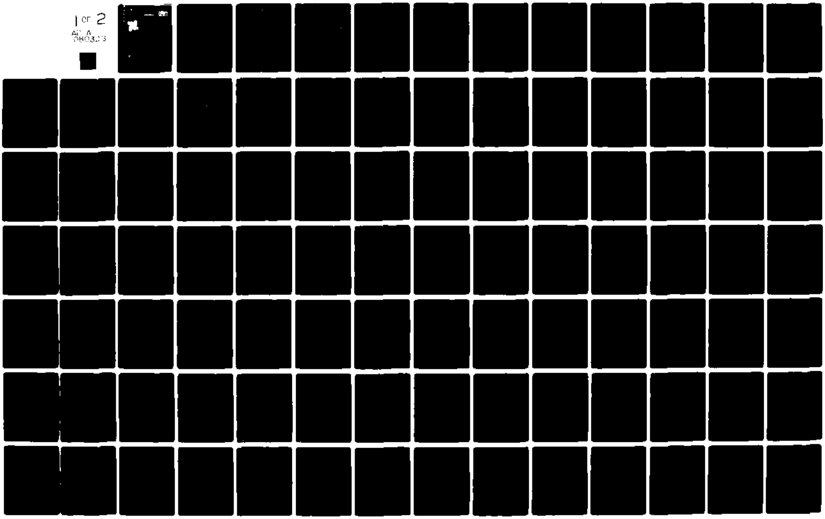
AD-A080 323 MATHTECH INC PRINCETON NJ F/G 5/9
ANALYSIS OF ENLISTMENT IN THE NAVY AS AN OCCUPATIONAL CHOICE: P--ETC(U)
JAN 80 D E WISE, R L HORST N00014-78-C-0488
UNCLASSIFIED 2033 NL

1 of 2
AD A
080323



LEVEL II (12)

This report was prepared under the Navy Manpower R&D Program of the Office of Naval Research under Contract No. N00014-78-C-0488.

ANALYSIS OF ENLISTMENT IN THE NAVY AS AN OCCUPATIONAL CHOICE: PRELIMINARY RESULTS

Prepared For:

Office of Naval Research
U.S. Department of the Navy
Washington, D.C. 22217

Prepared By:

MATHTECH, Inc.
P. O. Box 2392
Princeton, New Jersey 08540

January 28, 1979

LEVEL II (12)

This report was prepared under the Navy Manpower R&D Program of the Office of Naval Research under Contract No. N00014-78-C-0488.

ANALYSIS OF ENLISTMENT IN THE NAVY AS AN OCCUPATIONAL CHOICE: PRELIMINARY RESULTS

Prepared For:

Office of Naval Research
U.S. Department of the Navy
Washington, D.C. 22217

Prepared By:

MATHTECH, Inc.
P. O. Box 2392
Princeton, New Jersey 08540

January 28, 1979

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER — (9) Final rept. 15 Jun 78 - 28 Jan 80

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle) — (6) Analysis of Enlistment in the Navy as an Occupational Choice: Preliminary Results

5. TYPE OF REPORT & PERIOD COVERED — Final Report 6/15/78 - 1/28/80

6. PERFORMING ORG. REPORT NUMBER — (14) 2033

7. AUTHOR(s) — (10) Donald E. Wise, Robert L. Horst, Jr.

8. CONTRACT OR GRANT NUMBER(s) — (15) N00014-78-C-0488

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Mathtech, Inc.
P. O. Box 2392
Princeton, NJ 08540

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
62763N, RF63-521-002
NR 170-791

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217
(11) 28

12. REPORT DATE — Jan[redacted]80

13. NUMBER OF PAGES — 97

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) — (12) 100

15. SECURITY CLASS. (of this report) — DO-C9

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

(16) F63521

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(17) RF63521002

18. SUPPLEMENTARY NOTES

Supported by Office of Naval Research Manpower R&D Program.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Occupational Choice, Enlistment, Discrete Event Estimation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report presents the results of a project designed to predict the probability of an individual joining the Navy. The data was taken from the National Longitudinal Study of the High School Class of 1972, the Dictionary of Occupational Titles, and the 1970 Census of Population. A model of initial occupational choice was developed and estimated using maximum likelihood techniques. The findings and implications of this research are discussed.

394 224

TABLE OF CONTENTS

I. INTRODUCTION AND CONCLUSIONS

This paper describes the process and results of Mathtech's project to develop a model of initial occupational choice which would help the Navy to determine likely candidates for enlistment in the Navy. The results of this study can help the Navy to direct its recruitment programs toward those individuals who are most likely to join the Navy.

The object of the study has been to develop a model of initial occupational choice which could be used to predict the probability that a given individual would enlist in the military service and, in particular, to enlist in the Navy. In this study we have given equal attention to military and civilian choices. This has been done in order to make use of the richness of our data on civilian job choices in modeling the enlistment decision. It may appear, therefore, that this study concentrates on the civilian labor market and not on military enlistments. The entire study, however, has been directed at shedding light on the individual's decision process when considering enlistment.

The primary data source used in this study is the National Longitudinal Study of the High School Class of 1972. This survey was conducted for the National Center for Education Statistics and contains detailed data on 1972 high school graduates. These individuals were surveyed about their work and educational experience in 1972, 1973, 1974, and 1976. This data set is, therefore, ideal for studying the initial occupational choice decisions for young men and women. It should be noted, however, that the survey contains information on 1972 high school graduates only, and therefore the results reported here should not be extrapolated to the general population without a great deal of care.

There are two primary findings of this study.

- Individuals most receptive to Navy recruitment come from middle to lower middle income families and have average to slightly below average mathematical and vocabulary skills.
- Future income considerations are of primary importance in initial choice of occupation. Individuals in our study seem to give little or no weight to current or very near future earnings.

The first of these results is certainly not surprising and, in fact, this work does not suggest a change in the existing recruiting policies of any of the military services. Indeed, this work suggests that existing recruiting policies appear to be directed at the appropriate segments of the population.

The description of an individual most likely to be receptive to Navy recruiting would be a single 17 year old black with parents' income slightly (10 percent) below average who scores slightly (10 percent) below average on mathematics and vocabulary ability exams. While such individuals would be the most likely to join the Navy, the probability of their joining would be only about 5 percent; i.e., only one in twenty would join.

The second primary finding of the study is that initial wages are not important in choosing an initial occupation. The individuals in our data set seem to choose their initial occupation on the basis of future income considerations in conjunction with job quality considerations.

Current income has little or no effect. We find this result to be particularly interesting since it indicates that, in recruiting, emphasis should be placed on the skills that will be learned in the Navy and how these skills will prepare the individual for future private sector careers if the individual does not wish to make a career of the Navy.

The remainder of this report describes the research we have conducted. Section II describes the occupational choice model and estimation procedure that we have used. Section III describes the data sources that we have used. Section IV describes our estimates of wage rates, both immediate and future, for the various occupational classes we have identified. Section V reports on job characteristic data and procedures we have used to aggregate the data from thousands of detailed job titles up to the major occupational grouping we planned to use in the estimation phase of our work. Finally, Section VI details the estimation phase of our work. This section describes the modifications we were forced to make in our procedures, the estimated parameters, and the conclusions that we have drawn from those procedures.

II. ESTIMATION OF A MODEL OF DISCRETE CHOICE

A. Theoretical Underpinnings of Estimation Technique

The individual's decision to enter an occupation can be viewed as a problem in discrete choice; i.e., the individual must choose from among a finite number of occupational alternatives, which, in this case are differentiated by the combinations of earnings and job characteristics each offers. In order to simplify the discussion to follow we assume that the individual is faced with only three such possible alternatives and that each alternative is associated with an income level and an index of "working conditions." This situation is represented graphically in Figure II-1.

In theory the solution to this problem is straightforward: the individual chooses the occupation which provides the greatest level of satisfaction; i.e., the one which maximizes his utility function over the set of possible occupations. Figure II-2 adds a set of indifference curves to Figure II-1 and indicates the chosen occupation, given the particular set of indifference curves.

Of course, different individuals have different indifference curves and as a result make different choices. The individual depicted in Figure II-2, for example, is very sensitive to "working conditions" whereas the individual shown in Figure II-3 is particularly sensitive to income.

While the problem of occupational choice is quite simple in theory, it poses serious difficulties for empirical work. In contrast to problems in continuous choice, there are no continuous first order conditions to estimate. As a result, the problem is not easily reduced to estimation of a demand curve, instead the analyst is left trying to estimate the

