044	.395	.062	.149
South[c]	Black female	.546	.250	.081	.061	.053	.008
	Black male	.212	.488	.046	.174	.045	.035
	White female	.150	.273	.108	.326	.088	.055
	White male	.031	.282	.032	.492	.040	.124
West[d]	Black female	.444	.182	.174	.091	.087	.022
	Black male	.279	.419	.063	.121	.062	.057
	White female	.106	.112	.243	.325	.131	.085
	White male	.058	.222	.076	.374	.080	.190

[a] Average education is 11.552, average experience is 24.730.
[b] Average education is 11.458, average experience is 24.245.
[c] Average education is 10.807, average experience is 24.571.
[d] Average education is 12.219, average experience is 23.265.

can take on an arbitrary number of values. The model was then applied to the prediction of occupation and industry, and reasonably plausible results were obtained.

University of North Carolina.

Manuscript received February, 1974

REFERENCES

[1] GOODMAN, L.: "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables," *American Sociological Review*, 37 (1972), 28–46.

[2] GRIZZLE, J., D. F. STARMER, AND G. KOCH: "Analysis of Categorical Data by Linear Models," *Biometrika*, 25 (1969), 489–504.

[3] NERLOVE, M., AND S. J. PRESS: *Univariate and Multivariate Log-Linear and Logistic Models*. Santa Monica, Calif.: RAND Corporation Report R-1306, 1973.

[4] PRESS, S. J.: *Applied Multivariate Analysis*. New York: Holt, Rinehardt, and Winston, 1972.

[5] SCHMIDT, P., AND R. P. STRAUSS: "The Prediction of Occupation Using Multiple Logit Models," *International Economic Review*, forthcoming.

[6] THEIL, H.: "A Multivariate Extension of the Logit Model," *International Economic Review*, 10 (1969), 251–259.