Determinants of the Female Occupational Distribution: A Log-Linear Probability Analysis

Evelyn L. Lehrer; Houston Stokes

The Review of Economics and Statistics, Vol. 67, No. 3 (Aug., 1985), 395-404.

Stable URL:
http://links.jstor.org/sici?sici=0034-6535%28198508%2967%3A3%3C395%3ADOTFOD%3E2.0.CO%3B2-O

The Review of Economics and Statistics is currently published by The MIT Press.

DETERMINANTS OF THE FEMALE OCCUPATIONAL DISTRIBUTION: A LOG-LINEAR PROBABILITY ANALYSIS

Evelyn L. Lehrer and Houston Stokes*

Abstract—The present study examines the determinants of two aspects of occupation which have been found to have important influences on female wages: the skill level associated with the occupation and the sex composition of the occupation (typically female versus male or integrated). Using a log-linear probability technique and data from the National Longitudinal Surveys, Young Women Cohort, hypotheses drawn from the economic and sociological literature are tested.

Introduction

RESEARCH on the relative positions of males and females in the labor market has uncovered many similarities: (a) men and women in the labor force have the same median years of schooling completed (U.S. Department of Labor, 1977); (b) they are distributed in occupations requiring, on average, the same amount of general skills (England et al., 1982); (c) they attain the same mean level of occupational status (as measured by the Duncan index and other indicators), and do so through similar processes (Treiman and Terrell, 1975; McClendon, 1976; Featherman and Hauser, 1976).

Despite these similarities, the gap between the earnings of male and female workers is large. The median female full-time worker earns about 58% of the amount her male counterpart earns. An important explanation is the difference by sex in the amount and continuity of experience (Mincer and Polachek, 1974). Male-female differentials in occupational distribution also play a role. The heavy concentration of women in "female occupations" (e.g., secretaries, nurses) is well documented in the literature. These occupations share a common characteristic, namely, low pay (see, e.g., Smith, 1979). Recent studies (Jusenius, 1977; England et al., 1982; Zalokar, 1984) have emphasized another aspect of sex differences in occupational distribution: women are more likely than men to be in occupations which require or provide low levels of "specific vocational training" (*SVP*). As defined in U.S. Department of Labor (1981), *SVP* is "the amount of time required to learn the techniques, acquire the information and develop the facility needed for average performance in a specific job-worker situation. The training may be acquired in a school, work, military, institutional, or a vocational environment." Jusenius (1977) has shown that each of these dimensions of occupation—its sex composition and the *SVP* level associated with it—has an important impact on wages. The objective of the present paper is to examine empirically the *determinants* of these two aspects of occupation. The hypotheses outlined below are tested.

Hypothesis 1: Anticipating strong labor force attachment has a positive impact on the likelihood of having a high SVP occupation. This hypothesis is based on the notion that the optimal amount of investment in human capital is a positive function of the length of time that capital will be utilized (Becker, 1964). Those who expect long working lives will find it optimal to accumulate more job related human capital, because benefits will be reaped over a longer period of time.

Hypothesis 2: Anticipating strong labor force attachment has a positive impact on the likelihood of having a non-traditional occupation. From an economic perspective, the major difference between males and females in the labor market resides in their different patterns of labor supply. If the observed concentration of women in certain occupations results from rational economic calculations on their part, these occupations must have some characteristic that makes them attractive to individuals who expect interrupted careers and a relatively low duration of time in the labor force. Various possibilities have been offered in the literature. Polachek (1981) has suggested low atrophy rates: women choose the occupations they do because these occupations involve skills that do not depreciate much during periods of non-participation. Zellner (1975) has noted that female occupations may have higher starting wages (though flatter earnings profiles). According to Shaw (1983), female occupations may offer more possibilities of

Received for publication December 2, 1983. Revision accepted for publication October 29, 1984.

*University of Illinois at Chicago.

We have benefited from helpful comments on earlier drafts by Gilbert Bassett, Barry Chiswick, Carmel Chiswick, and two anonymous referees of this *Review*.

part-time employment. If the general statement that female occupations are attractive to individuals with low labor force attachment is true (be it à la Polachek, Zellner or Shaw), then, other factors held constant, those women who expect higher levels of labor supply should be less likely to be observed in the traditionally female occupations.

Hypothesis 3: Mother's occupation is a major predictor of the sex label of daughter's occupation. This hypothesis is based on the psychological and sociological literature: Since early childhood, girls learn that some activities and some types of behavior are appropriate for females, others are not; as adults they choose occupations consonant with these ideas (see, e.g., England, 1982a). This sex role socialization hypothesis suggests that, *ceteris paribus*, young women whose mothers did not work in the labor market, or who were in female dominated occupations, should be most likely to have traditional values and enter these occupations themselves.[1]

These hypotheses are examined using data from the National Longitudinal Surveys, Young Women Cohort (NLS), and a log-linear probability technique. Log-linear probability models have been used by social scientists to analyze a wide range of issues involving qualitative dependent variables (Bishop et al., 1975, Goodman, 1978). While initially only categorical exogenous variables could be considered, a procedure whereby jointly endogenous qualitative variables can be analyzed as functions of continuous or categorical explanatory factors was suggested by Nerlove and Press (1973, 1976). This technique, and a similar one developed independently by Schmidt and Strauss (1975a), have been employed in several studies on economic and demographic phenomena (e.g., Schmidt and Strauss, 1975b; Polachek, 1981; Lehrer, 1983; Koo and Janowitz, 1983; Lehrer and Nerlove, 1984). The Nerlove-Press methodology has recently been applied to the empirical study of jointly dependent variables with more than two states (Lehrer and Kawasaki, 1983; Lehrer, 1984). The present paper uses this technique.