Figure II-1

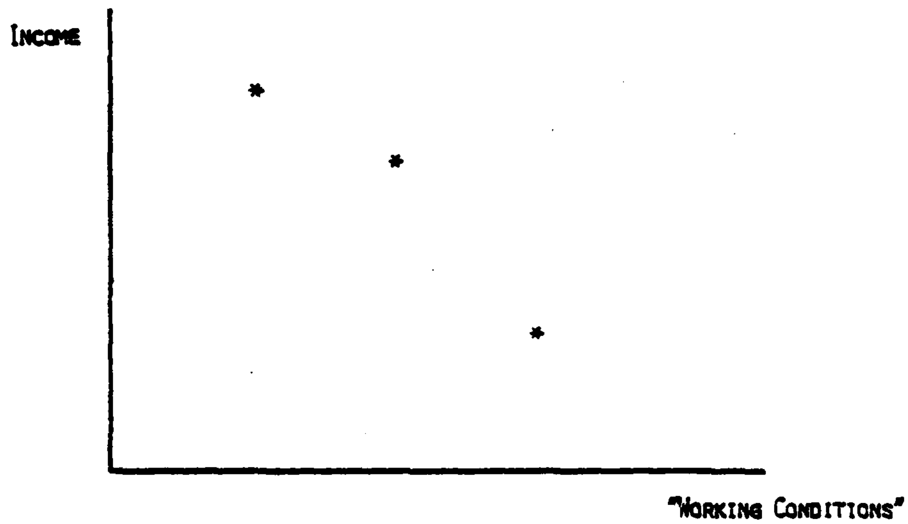


Figure II-2

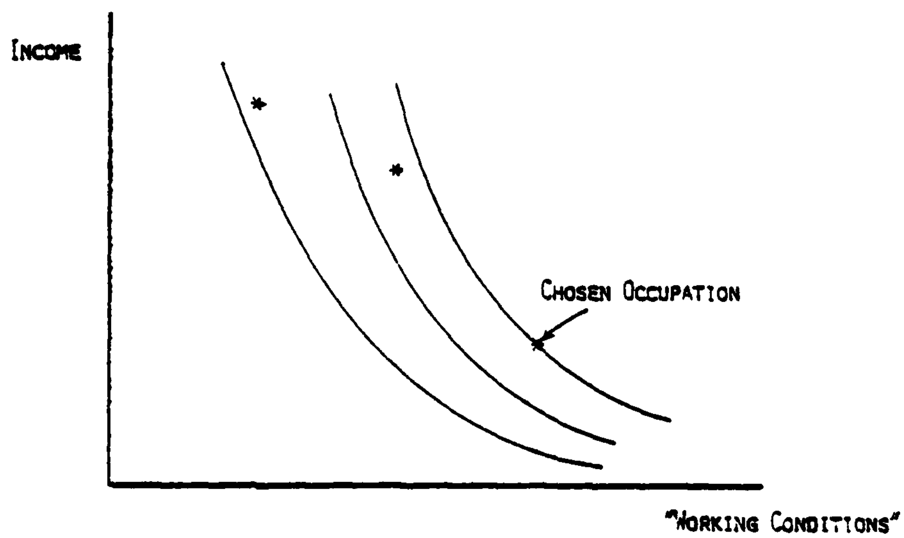
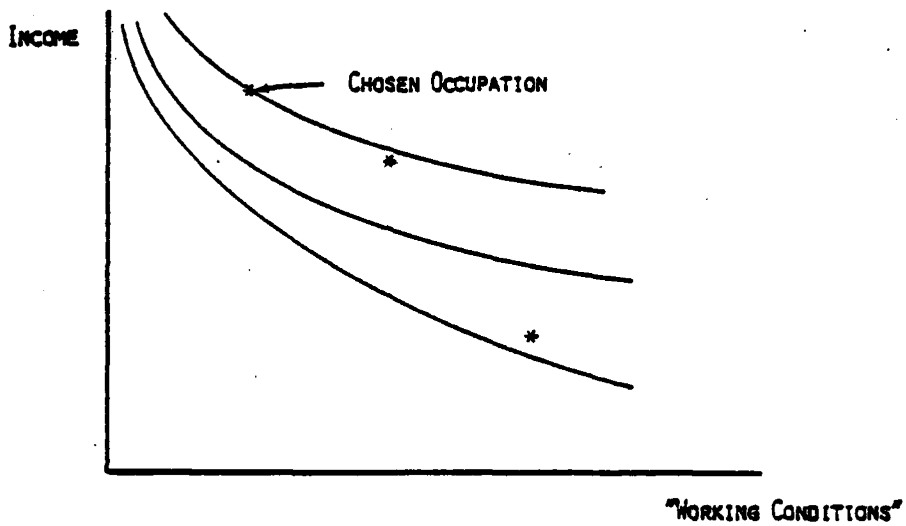


Figure II-3



probability of choosing one of the occupations. Generally, this is accomplished by use of a linear probability, probit, or logit model.

All of these methods have the drawback that the estimated functional form cannot be derived from a utility maximizing model of the type described above. At best, the resulting estimates can be regarded as conditional on the structure of the data used in estimation. A question of great analytical interest is one where shifts in the structure of the choice problem occur; i.e, where new occupational fields open and close, where relative income levels change, or where "working conditions" change (perhaps because of technological change or OSHA regulations). The approach we take in this work allows such changes and, indeed, takes them into account explicitly.

We begin by specifying an explicit utility function. Several possibilities for the fuctional form exist. Stone-Geary, Cobb-Douglas, CES, and Trans-log are but a few of the possibilities. For the purpose of this example we assume that utility is given by the Cobb-Douglas function:

$$(1) \qquad U = Y^{\alpha} W^{1-\alpha}$$

In this formulation Y is a measure of lifetime income for the occupation, W is a measure of "working conditions," and α is the weight the individual assigns income in the occupational choice decision.[1] Individuals with strong preferences for income will have α's near one, while those with strong preferences for pleasant "working conditions" will

1. While doing so at this point would only complicate the example, we could decompose Y into its various components — present direct earnings, future earnings, fringe benefits, etc. The components of W and α will, however, be considered here, because they aid in clarification of the estimation technique.

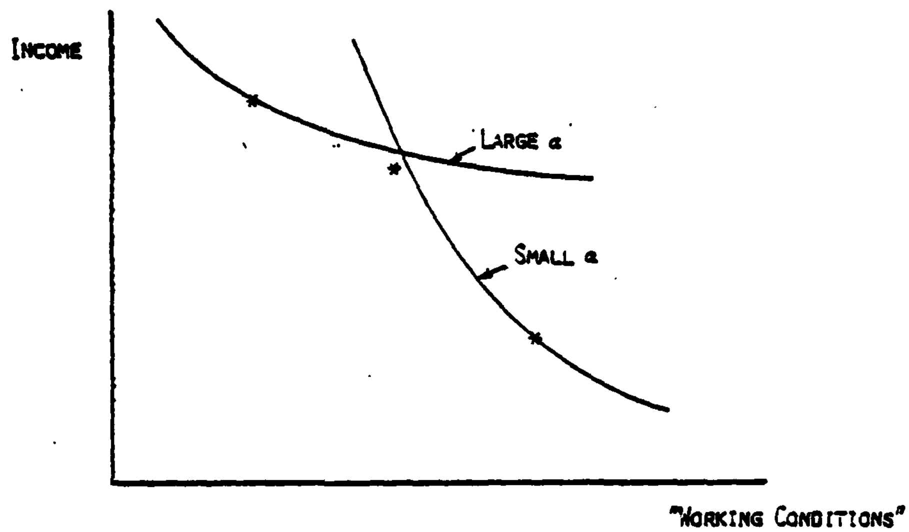
have small α's (see Figure II-4).

It is clear that different individuals assign different levels of importance to income over working conditions; i.e., that α differs from individual to individual. Not only does α differ across individuals, but it seems to differ in a somewhat systematic fashion. For example, older workers as well as wealthy workers seem to have stronger preferences for leisure than younger workers and/or poorer workers. In addition, we would expect cultural differences such as race, education, sex, and family composition, to have an impact on the value of α . For expository reasons we will simply assume that α is a linear function of these variables. Letting $\underline{X}$ represent the vector of individual characteristics, $\underline{A}$ a vector of coefficients, and "e" a residual which captures those differences purely related to individual differences in taste, we can write:

(2) $\alpha = \underline{AX} + e.$

In addition to the problems relating to differences in tastes mentioned above, the term "working conditions" needs to be operationalized. It is very clear that "working conditions" encompasses a large number of factors, including whether the work is outdoors, clean, physically demanding, or supervisory. Obviously, this list could be extended almost without limit. Assuming that occupations can be adequately characterized by a list of n job characteristics, we are still faced with the problem of representing these characteristics with a single index of job quality or working conditions. Again, we assume for simplicity of exposition that job quality can be approximated by a linear function of job characteristics (J's), so that $W = \underline{BJ}$, where $\underline{J}$ is a vector of weights and W is a scalar.

Figure II-4



The utility function given in equation (1) can now be written as:

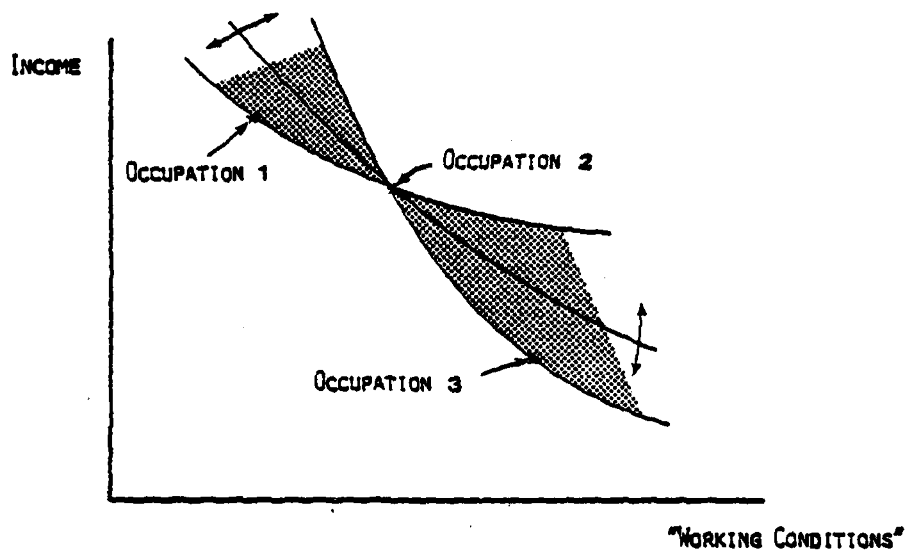
$$(3) \qquad U = Y^{(\underline{AX} + e)} \cdot (\underline{BJ})^{(1 - \underline{AX} - e)}$$

This function can be estimated by maximum likelihood methods if the distribution of e is known. For the purpose of this discussion we assume that e is distributed normally with mean zero and variance s^2. We also assume, temporarily, that $\underline{A}$ and $\underline{B}$ are known. We now ask the question: "What is the probability of observing an individual with $\underline{X}$ characteristics who chooses occupation j?"

Figure II-5 shows an example of the set of indifference curves that would lead an individual to choose occupation 2. Any indifference curve passing through the point representing occupation 2 and lying in the shaded area will lead the individual to choose occupation 2, since that choice maximizes the individual's satisfaction given the alternatives open to him. Indifference curves lying outside of this area will lead the individual to choose either occupation 1 or 3.

The location of an indifference curve will depend upon the assumed parameters $\underline{A}$ and $\underline{B}$, and the value of the stochastic term e. For given values of $\underline{A}$ and $\underline{B}$ we can calculate the highest and lowest values of e which lie in the shaded area of Figure II-5; i.e., the values of e which make the individual indifferent between occupation 2 and occupation 1 and 3. Since e is a stochastic variable with mean zero and variance s^2, it is an easy process to calculate the probability that the observed value will lie between any two given values. If, for instance, the extreme values of e (e_1

Figure II-5


INCOME
OCCUPATION 1
OCCUPATION 2
OCCUPATION 3
"WORKING CONDITIONS"

and e_2) are -12 and 34, respectively, and the variance s^2 is 100, then the probability of choosing occupation 2 is 0.8846; i.e., the probability of a standard normal deviate lying between -1.2 and 3.4.

Using the procedure described above to compute the probability P_i of an individual choosing occupation i, we can evaluate the likelihood of all individuals in the sample choosing the occupation class they are observed to have entered.

The likelihood function is defined as:

$$(4) \qquad \mathcal{L} = \prod_i P_i ,$$

or written in its more typical form

$$(4a) \qquad \log \mathcal{L} = \sum_i \log P_i \ .$$

The calculation of the likelihood function is, however, only the first step in the estimation of the parameters $\underline{A}$ and $\underline{B}$. Recall that to this point we have assumed $\underline{A}$ and $\underline{B}$ are given. The reason for this now becomes clear since estimation is accomplished by arbitrarily choosing values for $\underline{A}$ and $\underline{B}$, and evaluating the likelihood value $\mathcal{L}$. The parameters $\underline{A}$ and $\underline{B}$ are then altered iteratively until the set of parameters which maximize $\mathcal{L}$ is found.

B. A Numerical Example

To illustrate this procedure more completely, we turn now to a specific example where we assume that our sample contains 100 individuals with the personal characteristics X_1 and X_2, and the lifetime potential earnings Y_1, Y_2, and Y_3, shown in Table II-1.

The values X_1 and X_2 represent characteristics of the individual, say, education and a preference for out-of-doors employment; and the values Y_1, Y_2, and Y_3 represent the lifetime incomes each individual can expect to earn in, say, the Navy, the building trades, and as economic consultants, respectively. As can be seen from the data, each individual has different potential earnings in each occupation. While Navy personnel have the lowest earnings on average in this example, for some individuals this is not the case. For example, the fourth individual has potential earnings of $137.31 in the Navy, while earning only $116.65 in the building trades.

Example values of the characteristics of the three jobs (the J 's) are given in Table II-2 below.

Table II-2

Assumed Job Characteristics

Occupation	J_1 Index of Responsibility	J_2 % of Time Outdoors
1	0.0	0.6
2	0.3	0.3
3	0.6	0.1

Table II-1

Assumed Personal Characteristics and Potential Lifetime Incomes for Hypothetical Sample

X1	X2	Y1	Y2	Y3	OCCUPATION
8.96	1.00	100.05	124.80	164.40	3
7.81	1.00	88.45	127.45	154.36	3
8.01	1.00	113.07	123.06	165.14	3
11.42	1.00	137.31	116.65	179.04	3
9.40	1.00	94.23	125.03	140.65	3
11.03	1.00	113.27	129.15	159.81	3
6.20	1.00	92.99	135.81	158.84	3
7.95	1.00	106.63	124.20	147.00	3
11.59	1.00	113.45	101.09	172.11	3
6.85	1.00	103.62	99.06	170.50	3
4.72	1.00	103.14	134.24	155.06	3
6.02	1.00	106.15	112.86	165.74	3
5.97	1.00	110.74	110.50	157.56	3
8.97	1.00	111.62	125.30	149.39	3
6.65	1.00	93.04	119.37	149.64	3
8.06	1.00	85.74	125.30	159.87	3
7.58	1.00	107.50	110.27	159.98	3
6.17	1.00	124.22	113.65	146.47	3
5.92	1.00	123.13	128.23	135.71	3
8.35	1.00	93.38	131.75	151.51	3
7.97	0.00	97.73	122.21	145.32	3
8.93	0.00	108.60	124.02	156.59	3
9.58	0.00	108.27	122.39	154.05	3
11.58	0.00	109.71	123.23	160.81	3
7.87	0.00	111.71	123.36	172.83	3
6.00	0.00	105.46	127.60	156.45	3
8.50	0.00	97.09	113.61	153.72	3
8.57	0.00	104.29	100.29	151.57	3
4.99	0.00	111.98	124.91	152.86	3
11.46	0.00	109.75	134.50	173.27	3
12.20	0.00	94.44	107.38	153.15	3
12.35	0.00	113.18	123.55	161.74	3
7.78	0.00	117.70	109.97	158.89	3
6.53	0.00	115.07	129.82	152.76	3
7.43	0.00	101.75	124.98	158.84	3
9.92	0.00	93.51	122.28	141.12	3
9.31	0.00	117.77	113.34	159.69	3
4.40	0.00	85.48	107.14	165.77	3
8.30	0.00	101.93	111.96	164.58	3
10.41	0.00	108.24	125.09	154.06	3
10.18	1.00	85.00	111.25	155.78	1
12.64	1.00	102.75	130.18	161.57	1
8.12	1.00	90.79	122.25	141.93	1
6.87	1.00	108.03	115.34	160.30	1
12.91	1.00	96.98	115.59	175.67	1
5.19	1.00	109.50	130.20	138.49	1
7.27	1.00	112.26	125.58	145.97	1
7.45	1.00	94.99	100.51	163.33	1
11.51	1.00	116.01	117.68	138.92	1
10.19	1.00	98.37	109.00	148.30	1

Table II-1 (Continued)

X1	X2	Y1	Y2	Y3	OCCUPATION
5.10	0.00	115.36	129.61	167.21	1
13.53	0.00	92.68	101.96	149.60	1
13.56	0.00	94.29	125.94	158.67	1
7.76	0.00	85.70	121.87	152.06	1
7.44	0.00	100.84	121.79	159.08	1
6.57	0.00	96.29	114.50	150.22	1
9.03	0.00	110.34	119.03	164.08	1
14.26	0.00	88.76	116.79	175.98	1
11.44	0.00	90.74	120.84	134.82	1
8.57	0.00	121.02	134.78	164.46	1
9.54	0.00	102.39	117.71	183.26	1
10.02	0.00	111.41	118.92	143.97	1
13.59	0.00	105.04	121.26	166.14	1
8.33	0.00	95.65	117.50	143.05	1
8.73	0.00	101.35	107.34	176.26	1
8.15	0.00	119.74	127.02	171.75	1
10.93	0.00	109.25	128.07	176.44	1
12.66	0.00	118.80	132.00	167.32	1
8.76	0.00	84.97	122.19	163.89	1
10.45	0.00	120.49	137.51	142.82	1
9.63	0.00	116.66	135.85	159.61	1
8.93	0.00	108.13	130.22	164.32	1
7.41	0.00	103.82	119.04	164.32	1
8.45	0.00	99.80	106.70	167.14	1
7.53	0.00	97.81	135.30	171.10	1
13.76	1.00	106.47	131.80	165.91	2
12.29	1.00	108.43	124.33	163.83	2
10.86	1.00	105.11	126.51	179.09	2
16.29	0.00	108.16	115.79	182.95	2
15.84	0.00	94.13	97.70	177.75	2
16.00	0.00	89.21	105.03	149.00	2
11.30	0.00	107.20	124.33	158.81	2
16.08	0.00	106.64	114.70	145.77	2
11.21	0.00	98.00	130.33	162.99	2
11.81	0.00	92.58	103.16	160.14	2
11.86	0.00	110.35	116.68	160.07	2
11.83	0.00	105.89	121.18	160.26	2
15.82	0.00	99.51	119.73	160.86	2
14.84	0.00	104.58	126.30	163.91	2
15.93	0.00	96.48	129.07	151.32	2
11.24	0.00	113.28	100.13	142.99	2
11.29	0.00	108.46	119.45	165.15	2
14.39	0.00	95.33	110.80	164.62	2
14.55	0.00	97.89	124.80	169.79	2
15.27	0.00	97.21	128.82	166.54	2
16.11	0.00	134.34	124.49	15[illegible].17	2
12.54	0.00	104.63	116.07	165.79	2
9.45	0.00	97.64	128.16	147.05	2
11.19	0.00	94.49	109.50	151.30	2
12.58	0.00	87.73	112.77	156.43	2

The utility function estimated for this example was:

(5) $U = Y^{\alpha} W^{1-\alpha}$

where:

$\alpha = a_0 + a_1X_1 + a_2X_2 + e$

$W = b_1J_1 + b_2J_2$

with e distributed normally with zero mean and variance s^2.

Using a variant of the Davidson-Fletcher-Powell optimization algorithm,[2] the likelihood function corresponding to this problem,

(6) $\mathcal{L} = \prod_i P_i(e_0 \leq e \leq e_1)$,[3]

was estimated. The resulting values are given in Table II-3.

---

2. See Goldfeld, Stephen M. and Quandt, Richard E., Nonlinear Methods in Econometrics, North Holland, 1972, for a description of this procedure.

3. $P_i(e_0 \leq e \leq e_1)$ refers to the probability that the true value of e was within the limiting values e_0 and e_1 for the occupation observed to be chosen.

Table II-3

Estimated Parameters

Parameter	Value
a_0	0.849
a_1	0.1255
a_2	-0.1567
b_1	5.844
b_2	19.780
s^2	0.1558

In order to illustrate the use of these parameters we will calculate the probabilities of choosing occupations 1, 2, and 3 for the four individuals shown in Table II-4.

Table II-4

Example Data

X_1	X_2	Y_1	Y_2	Y_3
9.	1.	105.	140.	150.
13.	1.	95.	145.	165.
15.	0.	100.	170.	185.
8.	0.	120.	140.	145.

These probabilities are computed by calculating the values of e (e_{12}, e_{13}, e_{23}) that leave the individual indifferent between the three pairs of occupations (1,2), (1,3), and (2,3). This is done by making use of the fact that $U_1 = U_2$ if the individual is indifferent between occupations 1 and 2.

Defining α_{ij} as the value of α which corresponds to $U_i = U_j$, we can write:

$$(7) \qquad U_i = U_j = Y_i^{\alpha_{ij}} W_i^{1-\alpha_{ij}} = Y_j^{\alpha_{ij}} W_j^{1-\alpha_{ij}} .$$

To solve for α_{ij}, we note that

$$(8) \qquad \alpha_{ij} \log Y_i + (1-\alpha_{ij}) \log W_i = \alpha_{ij} \log Y_j + (1-\alpha_{ij}) \log W_j$$

Therefore,

$$(9) \qquad \alpha_{ij} = \frac{\log W_j - \log W_i}{\log Y_i - \log Y_j + \log W_j - \log W_i}$$

Now since

$$(10) \qquad \alpha_{ij} = a_0 + a_1 X_1 + a_2 X_2 + e_{ij} ,$$

we can solve for e_{ij}, the "taste" term which leaves the individual indifferent between occupations i and j.

(11) $$e_{ij} = \frac{\log W_j - \log W_i}{\log Y_i - \log Y_j + \log W_j - \log W_i} - a_0 - a_1X_1 - a_2X_2$$

Referring back to Figure II-5, we see that e_{12} and e_{23} define the extreme values for e such that the individual will choose occupation 2. The values of e_{12} and e_{23} for the example data in Table II-4 are given in Table II-5, as well as the values for e_{13} .

Table II-5

Values for e_{ij}

Observation	e_{12}	e_{23}	e_{13}
1	.4777	.5600	.7065
2	-.1191	-.0430	.0974
3	-.5834	-.4770	-.2338
4	.5830	.6480	.7509

The probability that the first individual in our example will choose occupation 2 is the probability that a random normal variable with variance of 0.1558 (this was one of the estimated parameters given in Table II-3) takes a value between 0.4777 and 0.5600. This is 3.51%, the shaded area under the normal curve shown in Figure II-6. The probability that this individual chooses occupation 1 is the area to the left of the shaded area in the figure, with the probability of choosing occupation 3 being the area to the right of the shaded area.

Figure II-6

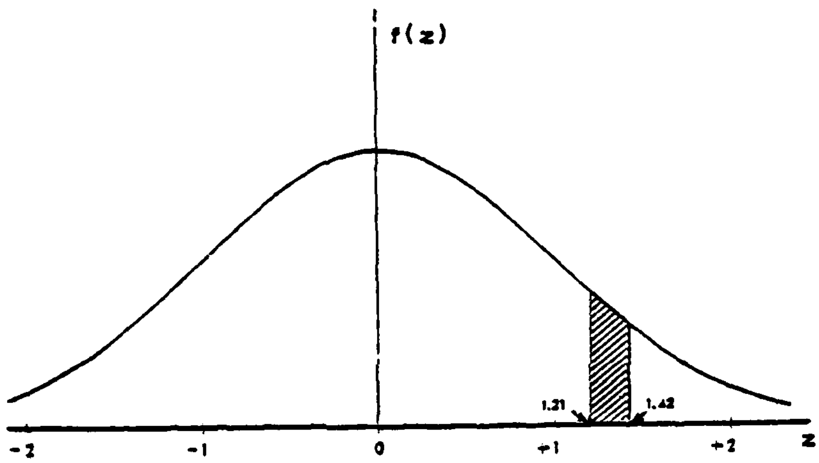


Table II-6 gives the probabilities that each occupation is chosen by the four individuals in the example.

Table II-6

Probabilities of Choosing Each Occupation

	Probability by Occupation		
	1	2	3
Observation			
1	88.69%	3.51%	7.78%
2	38.14	27.21	34.65
3	6.97	20.71	72.32
4	93.02	4.12	2.86

Now consider the effect on these probabilities of a change in the value of a job characteristic. For the purpose of the example, assume that J_1, the index of responsibility, is increased to .4 for occupation 2. We now repeat the procedure described above and find that the resulting e's and probabilities as given in Table II-7.