[1] Note that this hypothesis is related to the sex label of the occupation, not to the skill dimension. Suppose that traditional versus non-traditional orientation indeed has an important influence on occupational selection. Those women brought up with traditional values might, nevertheless, choose occupations involving substantial amounts of training and a relatively high skill level (e.g., elementary school teachers, librarians).

The data, variables and analytical framework are described in section A; section B presents the econometric procedure; section C reports the empirical results, and section D closes the paper with a discussion of the implications of our findings.

A. Data, Variables and Analytical Framework

The data employed in this study were collected by the Center for Human Resource Research of the Ohio State University. The initial sample consisted of 5,533 women aged 14 to 24 as of January 1, 1968. The women were interviewed in 1968, and in several subsequent years; the surveys are still in progress. Our sample is restricted to women who were in the labor force, not enrolled in school, during at least one of the interview years, 1975, 1977 or 1978. Following Polachek (1981), England (1982b) and others, attention is focused on white females only. This removes the possibility that race interactions may affect the coefficients of interest. After excluding also cases with invalid codes for relevant variables, the sample size is 1,929.[2]

Defining occupation as that which the respondent had in the last year she was observed in the labor force, not enrolled in school, the following model is estimated:

$$\text{Occupation} = f(\text{expected labor force attachment, mother's occupation, control variables}).$$

[2] We initially considered including in our sample all white respondents who were in the labor force, not enrolled in school, at some point in the 1968–1978 period. We decided against this possibility because our proxies for expected labor force attachment, namely, plans for age 35 and number of children expected are, as indicated later in the text, measured as of 1968 and 1973, respectively. Measuring occupation in the same year, or in a prior year, introduces causality problems. Considering a relatively long time period (1968–1978 versus 1975–1978) has the additional disadvantage that changes in socioeconomic conditions over time may contaminate the effects of interest.

For two reasons, the model presented in this paper was re-estimated with a smaller sample ($n = 1145$) obtained by eliminating respondents younger than 18 years of age as of 1968. (a) We reasoned that women this young might not yet have serious plans regarding future labor supply and family size. (b) If women of all ages are considered, the age range as of 1978 is 24–34. By excluding those who were still enrolled in school, a censorship bias could be introduced. Limiting the sample to those aged 18 and above in 1968 narrows the age range in 1978 to 28–34 and thus lessens the danger of any censorship bias. The results based on the reduced sample (which are obtainable from the authors on request) are similar to those reported here.

Occupation is indexed by two variables, *LABEL* and *SKILL*. *LABEL* equals 2 if the respondent's occupation is "typically female," 1 otherwise. If men and women were equal in all respects relevant to the labor market, the expected fraction of women in any given occupation would be approximately equal to the fraction in the labor force. Using the proportion of the labor force which was female in 1970 as the reference point (38.1%), and allowing 10 percentage points for random deviations, a female occupation is defined as one in which women's share of employment exceeds 48.1%. Similar definitions have been used in earlier studies (Jusenius, 1977; Beller, 1981).[3] *SKILL* is coded 1, 2 or 3 according to whether the *SVP* requirement of the occupation is high ($SVP = 5$ or higher), intermediate ($SVP =$ greater than 1, less than 5) or low ($SVP = 1$ or lower). The *SVP* values are measured in years.[4]

LABEL and *SKILL* are treated as jointly endogenous qualitative variables which depend on the right-hand variables specified above. A discussion of each follows.

Expected Labor Force Attachment

In the 1968 interview, women were asked what they planned to be doing at age 35. Two dummy variables are defined on the basis of this question. The first equals 1 if the respondent answered she planned to be working, 0 otherwise. The second takes the value 1 if the respondent indicated that she did not know. The answer "married, keeping house, raising family" is the omitted category. We use a measure of expected labor force attachment rather than retroactively examining labor force experience, because as Blakemore and Low (1984) emphasize, human capital theory predicts that *anticipated* labor supply influences occupational selection.

The 1973 interview included a number of questions on fertility.[5] Respondents were asked how many children, if any, they had had, and how many they expected to have in the future. The variable "expected family size" is constructed on the basis of this information. This is another, probably less precise, proxy for anticipated labor supply.

Based on hypotheses 1 and 2, respectively, strong attachment to the labor market is expected to be positively associated with the likelihood of observing the respondent in a high *SVP* occupation and in a non-traditional occupation.

Mother's Occupation

The NLS data document whether the respondent's mother was employed during the years when the respondent was a teenager, and, if so, what occupation she had. Two variables are constructed. The first takes the value 1 if the mother worked in an integrated or male occupation; the second equals 1 if she worked in a female-dominated occupation. Mothers who worked rarely or not at all constitute the base. The sex role socialization hypothesis predicts that, *ceteris paribus*, having a mother who worked in a male or integrated occupation should have a positive impact on the likelihood of choosing a non-traditional occupation.

Control Variables[6]

Many studies in the economics literature identify father's education as a proxy for household permanent income, and mother's education as a proxy for the quality of time inputs into the production of child services (e.g., Chiswick, 1973; Lehrer, 1984). It has also been suggested that "parental educational attainment influences family values and expectations concerning children's education. And in addition, parents serve, at least in part, as educational role models for their children" (Treiman and Terrell, 1975, p. 177). Whichever interpretation one takes, both father's and mother's education are expected to affect *SVP* positively. Because it has a positive impact on the benefits derived from human capital, ability (as measured

[3] The model reported in this paper was re-estimated allowing 20 percentage points for random deviations, i.e., using 58.1% as the dividing line between occupations that are typically female and those which are not. Essentially the same results were obtained.