As we see from these results, the probability that individual one will choose occupation 2 rises 6.96 percent, with the probabilities of choosing occupations 1 or 3 falling 2.34 and 4.60 percent.

Alternatively, consider the effect on these probabilities of a change in the value of a personal characteristic. For example, assume that individual one acquires 3 additional years of education. We might phrase the question as: "How much more likely would it be for individual one to choose occupation 2 if his educational level were 12 instead of 9?"

Table II-7

Values of e and Choice Probabilities

Observation	e_{12}	e_{13}	e_{23}	P_1	P_2	P_3
1	.4327	.5602	.7324	86.35%	10.47%	3.18%
2	-.1632	-.0428	.1350	33.78	29.61	36.62
3	-.6286	-.4770	-.2042	5.56	24.68	69.75
4	.5458	.6481	.7663	91.66	5.73	2.61

The procedure for calculating this is conceptually identical to the example above. In this case, we calculate the probability of choosing occupation 2, assuming individual one possessed the higher education level, and compare it to the base case probability shown in Table II-6. In this case it is 25.98 percent; 22.47 percent above the base case. Thus, an additional three years of education makes it 22.47 percent more likely that an individual like individual one will choose occupation 2.

As should be clear from the material above, the decision modeling procedure described above is quite powerful. It can be used to predict the probability that individuals with substantially different personal characteristics will make a given choice; i.e., join the Navy, as well as predict the effect on choices resulting from changes in work requirements.

III. DATA REQUIREMENTS

There are three major categories of data required for application of the theoretical model described in Section II to the question of enlistment.

1. Data on young males (of which a number are enlistees or potential enlistees) with variables pertaining to occupational choice, personal characteristics, and wage rates that can be used to predict present earnings and to estimate the probabilities of choosing each of the occupational classes we define.

2. Data on males of various ages with variables on occupation, personal characteristics, and wage rates from which lifetime wage profiles can be predicted for each occupational class defined.

3. Data describing job characteristics that can be used to compute indices of the job attributes of our occupational classes.

Three data sets have been identified which correspond to the data requirements listed above.

1. The National Longitudinal Study of the High School Class of 1972.

2. The Public Use Samples of Basic Records from the 1970 Census.

3. The Dictionary of Occupational Titles tape.[1]

In this section of the paper we describe these data sets in some detail.

A. The National Longitudinal Study of the High School Class of 1972

The National Longitudinal Study (hereafter NLS) consists of a base-year survey administered to a national sample of high school seniors in the spring of 1972 and three follow-up surveys, the first in the fall of 1973, the second in the fall of 1974, and the third in the fall of 1976. In addition, as part of the base-year survey, data were collected on the characteristics of the high schools represented in the sample, their guidance activities, and the academic profiles of the respondents (obtained from school records). Of the 22,652 persons surveyed, 14,112 participated in all four of the surveys, and of this latter group 6,809 were male.

This extremely rich data source allows us to estimate present earnings in each of the major occupation classes we have selected, and, in

1. The National Longitudinal Study of the High School Class of 1972, Center for Educational Research and Evaluation, Research Triangle Institute, Research Triangle Park, North Carolina: May 1977.
Public Use Samples of Basic Records from the 1970 Census: Description and Technical Documentation, U.S. Department of Commerce, Social and Economic Statistics Administration, Bureau of the Census, April 1972.
Dictionary of Occupational Titles, U.S. Department of Labor, Employment and Training Administration, Fourth Edition, 1977.

combination with data collected from the two other sources to be described below, to estimate the parameters of the utility function. Included in the base-year survey were variables reflecting ability, socioeconomic status, home background, community environment, ethnicity, educational attainment, school characteristics, school experience and performance, work history and future work plans (military or otherwise), as well as variables on the noncognitive traits of the individuals and their opinions and life goals. The follow-up surveys provide information on activities pursued over the time period and monitor changes in attitudes and goals. Table III-1 contains more detailed information on the types of information collected within each of the general areas mentioned above (pp. 6-7 in NLS Review and Annotation of Study Reports).

B. The Public Use Samples of Basic Records from the 1970 Census

The task of estimating lifetime wage profiles requires a large micro data set with information on the full range of socioeconomic cohorts. The data set must contain information on age, educational attainment, ethnic group, occupation, and, of course, wage rate. The 1970 Public Use Samples of the United States Census of Population meets these requirements rather well.

The sample of the Census data we have used, the 15%, One in a Thousand (1/1000) State Sample file, contains roughly 280,000 records on individuals living in the U.S. The data available on the tapes include the three-digit census occupation code, sex, race, age, education, normal hours worked, labor earnings, other income, family relationships, native tongue, veteran status and period of service, and place of work.

Table III-1

Content Areas in Each Survey

Content Category	Spring 1972 Base-Year Survey	Fall Winter 1973 74 1st Follow-Up Survey	Fall Winter 1974 75 2nd Follow-Up Survey	Fall Winter 1976 77 3rd Follow-Up Survey
Constitutional factors	Sex, birth date, physical handicap	Sex, birth date	Sex, birth date	Sex, birth date
Ability	SAT, ACT, and test scores *(vocabulary, reading, math, letter groups, mosaic comparison, picture-number)*			
Socioeconomic status	Parental income, education, occupation	Parental education, occupation		
Home background	Number of persons dependent on parents, number of siblings in college, objects in home, language at home, parental expectation	Parental encouragement	Birth order, number of siblings	
Community environment	Type of community, *distance of* home from postsecondary schools	*Type of community where* individual lives	*Type of community where individual* lives, mobility and reasons	Type of community where individual lives, mobility and reasons
Ethnicity	Race, religion		Race	
Significant others	*Relative importance of family, peers,* school personnel to secondary program, postsecondary plans, quality *of counseling* services			
Activity status	Activity plans for Fall of 1973	Activity status in Oct. 1972, Oct. 1973	Activity status in Oct. 1974	Activity status in Oct. 1975, Oct. 1976
Educational attainment	Educational plans (entry, financial support arrangement), factors interfered attainment, school choice	Educational status (degree, certificate received), factors interfered attainment	Educational status (degree, certificate received), factors interfered attainment	*Educational status, graduate school* application and entry, factors interfered attainment
School characteristics	Secondary school size, student counselor ratio, ability grouping or tracking, racial composition, college-going ratio, school SES, teacher qualifications, *school facilities,* counseling services	Postsecondary school control and type	Postsecondary school control and type	Postsecondary school control and type

Table III-1 (Continued)

Content Category	Spring 1972 Base-Year Survey	Fall-Winter 1973-74 1st Follow-Up Survey	Fall-Winter 1974-75 2nd Follow-Up Survey	Fall-Winter 1976-77 3rd Follow-Up Survey
School experience	Time spent on homework, program of study, participation in remedial and special services programs, school quality, courses taken, instructional strategies	Program type, major study area, full-time, financial aid programs, program duration	Program type, major study area, full-time, financial aid programs, program duration	Program type, major study area, full-time, financial aid programs, program duration
School performance	Grade average, extra curricular activities	Grade average, dropout, transfer, satisfaction with schooling, total credits earned	Grade average, dropout, transfer, satisfaction with schooling, total credits earned	Grade average, dropout, transfer, satisfaction with schooling, total credits earned
Work status	Type of work, hours of work, work plans for years after graduation	Resources used for job hunting, work type, hours/week, reasons for not working	Resources used for job hunting, work type, hours/week, reasons for not working	Resources used for job hunting, work type, hours/week, reasons for not working
Work performance and satisfaction	Factors in career selection	Income, pay, and work conditions, satisfaction	Income, pay, and work conditions, satisfaction, application of job training	Income, pay, and work conditions, satisfaction, supervision, application of job training
Noncognitive traits	Self concept, locus of control	Self-concept, locus of control	Self-concept, locus of control, maturity scale	Self concept, locus of control
Goal orientations	Work and educational aspirations, expectations, and plans; life goals	Work and educational aspirations, expectations, and plans; life goals	Work and educational aspirations, expectations, and plans; life goals	Work and educational aspirations, expectations, and plans; life goals
Marriage and family	Plans to be a full-time homemaker, number of dependents	Marital status, number of dependents, income, number of children had and expect to have	Marital status, number of dependents, spouse's education and occupation, income, number of children, items owned at home	Marital status, number of dependents, spouse's education and occupation, income, number of children had and expect to have, items owned at home
Opinions	Factors interfered with education, ratings of high school	Usefulness of specialized training in high school	Participation in political activities, consumerism, quality of life	Political activities, sex-role orientation, sex and race biases, rating of high school
Military	Plans for military service	Type, training, duration, satisfaction, plans	Type, training, duration, plans	Type, training, duration, plans

With these data, lifetime wage profiles can be estimated for individuals in the primary (NLS) data set, according to the major occupational class into which the individual's job falls. This information can then be entered into the lifetime income component of the utility function.

C. The Dictionary of Occupational Titles Tape

Identification of the characteristics of the various categories of occupations is also required before the final phase of the empirical part of the study; i.e., the estimation of the utility function parameters, can be undertaken. The Dictionary of Occupational Titles tape (hereafter DOT) fills this need.

The DOT tape contains a wide range of information on more than 12,000 different occupations. There are data on the physical and mental requirements of each job, as well as on the amount of training needed and various measures to reflect the compatibility of potential workers to each job. This information can be appropriately aggregated to determine the characteristics of much broader occupational categories. Since we are focusing on individuals' decisions to train for and enter particular occupations, it is the average characteristics of alternative occupations which are relevant, rather than the observed characteristics of the job eventually chosen.

While the DOT tape does not include military occupations, a methodology has been developed which we anticipate will allow for identification of job characteristics either for the military as a whole or for each of the separate branches. Information is currently being gathered on the distribution of jobs in the military. Combining these data and data from the Occupational Source Book, which provides a cross-reference between military occupational title codes and DOT codes, we will be able to estimate values for job characteristics for this final occupational category.

Education also represents an "occupational" choice that can be made by the individual coming out of high school. Since we will be looking at individuals who participated in the NLS survey, we will assume that post-high school education job characteristics can be identified by the characteristics of the traditional job the individual actually chose upon completion of the additional education.

Finally, a subset of NLS respondents may have decided either voluntarily or non-voluntarily to not work at all. At first glance this null decision may seem unimportant. However, from the perspective of the relationship of these unemployed people to an occupational choice involving the military, such a subgroup may be important in describing in complete detail those aspects that may lead to increased levels of enlistment in the military. We will attempt to bring this group into the analysis through the definition of appropriate dummy variables.

IV. ESTIMATION OF WAGE RATES

A. Introduction

As pointed out earlier, the decision to enlist in the military instead of choosing a civilian career or continuing schooling is based in part on the individual's perceptions, at that point in time, of the wages, both present and future, the individual could earn in various endeavors. Following a description of the occupational categories we have chosen, this section documents the estimation of present and future wages by occupation.

B. Occupational Categories

The individuality of people is a well-accepted idea, and economic theory typically relies on the somewhat nebulous concept of tastes to explain observed differences in individual behavior not accounted for through variations in prices or income. In a similar manner, jobs have their inherent individuality. Because of this, at one extreme, a model of occupational choice decisions could involve many thousands of job types. However, bounded rationality dictates some aggregation of data.

If we are to estimate a decision process, we must have a reasonably large number of observations on the decision. In the case at hand, we must have a reasonably large number of individuals who choose each of the occupations that we identify. Our desire in this process is to have as much occupational detail as possible without having so few observations in an occupation that our parameter estimates become imprecise. In line with this objective we have assumed that the occupation groupings shown in Table IV-1 represent a reasonable level of aggregation.

Table IV-1

Occupational Grouping

Number	Description	Census Codes
1	Professional, Technical and Kindred Workers	001-152,154-199
2	Managers and Administrators, Excluding Farm Managers	200-245
3	Sales Workers	246-285
4	Clerical and Kindred Workers	286-399
5	Craftsmen and Kindred Workers	400-429, 432-469, 500-599
6	Operations Except Transport	600-699
7	Transport Equipment Operations	700-719
8	Laborers, Except Farm	720-799
9	Farm Related Workers	800-899
10	Service Workers, Except Private Household	900-991
11	Military	992
12	Repairmen and Maintenance	470-499
13	Electrical Related Workers	153,430,431

In addition to the occupations listed in Table IV-1, we must add two additional categories -- education and no work -- to complete the choice set over which our problem is defined.

C. Predicting Current Wage Rates

1. Estimation Procedure

Estimates of the wage an individual could earn in each of the various occupations he could choose must be calculated to model the individual's choice process. The procedure for estimating expected wages is due to Hall (1973) and involves regressing the log of the wage rate on the characteristics of the individual:

$$\log(W) = BZ.$$

where B is a coefficient vector to be estimated, and Z is the vector of characteristics.

The National Longitudinal Study of the High School Class of 1972 contains a wide variety of information on personal attributes, ranging from labor supply data to a variety of tests of educational attainment and aptitude. Of the many variables available, we chose to include those listed in Table IV-2 in the regressions predicting earnings by occupation.

Regressions were run for each of the 13 wage and salary occupational categories we have defined. The explanatory variables in these regressions included measures of socioeconomic background (Black, Hispanic, and Income), age, handicap, a proxy for the size of the labor market (Urban), a measure of educational attainment (either READ, RANK,

Table IV-2

Glossary of Variables

Variable	Description
Age	Age at graduation from high school.
Handicap	A dummy variable equal to one if the individual has a handicap limiting the type of work he can do.
Black	A dummy variable equal to one if the individual is black.
Hispanic	A dummy variable equal to one if the individual reports that he is Chicano, Puerto Rican, or other Latin-American.
Income	Before-tax income of parents as reported in 1971 questionnaire. The original data is reported in ranges which have been translated as follows:

Range ($/yr)	Value Assigned ($/yr)
0-$2,999	1,500
3,000-$5,999	4,500
6,000-$7,499	6,750
7,500-$8,999	8,250
9,000-$10,499	9,750
10,500-$11,999	11,250
12,000-$13,499	12,750
13,500-$14,999	14,250
15,000-$18,000	16,500
18,001 +	25,000
NA	8,250

Variable	Description
Urban	A dummy variable equal to one if the community exceeds 100,000 in population
Reading Score (READ)	Score on a 15 minute reading comprehension test. scores range from 21 to 80.
Picture-Number Score (P-NUM)	Score on a 10 minute test of associative memory. Scores range from 21 to 70
Math	Score on 15 minute test of basic competence in mathematics. Scores range from 21 to 70.
Vocabulary Score (VOC)	Score on a 5 minute test of vocabulary. Scores range from 21 to 80.
Rank in Class (RANK)	Percentile rank in class.

Table IV-2 (continued)

Grade Point Average (GPA)	Imputed student grade point average (scaled 1-14, with 14 equivalent to A+).
Military Experience (M-EXP)	Time in military (years rounded up).
Job Tenure (J-EXP)	Time in months worked on current job.
Prior Work Experience (P-EXP)	Work experience in months prior to current job.
Additional Education (ED+)	Years of post high school education.

P-NUM, MATH, GPA, or VOC), and four variables which account for time spent gaining labor market experience or gaining additional education.

The latter four variables deserve additional comment. Upon graduation from high school there are three ways that an individual can gain additional experience. First, the individual can continue his education; second, he can join the military; and third, he can work at a civilian job. The wage regression attempts to measure the impact of these various types of training on market wages.

The estimated coefficient of ED+ measures the value the market places on an additional year of education for each of the various occupations. Our expectation is, of course, that education contributes positively to the worker's productivity. This effect is tempered, however, by the recognition that a young worker who is concurrently attending school may not have very strong job attachment and therefore employers may be unwilling to invest in training them for a position that they are likely to leave when their education is complete. The coefficient of EDUC in the wage regressions reported below are likely to be small and may even be negative, depending on the relative sizes of these effects.

The coefficient of M-EXP measures the value the market places on a year of military experience. Again, we would expect that the coefficients of this variable would be positive, measuring the impact of military experience on productivity. However, there are questions regarding the relevance of much of military experience to civilian jobs.

The final type of experience to be considered is labor market experience. This experience can be broken down into two separate types of experience -- on-the-job experience and experience on previous jobs. The returns to these types of labor market experience can be expected to differ substantially. On-the-job experience (J-EXP) will provide job specific skills which will generally be valued more highly by the employer than previous work experience (P-EXP) with other employers. Both of these types of experience can be expected to add to the worker's productivity, with the coefficient of JT expected to be larger than the coefficient of P-EXP.

2. Regression Results

The estimates of the regression parameters for these regressions are presented in Table IV-3.

Several interesting generalizations can be made from these results. First, the strongest effects are due to on-the-job experience and prior work experience. In all but two cases these coefficients are positive and they are generally significantly different from zero. Typically, an additional year on the job results in a 4.6% increase in real wages; i.e., after adjustment for inflation.

Prior work experience also adds to wages, though these coefficients are less closely estimated than those for on-the-job experience. Returns to prior work experience are typically about 3.8% per year, approximately 0.8% less than the return to on-the-job experience.

The estimated coefficient of military experience was generally negative, indicating that for individuals in a given occupation those with military experience receive lower wages. While this result seems counterintuitive, care must be taken in its interpretation. It does not necessarily imply that military training detracts from wages. It may be that military training allows the individual access to an occupation that was not previously open to him. As result, he may be in a higher paying occupation than otherwise, even though he is on the low end of the wage

Table IV-3

Young Worker Wage Regressions by Occupation

	OCC 1 Profess./ Tech. & Kindred	OCC 2 Mgrs. & Admin. Exc. Farm	OCC 3 Sales Workers	OCC 4 Clerical & Kindred Workers	OCC 5 Craftsmen & Kindred Workers	OCC 6 Operatives Except Transp.	OCC 7 Transport Equip. Operatives	OCC 8 Laborers Exc. Farm	OCC 9 Farm Related Workers	OCC 10 Service, Exc. Priv. Household	OCC 11 Military	OCC 12 Repairmen & Maint.	OCC 13 Electrical Related Employmt.
CONSTANT	0.5484 (.543)	1.0429[a] (.526)	1.3067[a] (.778)	1.0622[b] (.418)	1.8408[b] (.4504)	1.9260[b] (.3698)	.8678 (.8099)	1.8952[b] (.6012)	4.1856[b] (1.4632)	0.9871[b] (0.3272)	-.8669 (1.257)	1.2827[b] (0.5637)	2.7576[b] (1.0369)
AGE	0.0338 (.0298)	0.0098 (.0270)	-.0017 (.0414)	.0046 (.0231)	-.0241 (.0243)	-.0353 (.0194)	.0126 (.0411)	-.0279 (.0324)	-.1907 (.0783)	-.00004 (.0166)	.0992 (.0676)	-.0088 (.0291)	-.0815 (.0548)
HANDICAP	-.0534 (.0589)	-.0181 (.0578)	.0613 (.0808)	-.0324 (.0578)	-.0039 (.0460)	-.0342 (.0443)	-.1663[a] (.0904)	.0443 (.0792)	.1403 (.1314)	-.1062[b] (.0536)	-.4663[b] (.2207)	-.0181 (.0632)	-.0702 (.1130)
BLACK	-.0045 (.0537)	.0654 (.0699)	.1198 (.0965)	-.0143 (.0420)	-.1250[b] (.0495)	-.0458 (.0407)	.0394 (.0838)	-.1286[b] (.0589)	-.1105 (.2303)	-.0873[a] (.0484)	.1114 (.0906)	.0923 (.0672)	-.0637 (.1124)
HISPANIC	-.1259 (.0843)	-.0605 (.0641)	.2079[a] (.1084)	-.0309 (.0619)	.0559 (.0724)	.0047 (.0639)	.0911 (.1195)	-.0160 (.0919)	.0204 (.2183)	.1224 (.0734)	.0051 (.1662)	.0531 (.0820)	.0640 (.1149)
INCOME (000's)	.0043[a] (.0024)	.0092[b] (.0022)	.0062[a] (.0033)	-.0064[b] (.0024)	.0050[b] (.0025)	.0083[b] (.0025)	.0130[b] (.0048)	.0056 (.0032)	.0020 (.0063)	.0093[b] (.0026)	.0081 (.0058)	.0095[b] (.0031)	.0016 (.0056)
URBAN	-.0283 (.0413)	.0645 (.0415)	-.0753 (.0835)	.0317 (.0348)	.0757[a] (.0413)	-.0638 (.0435)	.1003 (.0752)	.0278 (.0575)		.0880[b] (.0426)	.1329 (.1072)	.0810 (.0485)	-.0339 (.0843)
READ												.0030 (.0019)	
RANK			.0001 (.0009)		.0008 (.0006)			.0006 (.0009)			.0010 (.0015)		
P-NUM						.0027[a] (.0014)				.0019 (.0018)			.0008 (.0033)
MATH		0.0005 (0.0018)							.0066 (.0050)				
GPA	.0228 (.0053)			.0112[b] (.0054)									
VOC							.0032 (.0031)						
M-EXP	-.0611 (.0375)	-.0537[b] (.0267)	.0191 (.0710)	-.0266 (.0233)	-.0170 (.0223)	-.0417[b] (.0186)	-.0320 (.0276)	-.0030 (.0305)	-.0495 (.0869)	.0130 (.0189)	.0241 (.0383)	-.0101 (.0233)	.0008 (.0353)
J-EXP	.0032[b] (.0011)	.0034[b] (.0008)	.0008 (.0013)	.0053[b] (.0007)	.0039[b] (.0008)	.0035[b] (.0007)	.0048[b] (.0014)	.0028[b] (.0099)	.0002 (.0017)	.0054[b] (.0010)		.0041[b] (.0009)	.0074[b] (.0015)
P-EXP	.0074[a] (.0040)	.0088[a] (.0045)	.0002 (.0025)	.0005 (.0023)	.0026 (.0033)	-.0017 (.0023)	.0001 (.0041)	.0007 (.0032)	.0010 (.0026)	.0101[b] (.0039)		.0063[b] (.0029)	-.0021 (.0053)
ED:	.0138 (.0130)	-.0115 (.0125)	.0013 (.0191)	-.0291[b] (.0134)	.0074 (.0181)	-.0304 (.0185)	-.0789[b] (.0346)	-.0010 (.0211)	-.0311 (.0397)	-.0034 (.0164)	.0458[a] (.0278)	-.0154 (.0223)	.0531 (.0337)
R^2	.0969	.0995	.0381	.1419	0.0782	.0888	.1298	.0609	0.1110	0.1554	0.1434	0.1003	0.2096
$\bar{R}^2$	.0780	.0776	-.0143	.1227	.0621	.0743	.0924	.0298	0.0039	0.1285	0.0470	0.0731	0.1395
F (v_1, v_2)	5.114 (11,524)	4.531 (11,451)	0.73 (11,202)	7.38 (11,491)	4.853 (11,629)	6.124 (11,691)	3.472 (11,256)	1.961 (11,333)	1.036 (10,83)	5.787 (11,346)	1.488 (9,80)	3.697 (11,365)	2.990 (11,124)
$P(F(v_1, v_2))$	0.1238×10^{-6}	$.17 \times 10^{-5}$	0.711	$.27 \times 10^{-10}$	$.32 \times 10^{-6}$	$.12 \times 10^{-8}$	$.16 \times 10^{-3}$	$.31 \times 10^{-1}$	.421	$.13 \times 10^{-7}$	.17	$.52 \times 10^{-4}$	$.15 \times 10^{-2}$
N	536	463	214	503	641	703	268	345	94	358	90	377	136
EXP $(\overline{\log(W)})$	4.33	4.23	4.02	3.94	4.85	4.57	4.26	4.50	2.94	3.60	3.17	4.31	4.33

scale. Another possibility is that we are seeing the effects of the rumored backlash to the Vietnam conflict. The most likely explanation stems from the fact that we are observing individuals who have been out of high school only four years. Those who have had military experience and are already back in the civilian labor force may well be the least desirable potential employees; e.g. they may have been discharged from the service because of poor attitudes or other undesirable traits. As a result, the military experience variable is probably serving as a proxy for someone who has a wide range of undesirable personality characteristics.