[4] In the NLS data, Young Women Cohort, occupations are coded in three-digit 1960 Census categories. Procedures similar to those described by Scoville (1966) were employed to assign an *SVP* value to each of these categories.

[5] The 1968 survey also documents expected fertility. Due to some coding errors, however, this information is not reliable.

[6] Variables unrelated to the central hypotheses are labelled "controls." These are variables which, we believe, cannot explain the bulk of male–female differences in occupational choice, but which should be held constant in the analysis in order to obtain clean tests of the hypotheses.

by results from IQ tests) is expected to influence *SVP* in the same direction.[7]

Respondents with no siblings are expected to attain higher levels of training because, all other factors the same, their parents are likely to invest more time and financial resources in them. Further, as Chiswick et al. (1974, p. 226) have noted:

> Although on the average, men and women have the same family background it has been alleged that families do more to promote labor market oriented human capital investments in their sons than in their daughters. If so, differential intra-family treatment could explain a different distribution of occupation and wage rates for women than for men.

This observation also suggests that the absence of siblings should have a positive influence on the amount of training acquired by the respondent.

The final control variable is the respondent's age, as of the year she was last observed in the labor force, not enrolled in school. To the extent that female intra-generational mobility takes the form of movement to higher skilled and male occupations, age may be expected to have a positive influence on the odds that *SVP* will have a high value and that *LABEL* will equal 1.

Table 1, which shows a cross-tabulation of the endogenous variables, confirms the well-known fact that the vast majority of women are in typically female occupations (70.9%), and in the low (51.1%) or intermediate-skill (36.0%) categories. Women in non-traditional occupations are more likely than their counterparts in female occupations to be in the high-skill category (18.9% as opposed to 10.5%). Looking at the figures from another angle, among those in the high-skill category, 57.4% are in female occupations; this contrasts with 73.3% for those in the low-skill group. A pronounced difference is also observed when comparing the high and intermediate-skill categories, but not when comparing the intermediate and low-skill groups. Whether these relationships still hold up when other variables are held constant is examined in section C.

TABLE 1.—CROSS-TABULATION OF *SKILL* AND *LABEL* (ROW PERCENTAGES IN PARENTHESES, COLUMN PERCENTAGES IN BRACKETS)

SKILL \ *LABEL*	Male or Integrated	Female	
High	106 (42.6%) [18.9%]	143 (57.4%) [10.5%]	249 [12.9%]
Intermediate	193 (27.8%) [34.3%]	501 (72.2%) [36.6%]	694 [36.0%]
Low	263 (26.7%) [46.8%]	723 (73.3%) [52.9%]	986 [51.1%]
	562 (29.1%)	1,367 (70.9%)	1,929 100.0%

B. The Econometric Technique

Suppose A is a trichotomous and B a dichotomous endogenous variable. We define indices for A and B: $i_A = 1, 2, 3$; $i_B = 1, 2$. If we assume, for simplicity, that A and B are jointly dependent on only one continuous variable, x, the log-linear probability model may be written as follows:

$$\begin{aligned}
\log P[i_A = 1, i_B = 1] &= \mu + c_1 + a_1 x + d_1 + b_1 x + \beta_{AB}(1,1)\\
\log P[i_A = 2, i_B = 1] &= \mu + c_2 + a_2 x + d_1 + b_1 x + \beta_{AB}(2,1)\\
\log P[i_A = 3, i_B = 1] &= \mu + c_3 + a_3 x + d_1 + b_1 x + \beta_{AB}(3,1)\\
\log P[i_A = 1, i_B = 2] &= \mu + c_1 + a_1 x + d_2 + b_2 x + \beta_{AB}(1,2)\\
\log P[i_A = 2, i_B = 2] &= \mu + c_2 + a_2 x + d_2 + b_2 x + \beta_{AB}(2,2)\\
\log P[i_A = 3, i_B = 2] &= \mu + c_3 + a_3 x + d_2 + b_2 x + \beta_{AB}(3,2).
\end{aligned} \tag{1}$$

Identification restrictions are necessary. Nerlove and Press (1973, 1976) impose the constraint that if any one of the effects is summed over all the values of one of the indices on which it depends, the sum should equal 0. Because it facilitates the interpretation of the results (see appendix), in the analysis below we make the following alternative assumptions:

$$a_3 = 0 \quad b_2 = 0 \quad c_3 = 0 \quad d_2 = 0$$
$$\beta_{AB}(3,2) = 0 \quad \beta_{AB}(2,2) = 0$$
$$\beta_{AB}(1,2) = 0 \quad \beta_{AB}(3,1) = 0.$$

Analogously to the familiar ANOVA models, the log probabilities are decomposed into main

[7] Where an IQ score is missing ($n = 492$), the mean IQ score in the population is imputed. IQ is thus measured with error, and, in principle, this should bias the estimated coefficients downward. Other coefficients are also biased in directions which depend partly on the nature of the correlation between IQ and the variable in question.