The other variable that deserves particular attention is AGE. The first effect of age on earnings that comes to mind is that older workers earn more, because they are more mature and have had more experience. This effect is countered by a second effect, however. All of the individuals in our sample graduated from high school at the same time. Those who are older graduated from high school later in life. This may have been due to many things, but the predominant reason will be that the oldest of these workers repeated a grade, whereas the youngest were advanced a grade. Thus, age may serve as an indication of performance in school, with the best performers being the youngest.

These two effects of age on wage rates result in coefficients which may be either positive or negative in sign, depending on the relative strengths of these effects. In general, we would expect those occupations which have rapidly rising wage profiles or which value physical maturity highly to show positive coefficients for age. For those occupations which have relatively flat wage profiles we would expect negative coefficients. To some degree this is borne out by the results, but the evidence is sufficiently weak that no strong statements can be made other than to note that either sign is possible. In addition, this coefficient has a relatively large standard error so that our estimate of the net effect is very imprecise.

The variable INCOME performs basically as expected. Individuals coming from higher socioeconomic strata find higher paying jobs. The one exception to this is the regression for clerical and kindred workers, where a strong negative effect is found. We see no obvious explanation for this, however.

The final variable that we discuss is ED+, the measure of post high school education. In eight out of our 13 occupational groups this coefficient has a negative sign. As was the case with AGE, ED+ is serving as a measure of two different effects. The obvious one is that more education means the worker becomes more productive, he does his job better because he has more skills. The other effect is that a worker who is simultaneously attending school is unlikely to stay with the firm for as long as one who is not attending school. In effect, school attendance signals that the worker hopes to improve his skills and move on to a better job, very likely with another firm. Those occupations which have positive coefficients are also those where one would expect education to be most highly valued; e.g., professionals, craftsmen, military, and electrical.

These regressions provide us with a means of estimating the wage rate for an individual going into each of 13 broad occupations. Within a single occupation class there is little variation in wages which is explained by the regression; the highest R^2 being only 0.1395.

Substantial differences in wage rates do exist from occupation to occupation, however, and it is these differences which will be important in the occupational choice decision which we will be modeling in the next phase of our work.

D. Estimation of the Long-Term Wage Equation

1. Preparation of the Data for Analysis

Multiple regression wage equations were run using the 15%, One in a Thousand (1/1000) State Sample of the 1970 Public Use Sample. In line with the theory underlying hedonic wage equations, the expressions were specified in semi-log form with the natural log of the hourly wage rate for an individual respondent being regressed on various characteristics of the individual. These characteristics included variables such as education, age, ethnicity, marital status, and veteran status.

Since it was believed a priori that age and years of education would be non-linearly related to the log of hourly wage, it was necessary to respecify the reported census variables in a form that would better reflect these expected non-linearities within the simple linear regression framework utilized.

The years of education variable was redefined in terms of three variables. These variables represent slope identifiers for three distinct periods of schooling. The first variable is defined to take on values ranging from zero to twelve. It represents years of education up to completion of the normal high school education. The second education variable ranges from zero to four, and represents the years of education completed by the individual in the course of a typical college education. Finally, the third education variable is a simple binary variable reflecting whether or not any post-college education was attempted. Note, the census tapes do not have any information available on major field of study. Furthermore, the public use sample we selected did not allow us to identify types of schooling undertaken. In particular, we could not uniquely identify years of attendance at vocational or technical schools.

The age variable was fitted across three intervals, representing years lived between 22 and 35, 36 and 50, and 51 to 65 years. Thus, if an individual were 44 years old, the values associated with the age variables for this person would be 13, 9, and 0, respectively. Although the intervals are somewhat arbitrary, they were chosen to reflect stages of earning power. First, there is typically a period of rapid rise in income, then a period of more modest rise, and finally a leveling off or perhaps even a decrease in wage rate. With the three age variables we have defined we hope to isolate the segments of this expected non-linear effect.

Finally, it seemed plausible that an interaction effect of education and age on wage rate might exist. That is, for example, the marginal effect of the change in years lived on the log of the hourly wage rate could depend on the level of educational attainment. In order to account for this possibility, three age-education interaction variables were defined. They were formed by multiplying the natural log of actual age by each of the three education dummy variables defined above. The log of age was used rather than simply arithmetic age in order to inject a notion of diminishing returns of aging as it affects the marginal relationship between education and wage rate; that is, it allows the effect of education on wages to decay over time.

Other variables included in the various regression specifications were binary variables. They included marital status, ethnicity, locational parameters, and veteran status. Since separate regressions were run for each of 13 different occupation classes, occupation dummy variables were not defined.

In forming the observations to be included in the analysis, several assumptions were necessary concerning the inclusion or exclusion of specific census data. There were three levels of data checks. First, we wanted observations only for male heads of household between 22 and 65 years of age inclusive for the specific occupation class under review.

Those observations meeting this first criteria were then looked at to see if they were "allocated." Allocation is a process used by census personnel to fill in missing pieces of data information. A brief description of the allocation methodology and rationale used by the Census Bureau can be found in the Public Use Sample of the 1970 Census publication (1972).

The presence of variables that are allocated can present problems for a statistically meaningful analysis. The decision facing the researcher involves weighing the trade-offs in the gains or losses associated with either retaining or eliminating observations containing relevant (i.e., used in the analysis) allocated variables. If the allocation scheme used is not a good predictor of the true, unreported data, retention of observations with allocated variables can lead to a systematic bias in the constructed data items and inconsistent coefficient estimates. In addition, it may be that the inability or unwillingness of the survey respondent to answer these questions may indicate a systematic relationship between the true answer and other included variables in the analysis. If this were true, use of this data may bias other important parameters. On the other hand, elimination of such observations results not only in a loss of degrees of freedom, but may also impart a bias to the final coefficient estimates if the distribution of observations of a particular allocated item is non-random; i.e., correlated with the other right-hand side variables. This latter occurrence implies that the sample remaining after an elimination of allocated observations is not a random sample of the general population.

With these considerations in mind, a review of the complete data file indicated that no severe statistical problems would be likely if all observations with relevant allocated items were eliminated from the final sample. A tabular breakdown of allocated items revealed only two variables with more than 2 percent allocations. These were the dummy variables for place of work and veteran status.

Rather than relying on the allocation scheme used by census personnel, we attempted our own allocation procedure that would take into account only those variables considered relevant for the present study. The idea was to predict values for the two frequently missing data items by regressing valid observations of these variables on the corresponding valid observations of all other independent variables in the system. Such regressions would necessarily be run for each of the occupation classes.

Unfortunately, the explanatory power of several trial runs of these regressions was quite low, on the order of 10-15 percent. This poor predictive behavior, coupled with the fact that a breakdown of the allocated variables by occupational category did not reveal significant departures from the distribution obtained with only "good" observations, indicated that simple elimination of the allocated items would not pose a serious problem.

Finally, after the allocation checks were made, additional checks were undertaken in order to guarantee that all responses fell within the numerical limits of that variables' code. Within this check, several recodes of variables were generated. These included:

- For the locational parameters identifying place of residence, all recorded responses greater than "1" were set equal to "1." This was done for the variables describing whether or not the respondents lived in an urban area, a metropolitan area, or a central city as these variables are defined in the questionnaire. The resulting dummy variables were defined as:

Urban:	0 = In urban aea	1 = Otherwise
Metropolitan:	0 = In SMSA	1 = Otherwise
Central City:	0 = In urban part of central city of an SMSA 1 = Otherwise	

- Marital status was recoded as a single dummy variable.

  0 = Never married   1 = All else

- The questionnaire category ranges for hours worked and weeks worked were replaced by a point estimate at the range midpoint.

- All calculated average hourly wage rates below $1.60 were deleted from the sample. This was done in an attempt to eliminate reporting and keypunch errors from the data used in the regressions.

- All incomes recorded at the questionnaire upper limit of $50,000 annually were deleted from the sample. Inclusion of these observations would seriously understate the true wage rate for these individuals.

- All self-employed workers were eliminated from the sample, due to the problem of distinguishing between labor income and returns to capital.

- Race variables were redefined as:

  1 = Black   0 = Otherwise
  1 = Other non-white   0 = Otherwise

- Veteran status variables were redefined as

  1 = Veteran   0 = Otherwise
  1 = In military during wartime   0 = Otherwise

- Workplace recode was redefined as a binary variable

  1 = Outside of SMSA   0 = Inside SMSA

- Thirteen occupation classes were defined, corresponding to major groupings in the 3-digit occupation code lists.

With the conditions as outlined above for including or excluding observations from the sample, we obtained a sample of 18,540 acceptable observations. This sample is drawn from a census "population" of 41,179 male heads of household with 16,518 records eliminated because of the various data checks and an additional 6,121 records eliminated because they contained at least one allocated variable of interest.[1] Table IV-4 gives the breakdown of the resultant sample size by occupation class.

2. Regression Results

Thirteen regression equations were specified, one for each of the occupational classes. Table IV-5 presents a glossary of the variable acronyms. Note that the criteria underlying the formation of several of these variables have been described in some detail in the previous section. The regression results for the long-term wage equations are recorded in Table IV-6. Given the semi-log nature of the specifications, the estimated coefficients can be directly interpreted as the percent change in the wage rate due to a one-unit increase of the associated variable. In general, neither the magnitudes nor the signs of the coefficients are offensive to expectations. The size of the percent changes for the ED4 variable are high for two occupation categories, but in both instances this can be traced to the small number of people with graduate level education in these occupations.

The explanatory power of the individual equations is certainly not overwhelming; however, they are typical of the levels of $\bar{R}^2$ usually

1. The 16,518 records which were eliminated in the data check include those records where the male head of household was younger then 22 or older than 65.

Table IV-4

Sample Size by Occupation Class

Occupation	Sample Size	Percent of Total
1. Professional/Technical	2,887	.156
2. Manager/Administrator	2,002	.108
3. Sales	1,206	.065
4. Clerical	1,457	.079
5. Craftsmen	3,210	.173
6. Operatives Except Transport	2,731	.147
7. Transport Operatives	1,216	.066
8. Laborers Except Farmworkers	832	.045
9. Farm Laborers	95	.005
10. Service	1,097	.059
11. Military	402	.022
12. Repair/Maintenance	1,086	.059
13. Electrical Related	319	.017
TOTAL	18,540	1.001 [a]

a. Sums to more than 1 due to rounding.

Table IV-5

Glossary of Variables

Acronym	Description
ED2	Years of education completed from 1st grade through 12th.
ED3	Number of years of college completed (i.e., 12-16 years of education).
ED4	More than 16 years of school? (YES = 1)
AGE1	Number of years lived between 22 and 35.
AGE2	Number of years lived between 36 and 50.
AGE3	Number of years lived between 51 and 65.
MARR	Marital status: 1 = Never Married, 0 = Otherwise.
RAC1	Ethnicity: 1 = Black, 0 = All Else.
VET1	Veteran status: 1 = Veteran, 0 = Otherwise.
URB	0 = Urban Area for Residence, 1 = Rural Residence.
WPL	Place of work: 1 = In Ring of or Outside SMSA, 0 = Otherwise.
WKNR	Place of work not reported: 1 = Yes, 0 = No.
AGED2	Log of actual age time's ED2 (see above).
AGED3	Log of actual age times ED3 (see above).
AGED4	Log of actual age times ED4 (see above).
LDOL	Log of hourly wage rate.

Table IV-6

Long-Term Wage Equations by Occupational Class

(Dependent Variable LDOL)

Independent Variable \ Occupation	OCC. 1 Profess/ Technical	OCC. 2 Manager/ Admin.	OCC. 3 Sales	OCC. 4 Clerical	OCC. 5 Craftsmen	OCC. 6 Operatives Except Transport	OCC. 7 Transport Operative	OCC. 8 Laborers Except Farm	OCC. 9 Farm	OCC. 10 Service	OCC. 11 Military	OCC. 12 Repair/ Maint.	OCC. 13 Electrical
CONSTANT	.734	.685	.229	.957	1.029	.673	.564	1.022	.667	.659	2.994	.700	1.099
	(.370)	(.351)	(.418)	(.236)	(.119)	(.094)	(.204)	(.204)	(.572)	(.180)	(.462)	(.174)	(.398)
ED2	.025	-.030	.193	-.023	-.057	.081	.117	-.037	.091	.051	-.513	.065	-.149
	(.161)	(.145)	(.178)	(.101)	(.048)	(.041)	(.084)	(.088)	(.254)	(.075)	(.307)	(.073)	(.179)
ED3	-.062	.038	-.185	-.043	.206	-.051	.028	-.157	.045	.357	-.591	-.105	.374
	(.075)	(.102)	(.115)	(.104)	(.124)	(.166)	(.317)	(.390)	(.790)	(.188)	(.186)	(.271)	(.301)
ED1	-.201	.528	-.546	-.371	-.132	-.499	1.828			-.354	-.341	3.521	
	(.265)	(.523)	(.817)	(.865)	(1.294)	(1.593)	(7.691)			(2.053)	(.965)	(3.370)	
AGE1	.026	.028	.050	.009	.017	.020	.028	.004	.033	.013	-.016	.025	-.0007
	(.018)	(.016)	(.020)	(.012)	(.006)	(.005)	(.010)	(.011)	(.032)	(.009)	(.037)	(.009)	(.021)
AGE2	.005	.005	.008	.0006	-.008	.003	.007	-.004	-.011	.0002	-.042	.002	-.004
	(.012)	(.010)	(.013)	(.007)	(.004)	(.003)	(.006)	(.006)	(.023)	(.006)	(.024)	(.006)	(.013)
AGE3	-.011	-.001	-.007	-.013	-.003	-.0002	-.006	-.009	.027	-.004	-.027	-.004	-.025
	(.009)	(.008)	(.010)	(.006)	(.003)	(.003)	(.006)	(.006)	(.020)	(.005)	(.050)	(.005)	(.012)
AGED2	.005	.018	-.041	.014	.021	-.016	-.028	.016	-.029	-.007	.152	-.012	.048
	(.042)	(.037)	(.047)	(.027)	(.013)	(.011)	(.023)	(.023)	(.069)	(.020)	(.088)	(.019)	(.048)
AGED3	.032	.012	.072	.019	-.046	.022	-.008	.046	.022	-.082	.204	.039	-.087
	(.021)	(.027)	(.032)	(.028)	(.034)	(.046)	(.089)	(.110)	(.221)	(.052)	.055	(.076)	(.082)
AGED4	.060	-.123	.155	.176	.064	.116	-.445			.168	.099	-.941	
	(.073)	(.148)	(.023)	(.235)	(349)	(428)	(2.347)			(.562)	(.278)	(.896)	
MARR	.066	.127	.170	.052	.125	.194	.267	.093	.182	.148	.075	.166	.210
	(.035)	(.068)	(.071)	(.045)	(.058)	(.039)	(.110)	(.101)	(.362)	(.060)	(.088)	(.092)	(.145)
RAC1	-.126	-.075	.014	-.140	-.185	-.120	-.169	-.132	-.054	-.203	-.090	-.060	-.209
	(.050)	(.075)	(.099)	(.039)	(.032)	(.024)	(.039	(.040)	(.163)	(.032)	(.074)	(.048)	(.110)
VET1	-.031	.005	-.0008	-.007	.077	.030	.045	.001	.021	-.015		-.003	.029
	(.019)	(.023)	(.029)	(.023)	(.016)	(.015)	(.027)	(.033)	(.145)	(.026)		(.025)	(.042)
URB	-.134	-.106	-.128	-.055	-.118	-.076	-.098	-.089		-.096	.026	-.073	-.063
	(.021)	(.025)	(.034)	(.026)	(.016)	(.015)	(.029)	(.035)		(.030)	(.049)	(.025)	(.040)
WPL	-.032	-.108	-.068	-.048	-.039	-.027	-.041	-.076		-.019	-.039	-.066	-.040
	(.017)	.022	(.028)	(.022)	(.016)	(.016)	(.028)	(.036)		(.025)	(.043)	(.025)	(.039)
WKNR	-.046	-.141	-.092	.017	.106	.001	-.076	.084		-.0004	-.180	.036	.057
	(.056)	(.065)	(.079)	(.069)	(.046)	(.042)	(.067)	(.080)		(.007)	(.091)	(.078)	(.120)
$\bar{R}^2$	.168	.227	.179	.103	.090	.068	.053	.060	-.030	.114	.492	.069	.115
$F(v_1, v_2)$ -	F(15, 2871)	F(15, 1986)	F(15, 1190)	F(15, 1441)	F(15, 3194)	F(15, 2715)	F(15, 1200)	F(13, 818)	F(10, 84)	F(15, 108)	F(14, 447)	F(15, 1070)	F(13, 305)
F-Statistic	39.95	40.25	18.48	12.09	22.21	14.24	5.53	4.01	.725	10.44	32.84	6.38	4.17
EXP(Mean of LDOL)	5.34	5.23	4.35	3.98	4.22	3.61	3.55	3.[illegible]	2.92	3.21	3.50	3.69	4.49

obtained in hedonic wage equations. The relatively high $\bar{R}^2$ for the military occupation class is likely due to the more structured orientation of military pay scales vís-a-vís the age and educational backgrounds of military service personnel. Note, however, that the equation F-statistics, which test the hypothesis that <u>none</u> of the explanatory variables has an influence on the mean of the logged wage rate, is rejected at the 99 percent significance level for all occupations except farm workers.

Although the statistical significance of the individual paramaters is not an issue for present purposes, it should be pointed out that the appropriate test of significance for the education and age variables involves structuring the null hypothesis so as to include both the untransformed variable of interest as well as the interaction variable. Consequently, one cannot discern statistical significance or insignificance for age and education directly from Table IV-6.

One aspect of some interest is the profiles for age and education that are indicated by the regression results. Table IV-7 is a listing of derived values representing the predicted wage rates for different levels of education and age. For purposes of presentation, we have limited the table to three occupation classes. The occupations chosen represent a range of types of occupations and give a reasonable picture of typical results for the different occupation classes. Figures IV-1a, IV-2a, and IV-3a are the graphical portrayals of the information contained in Table IV-7 when age is taken as given, while Figures IV-1b, IV-2b, and IV-3b use the data in Table IV-7 conditional on a given education level. In predicting the wage rates from these profiles, all other variables in the system apart from age and education are assumed to take a zero value. Consequently, any comparisons across occupations are valid only for the

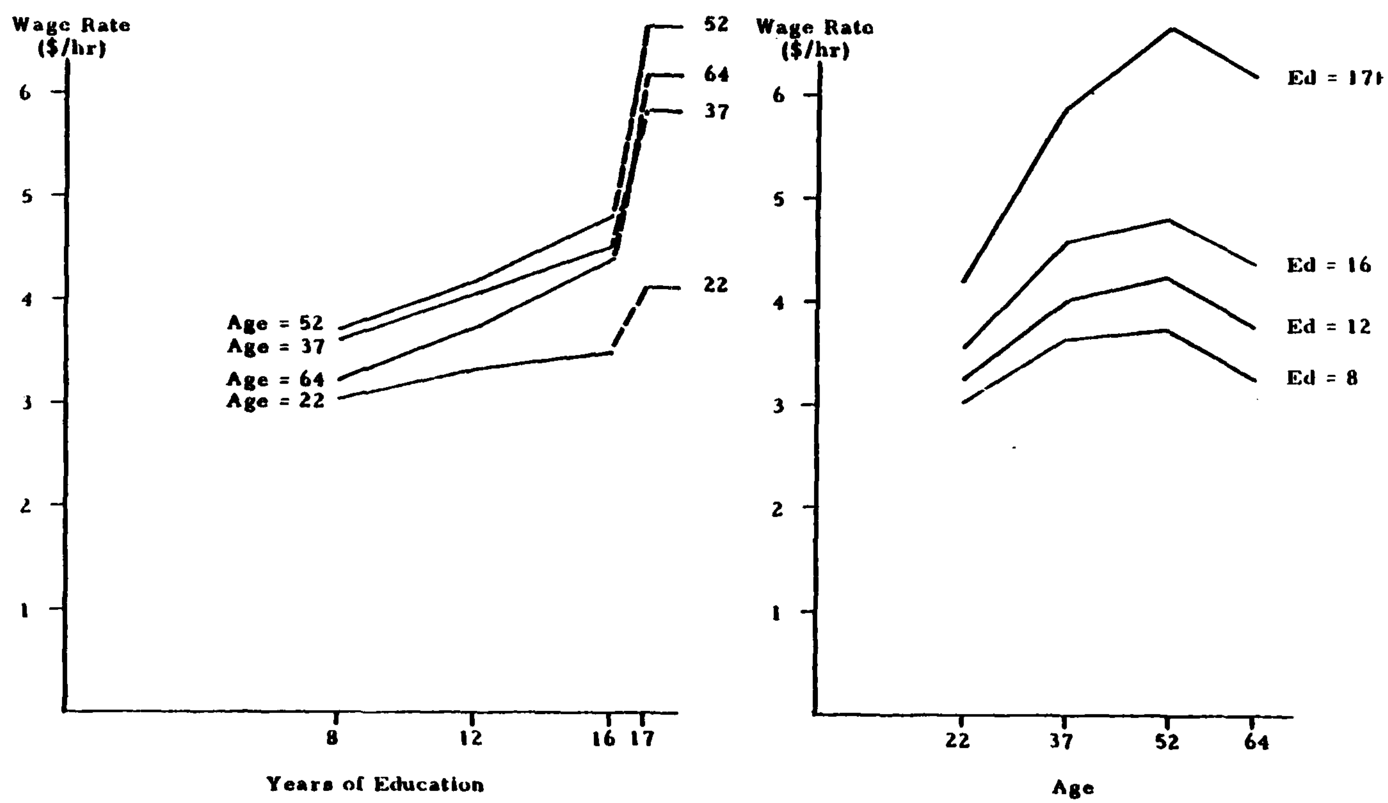
Table IV-7

Predicted Wage Rates for Given Levels of Age and Education ($/hr)