effects ($c_j + a_j x$, $j = 1, 2, 3$, and $d_k + b_k x$, $k = 1, 2$), interaction effects (β_{AB}) and the grand mean, μ. μ is defined so as to ensure that $\Sigma_{i_A}\Sigma_{i_B}P[i_A, i_B] = 1$. Algebraic manipulation of (1) yields the following interpretations for the parameters:

$$\begin{aligned}
a_1 &= \frac{\partial}{\partial x}\log\frac{P[i_A = 1, i_B = 2]}{P[i_A = 3, i_B = 2]} \\
&= \frac{\partial}{\partial x}\log\frac{P[i_A = 1, i_B = 1]}{P[i_A = 3, i_B = 1]} \\
a_2 &= \frac{\partial}{\partial x}\log\frac{P[i_A = 2, i_B = 2]}{P[i_A = 3, i_B = 2]} \\
&= \frac{\partial}{\partial x}\log\frac{P[i_A = 2, i_B = 1]}{P[i_A = 3, i_B = 1]} \\
b_1 &= \frac{\partial}{\partial x}\log\frac{P[i_A = 3, i_B = 1]}{P[i_A = 3, i_B = 2]} \\
&= \frac{\partial}{\partial x}\log\frac{P[i_A = 2, i_B = 1]}{P[i_A = 2, i_B = 2]} \\
&= \frac{\partial}{\partial x}\log\frac{P[i_A = 1, i_B = 1]}{P[i_A = 1, i_B = 2]} \\
\beta_{AB}&(1,1) \\
&= \log\frac{P[i_A = 1, i_B = 1]/P[i_A = 1, i_B = 2]}{P[i_A = 3, i_B = 1]/P[i_A = 3, i_B = 2]} \\
\beta_{AB}&(2,1) \\
&= \log\frac{P[i_A = 2, i_B = 1]/P[i_A = 2, i_B = 2]}{P[i_A = 3, i_B = 1]/P[i_A = 3, i_B = 2]}.
\end{aligned} \tag{2}$$

Thus, a_1 represents the change in the log odds that A will take the value 1 rather than 3, variable B held constant, associated with a unit change in x, i.e., a_1 indicates the *direct* impact of x on the likelihood that A will equal 1 as opposed to 3. If another exogenous variable, y, were introduced, a_1 would indicate the change in the log odds, B and y held constant. A similar interpretation applies to a_2 and b_1. The interaction term $\beta_{AB}(1,1)$ measures the relationship between levels 1 and 3 of variable A on the one hand and levels 1 and 2 of variable B on the other, net of exogenous effects. For instance, the odds that i_B will equal 1 given $i_A = 1$ are $P[i_A = 1, i_B = 1]/P[i_A = 1, i_B = 2]$. If A and B are not associated, the odds that i_B will equal 1 given $i_A = 3$ should be the same. Thus, if $\beta_{AB}(1,1) = 0$, this means that the value of B is not associated with whether variable A takes the value 1 or 3. Departures from 0 indicate lack of independence. $\beta_{AB}(2,1)$ may be interpreted in an analogous manner.

The likelihood function is

$$L = \prod_{n=1}^{N}\prod_{j=1}^{3}\prod_{k=1}^{2} p_{njk}^{m_{njk}}, \tag{3}$$

where N is the number of observations, m_{njk} equals 1 if for the n^{th} observation $i_A = j$ and $i_B = k$; m_{njk} equals 0 otherwise. p_{njk} is the probability that for the n^{th} observation $i_A = j$ and $i_B = k$. These probabilities may be obtained from exponentiation of (1). Maximizing the log of L (using a numerical maximization program) yields maximum-likelihood estimates of a_1, a_2, b_1, $\beta_{AB}(1,1)$ and $\beta_{AB}(2,1)$.

Two other parameters of interest are

$$\begin{aligned}
&\frac{\partial}{\partial x}\log\frac{P[i_A = 1, i_B = 2]}{P[i_A = 2, i_B = 2]} \\
&\quad = \frac{\partial}{\partial x}\log\frac{P[i_A = 1, i_B = 1]}{P[i_A = 2, i_B = 1]},
\end{aligned}$$

and

$$\log\frac{P[i_A = 1, i_B = 1]/P[i_A = 1, i_B = 2]}{P[i_A = 2, i_B = 1]/P[i_A = 2, i_B = 2]}.$$

These are automatically determined from the equations in (1). Since their t-ratios, however, cannot be readily inferred, the estimates of these parameters and the corresponding t-scores are also reported in the empirical section.

Once parameter estimates are obtained, predicted probabilities may be computed. For instance, one can compute the probability that A equals 1 setting x first at m and then at n, where m and n are constants. The two computed probabilities will differ because (a) x affects A, (b) x affects B, and (c) A and B interact. Comparison of these probabilities thus provides a quantitative measure of the *total* impact of x on A.[8]

[8] It should be emphasized that the parameter estimates and the predicted probabilities provide different types of information. For example, the first equation in system (2) shows that a_1 reflects the impact of x on the odds that A will equal 1 rather than 3, *variable B held constant*; thus, a_1 represents a *direct* effect. On the other hand, if we were to compute the predicted probability that A will equal 1 setting x first at, say, $x = m$ and then at $x = n$, comparison of these two probabilities would provide information regarding the *total* impact of x on A, i.e., the direct impact of x on A *plus* the indirect impact of x on A through its influence on B. The predicted probabilities that A will equal 1 are computed by adding the estimated values of $P[i_A = 1, i_B = 1]$ and $P[i_A = 1, i_B = 2]$ (see system (1)), setting x first at $x = m$ and then at $x = n$. This distinction between direct and total effects has been used by Lehrer and Nerlove (1984) and Lehrer (1984).

TABLE 2.—MAXIMUM-LIKELIHOOD ESTIMATES
(t-RATIOS IN PARENTHESES)

	High Skill versus Low Skill[a]	Intermediate Skill versus Low Skill[b]	High Skill versus Intermediate Skill[c]	Male or Integrated Occupation versus Female Occupation[d]
Plans for age 35 = working	0.3328	−0.0593	0.3921	0.1213
	(2.0)	(−0.5)	(2.3)	(1.1)
Plans for age 35 = don't know	−0.5442	−0.3841	−0.1600	0.0504
	(−1.5)	(−1.9)	(−0.4)	(0.2)
Expected family size	−0.0111	−0.0832	0.0721	−0.0682
	(−0.2)	(−1.8)	(1.0)	(−1.5)
Mother worked/male occupation	0.3007	−0.1127	0.4134	−0.0553
	(1.3)	(−0.7)	(1.8)	−(0.3)
Mother worked/female occupation	0.0038	−0.1564	0.1602	−0.0452
	(0.02)	(−1.3)	(0.9)	(−0.4)
Control Variables				
Father's education	0.0348	0.0098	0.0250	0.0045
	(1.3)	(0.6)	(1.0)	(0.3)
Mother's education	0.2177	0.0807	0.1370	0.0085
	(5.9)	(3.5)	(3.7)	(0.4)
IQ score	0.0526	0.0239	0.0287	0.0112
	(7.7)	(5.0)	(4.2)	(2.4)
Only child	0.3777	0.0539	0.3238	0.1239
	(1.3)	(0.2)	(1.1)	(0.6)
Age	0.0208	0.0020	0.0187	0.0383
	(0.8)	(0.1)	(0.8)	(2.3)
Constant	−10.80	−3.605	−7.192	−3.273
	(−10.1)	(−5.1)	(−6.7)	(−4.7)
Interaction Terms				
[High Skill versus Low Skill] with Male or Integrated Occupation[e]				0.5619
				(3.6)
[Intermediate Skill versus Low Skill] with Male or Integrated Occupation[f]				−0.0045
				(−0.04)
[High Skill versus Intermediate Skill] with Male or Integrated Occupation[g]				0.5664
				(3.6)

Note: $N = 1{,}929$.
[a] The estimates in this column (rows 1–10) correspond to a_1.
[b] The estimates in this column (rows 1–10) correspond to a_2.
[c] The estimates in this column (rows 1–10) correspond to

$$\frac{\partial}{\partial x}\log\frac{P[i_A=1,i_B=2]}{P[i_A=2,i_B=2]} = \frac{\partial}{\partial x}\log\frac{P[i_A=1,i_B=1]}{P[i_A=2,i_B=1]} = a_1 - a_2.$$

[d] The estimates in this column (rows 1–10) correspond to b_1.
[e] The estimate in this row represents $\beta_{AB}(1,1)$.
[f] The estimate in this row represents $\beta_{AB}(2,1)$.
[g] The estimate in this row represents

$$\log\frac{P[i_A=1,i_B=1]/P[i_A=1,i_B=2]}{P[i_A=2,i_B=1]/P[i_A=2,i_B=2]} = \beta_{AB}(1,1) - \beta_{AB}(2,1).$$

For the interpretation of a_1, a_2, b_1, $\beta_{AB}(1,1)$ and $\beta_{AB}(2,1)$ see equation system (2). In the present case, a_1, a_2 and b_1 are, of course, vectors.

C. Empirical Results

The empirical findings are reported in tables 2 and 3.[9] Table 2 shows the direct influence of each explanatory variable on *SKILL* and *LABEL*. Measures of total impacts are presented in table 3.

[9] Our estimates are based on LOGLIN, Seiichi Kawasaki's version of the latest Nerlove-Press computer program, as converted to IBM and extended by Houston Stokes. This program is now an option in the B34S package (Stokes, 1983).

The first two rows of table 2 indicate the direct effects associated with the respondent's plans for age 35. Mixed results emerge. Hypothesis 1 predicts that plans to be working at that time should have a positive impact on *SVP*. The estimated coefficients show that, indeed, plans to be engaging in market activity at age 35 (as opposed to being "married, keeping house, raising family") have a positive influence on the odds of having an occupation in the high-skill category rather than in

TABLE 3.—ESTIMATED PROBABILITIES AT SELECTED VALUES OF EXOGENOUS VARIABLES[a]

	High Skill[b]	Intermediate Skill	Low Skill	Male or Integrated Occupation	Female Occupation
Plans for age 35:					
Keeping house, raising children	.0956	.3892	.5152	.2769	.7231
Working	.1331	.3605	.5064	.3065	.6934
Number of children expected:					
1 child	.1007	.4008	.4984	.3030	.6969
2 children	.1022	.3818	.5160	.2892	.7109
3 children	.1036	.3631	.5333	.2756	.7244
Mother's education:					
8 years	.0605	.3390	.6004	.2750	.7250
10 years	.0858	.3647	.5495	.2815	.7185
12 years	.1196	.3858	.4947	.2891	.7110
14 years	.1634	.4000	.4365	.2980	.7020
IQ score:					
95 points	.0682	.3380	.5937	.2590	.7409
100 points	.0839	.3581	.5581	.2719	.7282
110 points	.1244	.3931	.4826	.2997	.7004
115 points	.1497	.4069	.4434	.3147	.6853
Age:					
26 years	.0969	.3779	.5252	.2649	.7351
28 years	.1013	.3769	.5218	.2804	.7195
30 years	.1058	.3759	.5182	.2967	.7033

[a] All other variables are set at their mean values. Probabilities may not add up exactly to 1 due to rounding errors.
[b] The probabilities in this column are obtained in each case by adding two estimated probabilities: $P[SKILL = 1$ and $LABEL = 1]$ and $P[SKILL = 1$ and $LABEL = 2]$. The figures in the other columns are computed similarly.

the low or intermediate categories (t-scores = 2.0 and 2.3); they have no influence, however, on the odds of having an occupation in the intermediate-skill category versus the low-skill category.

We expected "plans for age 35 = don't know" to have positive coefficients in the first three columns, reasoning that those who definitely plan to be working in the market should be the most likely to invest in job related skills, those who definitely plan to be engaging in home activities to be the least likely, with those undecided falling in between. The results, however, indicate that the "don't know" group is, if anything, *less* likely to invest in labor market skills than those who have a strong commitment to home activities (t-scores = -1.5 and -1.9 in columns 1 and 2). Perhaps these findings reflect the fact that women who intend to be home at age 35, keeping house and raising children, have at least some incentive to invest in job related skills—as an insurance against the possibility they may have to use them—whereas the don't know group is lacking in motivation.

The number of children expected, our other proxy for anticipated labor force attachment, has an effect in the expected direction on the odds in favor of being in the intermediate-skill rather than the low-skill group ($t = -1.8$); however, it does not affect the choice between high-skill versus low-skill occupations, or between high-skill and intermediate-skill occupations.

Table 3 shows that the total impact of plans to be engaging in market work at age 35 is to increase the probability of observing the respondent in a high skill occupation from 0.0956 to 0.1331. The probabilities of observing her in an intermediate or low skill occupation decrease, respectively, from 0.3892 to 0.3605, and from 0.5152 to 0.5064. As expected family size increases from 1 to 3, the probability of the respondent being in a low skill occupation rises from 0.4984 to 0.5333. These effects are in the expected direction; quantitatively, the magnitudes appear to be relatively small.

Does anticipated labor market attachment have a direct impact on *LABEL*? Expected family size has a negative influence on the likelihood of observing the respondent in a male or integrated occupation; however, the significance of this effect is very low ($t = -1.5$). The effects of plans for age 35 on this likelihood are insignificant at all conventional levels. Overall, it appears that anticipated labor supply has a stronger direct impact on the skill dimension of occupation than on the sex composition dimension.[10]

The mother's occupation variables, shown in rows 4 and 5 of table 2, have no impact on the sex label of the respondent's occupation. With one exception where significance is marginal ($t = 1.8$ in the third column), no effects on the skill level are found either.[11] It might be argued that, to the extent mother's occupation affects the respondent's expected labor force attachment, inclusion of the first three variables in table 2 clouds the testing of the sex role socialization hypothesis; however, the association between mother's occupation on the one hand, and respondent's plans for age 35 and number of children expected on the other, is weak. Thus, it is not surprising that omission of the first three variables from our model did not appreciably change the estimated effects of mother's occupation.[12] We conclude that the sex role socialization theory is not supported by these data.[13]

Turning to the control variables, while "only child" and father's education have effects on *SKILL* in the expected direction, these are not statistically significant. Mother's education has no direct impact on *LABEL*, but it affects *SKILL* significantly. In accordance with our expectations, daughters of more educated women are more likely to be in occupations with high *SVP*s (t-ratios = 5.9, 3.5, 3.7). A large total impact on *SKILL* emerges. An increase in mother's schooling from 8 to 14 years increases the probability of observing the respondent in a high and in an intermediate-level occupation from 0.0605 to 0.1634 and from 0.3390 to 0.4000, respectively. Ability also has the expected positive impact on *SKILL* (t-scores: 7.7, 5.0, and 4.2). Interestingly, an increase in the IQ score is associated with a rise in the probability that the respondent will be found in a male or integrated occupation ($t = 2.4$). Examination of the total effects shows that an increase in the IQ score from 95 to 115 raises the probability of finding the respondent in a high-skill occupation and in an intermediate-skill occupation, respectively, from 0.0682 to 0.1497, and from 0.3380 to 0.4069. At the same time, the probability that her occupation will be a male or integrated one rises from 0.2590 to 0.3147. The final control variable is age. Older women are, other factors held constant, less likely to be found in predominantly female occupations ($t = 2.3$). While age also has effects in the expected direction on *SKILL*, these are not significant.[14]

The interaction terms reported at the bottom of table 2 reveal that, holding other factors constant, a relationship between *SKILL* and *LABEL* per-

[10] The model displayed in table 2 contains two proxies for expected labor force attachment: plans for age 35 and anticipated family size. Would our conclusions be different if only one proxy were included? To ascertain this, we estimated the model two more times, first excluding the plans for age 35 dummies, and then excluding the expected number of children variable. The coefficients and t-ratios were substantially unchanged.

[11] In order to ascertain whether these results were sensitive to our definition of the mother's occupation variables, we re-ran our model with a somewhat different specification. The first occupation variable was defined to be equal to 1 if during the respondent's teenage years, the mother worked *all* of the time, in an integrated or male occupation; the second was made equal to 1 if she worked *all* of the time in a female occupation. This left mothers who worked part-time, rarely or not at all, as the base. The estimates obtained in this way were substantially the same as those reported in table 2.

[12] The t-ratios for the mother's occupation variables (when plans for age 35 and number of children expected were excluded) were

Mother worked/male occupation	1.5	−0.6	1.9	−0.3
Mother worked/female occupation	0.1	−1.3	1.0	−0.4

[13] An anonymous referee suggested that the sex role socialization hypothesis might also be tested using measures of attitudes toward the appropriateness of women working. We constructed an index, *ATT*, based on 12 attitudinal questions included in the 1972 survey; this is the first year such questions were asked. *ATT* ranges from −19 to 24, low values denoting traditional attitudes and high values, a non-traditional orientation. In a table similar to table 2, the t-ratios associated with *ATT* in columns 1, 2, 3 and 4, respectively, were 2.0, 2.6, 0.3, and 0.02; other coefficients and t-ratios in the model remained substantially unaffected. A second specification of the index yielded similar results. (A description of the attitudinal variables and indices is contained in an appendix obtainable from the authors.) Contrary to the predictions of the sex role socialization theory, *ATT* has no impact on *LABEL*. It does, however, significantly affect *SKILL*. Perhaps respondents who have a positive attitude towards women working also have a positive attitude towards human capital investments specific to the labor market. It may also be that ATT reflects a dimension of expected labor supply not captured by the first three variables in table 2. Respondents who regard women working favorably are more likely to foresee strong labor force attachment, and, as economic theory predicts, have a greater incentive to undertake the training required for the higher skill occupations. Because of the many difficulties involved in the use of attitudinal variables (e.g., see Festinger (1957) on cognitive dissonance), these results must be interpreted with caution.

[14] In preliminary runs we also controlled for calendar year. We included a dummy variable taking the value 1 if occupation was measured as of 1975, and 0 if it was measured as of 1977 or 1978. This variable was not significant and was dropped from the model.

sists. Being in the high rather than the low or intermediate skill categories is positively associated with the probability of being in a non-traditional occupation (t-scores = 3.6 and 3.6), *ceteris paribus*. The interaction between intermediate versus low-skill occupations and the sex label of the occupation is insignificant, however. Thus, the results discussed in connection with the simple cross-tabulation in table 1 continue to hold when the effects of exogenous variables are controlled for.

D. Discussion

Our estimates show that:

(a) Anticipated labor force attachment has a direct impact on the skill level of the occupation. Although the direct effects of this variable on the sex composition of the occupation are in the expected direction, they are not statistically significant at conventional levels.

(b) The skill and sex label dimensions of occupation interact. Being in the high skill category is positively associated with the probability of being in a male or integrated occupation.

These findings suggest that, *ceteris paribus*, women who anticipate a low (high) level of labor supply will tend to choose low (high) skill occupations, and, through the interaction effect, may be expected to be found in female (male or integrated) occupations. Indeed, the total effects reported in table 3 (first five rows), though relatively small in magnitude, point in this direction: women who express a weaker commitment to market activities are more likely to be observed in low skill and in typically female occupations. Thus, the nature of our results suggests that the sex differential in labor supply behavior (women, as a group, expecting comparatively low levels of employment) may explain, at least in part, the heavy concentration of women in low skilled, traditional occupations.

Can those female dominated occupations in the higher skill categories (e.g., elementary school teachers, librarians) be explained in terms of male–female differences in labor market attachment? If anticipated labor supply had a direct impact on the sex composition dimension of occupation in the expected direction, this would be evidence of an effect leading individuals with low labor force attachment to choose traditional occupations—other factors, including the skill level, held constant. As stated in (a) above, our results do show that the direct effects of anticipated labor supply on the occupation's sex label are in the predicted direction. In light of their low significance, however, future research on the occupational distribution might pay particular attention to the phenomenon of typically female occupations in the higher skill categories.

APPENDIX

Consider the case of one trichotomous variable, A, as a function of one continuous variable, x. The log-linear probability model may be written as follows:

$$\begin{aligned}\log P[i_A = 1] &= \epsilon + e_1 + f_1 x\\ \log P[i_A = 2] &= \epsilon + e_2 + f_2 x\\ \log P[i_A = 3] &= \epsilon + e_3 + f_3 x.\end{aligned} \qquad \text{(A-1)}$$

Imposing the constraints that $e_1 + e_2 + e_3 = 0$, and $f_1 + f_2 + f_3 = 0$, we may write:

$$\begin{aligned}&\log\left\{ P[i_A = 1]/\sqrt[3]{P[i_A = 1]P[i_A = 2]P[i_A = 3]} \right\}\\ &\quad = e_1 + f_1 x\\ &\log\left\{ P[i_A = 2]/\sqrt[3]{P[i_A = 1]P[i_A = 2]P[i_A = 3]} \right\}\\ &\quad = e_2 + f_2 x\\ &\log\left\{ P[i_A = 3]/\sqrt[3]{P[i_A = 1]P[i_A = 2]P[i_A = 3]} \right\}\\ &\quad = e_3 + f_3 x.\end{aligned} \qquad \text{(A-2)}$$

Differentiation of the first equation in (A-2) with respect to x indicates that f_1 represents the change in the log odds that A will take the value 1, *relative to the geometric average of the probabilities*, associated with a unit change in x. Similar statements may be made about f_2 and f_3.

If the constraints that $e_3 = 0$ and $f_3 = 0$ are imposed instead, from (A-1) we obtain:

$$\begin{aligned}\log\{ P[i_A = 1]/P[i_A = 3]\} &= e_1 + f_1 x\\ \log\{ P[i_A = 2]/P[i_A = 3]\} &= e_2 + f_2 x.\end{aligned} \qquad \text{(A-3)}$$

In this case, f_1 represents the change in the log odds that A will take the value 1 *rather than* 3 associated with a unit change in x. f_2 may be interpreted analogously.

In general, one approach is to impose the restriction that if any one of the effects is summed over all the values of one of the indices on which it depends the sum should equal 0; this leads to comparisons between the probability of being in a particular category and the geometric average probability. An alternative strategy is to constrain the last effects to be zero; this leads to comparisons between the probability of being in a particular category and the probability of being in the last category.

REFERENCES

Becker, Gary S., *Human Capital* (New York: Columbia University Press for the National Bureau of Economic Research, 1964).

Beller, Andrea, "Occupational Segregation by Sex: Determinants and Changes," *The Journal of Human Resources* 17(3) (1981), 371–391.

Bishop, Yvonne, Stephen E. Fienberg and Paul W. Holland, *Discrete Multivariate Analyses: Theory and Practice* (Cambridge: The MIT Press, 1975).

Blakemore, Arthur E., and Stuart A. Low, "Sex Differences in Occupational Selection: The Case of College Majors," this REVIEW 66(1) (1984), 157–163.

Chiswick, Barry R., "Schooling, Screening and Income," in L. C. Solmon and P. J. Taubman (eds.), *Does College Matter? Some Evidence on the Impacts of Higher Education* (New York: Academic Press, 1973), 151–158.

Chiswick, Barry R., June A. O'Neill, James S. Fackler and Solomon W. Polachek, "The Effect of Occupation on Race and Sex Differences in Hourly Earnings," *Business and Economic Statistics Section, Proceedings of the American Statistical Association* (1974), 219–228.

England, Paula, "Explanations of Occupational Sex Segregation: An Interdisciplinary Review," manuscript, University of Texas at Dallas (1982a).

______, "The Failure of Human Capital Theory to Explain Occupational Sex Segregation," *The Journal of Human Resources* 17(3) (1982b), 358–370.

England, Paula, Marilyn Chassie, and Linda McCormack, "Skill Demands and Earnings in Female and Male Occupations," *Sociology and Social Research* 66(2) (1982), 147–168.

Featherman, David L., and Robert M. Hauser, "Sexual Inequalities and Socioeconomic Achievement in the U.S., 1962–1973," *American Sociological Review* 41 (1976), 462–483.

Festinger, Leon, *A Theory of Cognitive Dissonance* (Evanston: Row, Peterson and Company, 1957).

Goodman, Leo, *Analyzing Qualitative / Categorical Data* (Cambridge: Abt Books, 1978).

Jusenius, Carol, "The Influence of Work Experience, Skill Requirement, and Occupational Segregation on Women's Earnings," *Journal of Economics and Business* 29(2) (1977), 107–114.

Koo, Helen P., and Barbara K. Janowitz, "Interrelationships between Fertility and Marital Dissolution: Results of a Simultaneous Logit Model," *Demography* 20(2) (1983), 129–145.

Lehrer, Evelyn L., "Determinants of Child Care Mode Choice: An Economic Perspective," *Social Science Research* 12 (1983), 69–80.

______, "Log-Linear Probability Models: An Application to the Analysis of Timing of First Birth," (1984) (forthcoming in *Applied Economics*).

Lehrer, Evelyn L., and Marc Nerlove, "The Impact of Expected Child Survival on Husbands' and Wives' Desired Fertility in Malaysia: A Log-Linear Probability Model," *Social Science Research* 13(3) (1984), 236–249.

Lehrer, Evelyn L., and Seiichi Kawasaki, "Child Care Arrangements and Fertility: An Analysis of Two-Earner Households" (1983) (forthcoming in *Demography*).

McClendon, McKee, "The Occupational Status Attainment Process of Males and Females," *American Sociological Review* 41(1) (1976), 52–64.

Mincer, Jacob, and Solomon Polachek, "Family Investments in Human Capital: Earnings of Women," *Journal of Political Economy* 82(2/2) (1974), S76–S108.

Nerlove, Marc, and S. James Press, *Univariate and Multivariate Log-Linear and Logistic Models* (Santa Monica: The Rand Corporation, 1973).

______, "Multivariate Log-Linear Probability Models for the Analysis of Qualitative Data," Northwestern University, Discussion Paper No. 1 (1976).

Polachek, Solomon, "Occupational Self Selection: A Human Capital Approach to Sex Differences in Occupational Structure," this REVIEW 63(1) (1981), 60–69.

Schmidt, Peter, and Robert P. Strauss, "Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach," *Econometrica* 43(4) (1975a), 745–755.

______, "The Prediction of Occupation Using Multiple Logit Models," *International Economic Review* 16(2) (1975b), 471–486.

Scoville, James G., "Education and Training Requirements for Occupations," this REVIEW 48(4) (1966), 387–392.

Shaw, Lois B., "Does Working Part Time Contribute to Women's Occupational Segregation?" presented at the annual meeting of the Midwest Economics Association, St. Louis, Missouri (1983).

Smith, Sharon, "Men's Jobs, Women's Jobs and Differential Wage Treatment," in *Job Evaluation and EEO: The Emerging Issues*, papers presented at the Industrial Relations Counselors Colloquium, September 14–15, 1978, Atlanta (New York: Industrial Relations Counselors, 1979), 67–84.

Stokes, Houston, "The B34S Data Analysis Program: A Short Writeup," College of Business Administration Working Paper Series, University of Illinois at Chicago, Report FY 77-1, revised (1983).

Treiman, Donald J., and Kermit Terrell, "Sex and the Process of Status Attainment: A Comparison of Working Women and Men," *American Sociological Review* 40 (1975), 174–200.

U.S. Department of Labor, *U.S. Working Women: A Databook* (Washington, D.C.: U.S. Government Printing Office, 1977).

______, *Selected Characteristics of Occupations Defined in the Dictionary of Occupational Titles*, Employment and Training Administration (Washington, D.C.: U.S. Government Printing Office, 1981).

Zalokar, Nadja, "Male-Female Differences in Occupational Choice and the Demand for General and Occupation-Specific Human Capital," manuscript, University of Florida, Gainesville (1984).

Zellner, Harriet, "The Determinants of Occupational Segregation," in C. Lloyd (ed.), *Sex, Discrimination and the Division of Labor* (New York: Columbia University Press, 1975), 125–145.