Occupation	Age	Years of Education			
		8	12	16	17
1. Clerical Workers	22	$3.06	$3.32	$3.54	$4.20
	37	3.65	4.08	4.52	5.89
	52	3.73	4.24	4.82	6.67
	64	3.26	3.76	4.34	6.23
2. Professional/Technical	22	2.88	3.39	3.92	3.86
	37	4.16	4.95	6.13	6.22
	52	4.41	5.27	6.82	7.07
	64	3.89	4.68	6.21	6.52
3. Military	22	2.12	1.78	2.09	2.02
	37	2.98	3.44	6.16	6.26
	52	2.47	3.51	8.30	8.73
	64	2.30	3.71	10.39	11.15

Figure IV-1

Education and Age Profiles for Clerical Workers

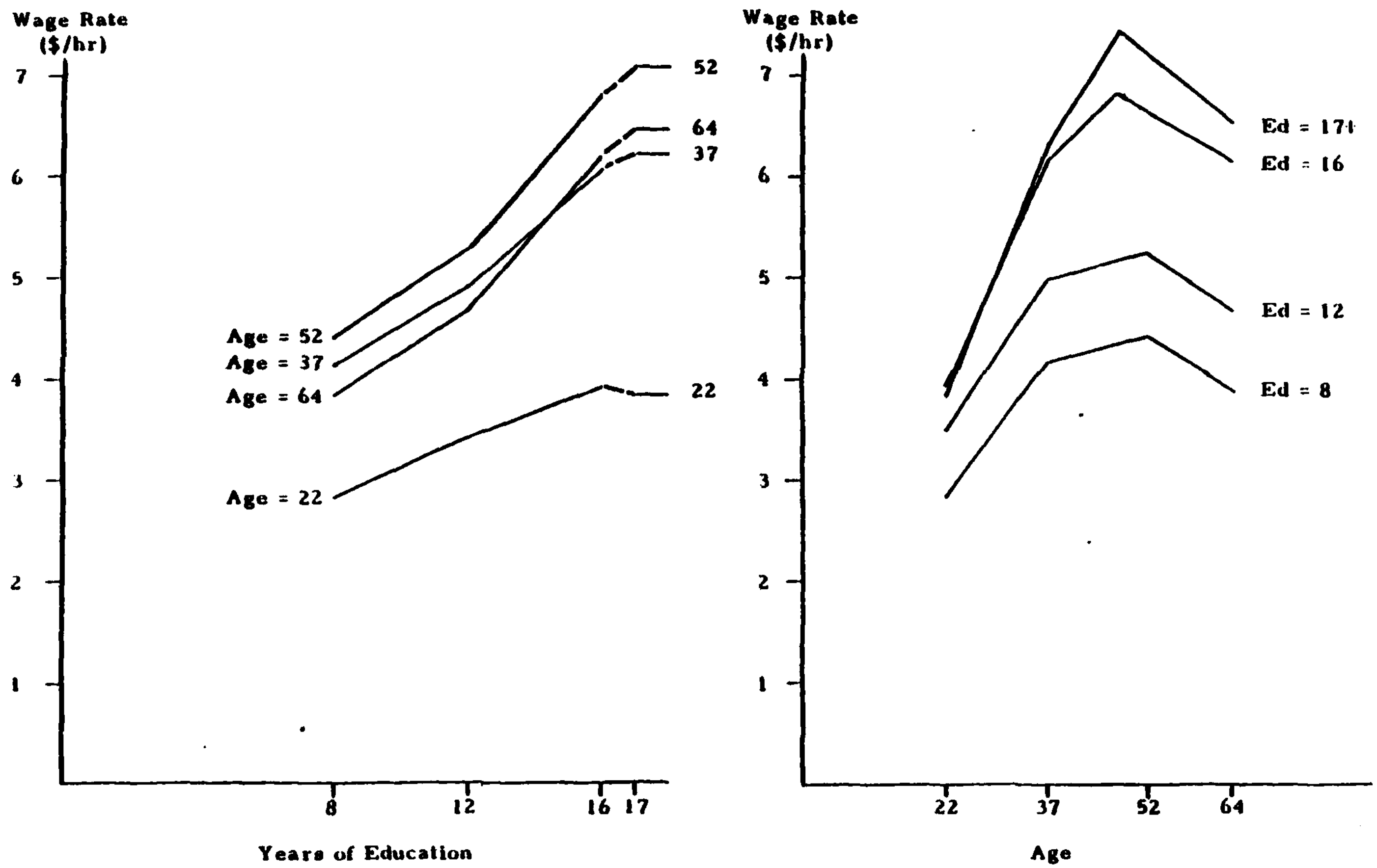


a) Education Profile Given Age

b) Age Profile Given Education

Figure IV-2

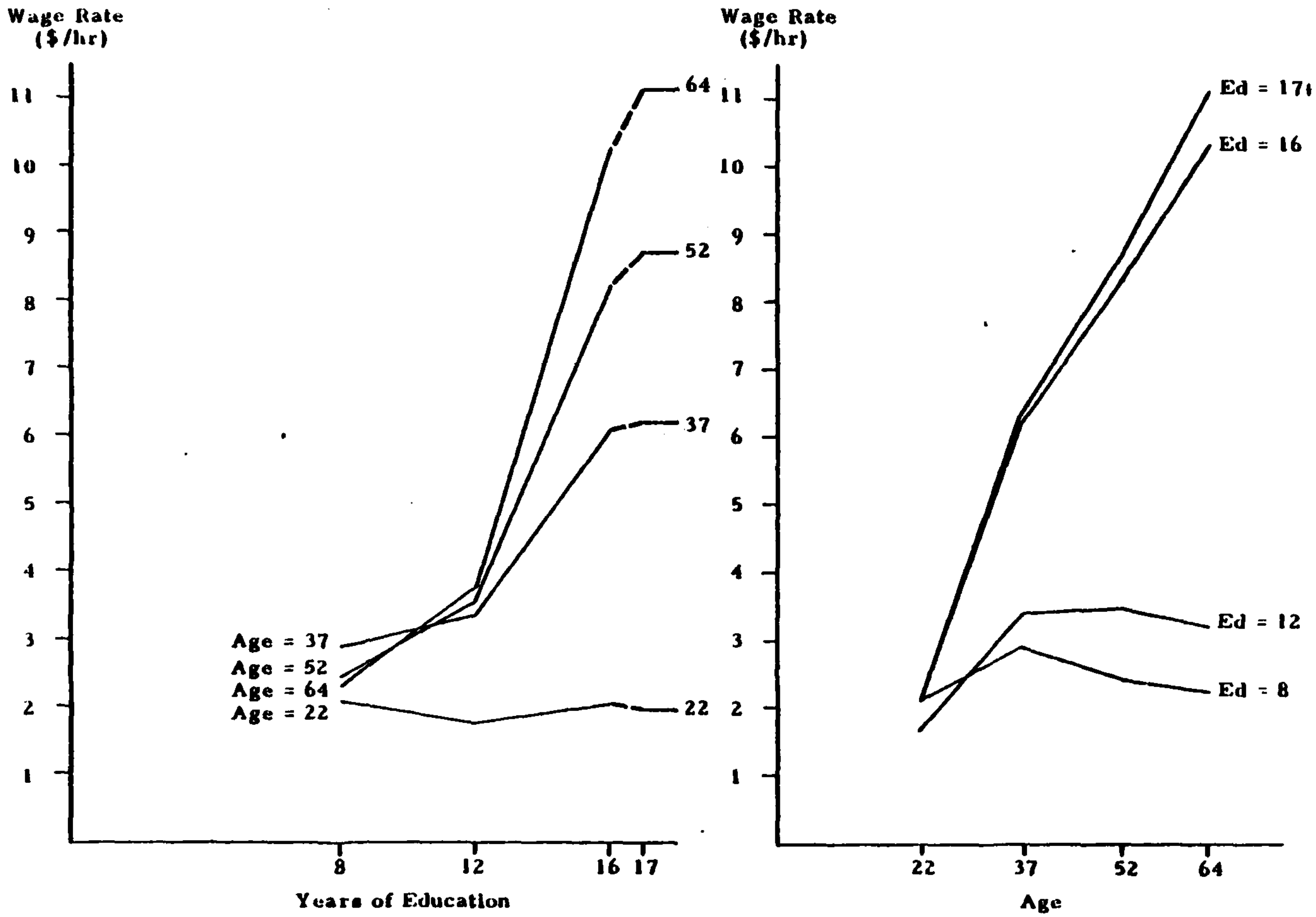
Education and Age Profiles for Professional/Technical Workers



a) Education Profile Given Age b) Age Profile Given Education

Figure IV-3

Education and Age Profiles for the Military



a) Education Profile Given Age

b) Age Profile Given Education

combination of assumed parameter values. With respect to Figures IV-1b, IV-2b, and IV-3b, note also that no growth factor has been introduced to reflect inherent wage increases associated with productivity gains over time. The figures simply portray expected wage profiles with all other variables held constant except education (Figures IV-1a, IV-2a, IV-3a) or age (Figures IV-1b, IV-2b, IV-3b).

The orientations of the various profiles tend to be supportive of our *a priori* beliefs concerning the relation between wages and education and age. In general, the figures show that both a high school education relative to an 8th grade education as well as four years of college relative to a high school diploma contribute positively to wage rate. Some puzzles do exist, however; for example, graduate level work yields impressive wage rate gains for someone in a clerical occupation, while more modest gains (and even a small loss for younger employees) are realized in the Professional/Technical category.

Perhaps the most interesting of the figures is Figure IV-3b, the Military Age Profile. For those people contemplating the military as a career, the figure indicates that if they have no more than a high school diploma, they can expect only modest increases in base wage rate after their first several years in the service. On the other hand, an investment of time and money in four years of college can contribute to impressive gains in the expected wage rate over time in a military career. This implies that any effort to model successfully a military occupational choice question should also address the issue of educational choice and possible delayed entry into the job market.

E. Conclusions

The results reported in this chapter are not particularly interesting when looked at in isolation. Instead, the regression expressions become useful and usable in the context of the occupational choice question. Our purpose here has been solely to devise a means whereby present and future expected wage rates could be ascertained for various occupational classes.

V. ESTIMATION OF NON-MILITARY JOB CHARACTERISTICS

A. Introduction

The focus of this chapter is on the non-pecuniary job characteristics that may influence an individual's decision to choose one occupation over another. The Dictionary of Occupational Titles (hereafter DOT) and the Public Use Samples of Basic Records from the 1970 Census (hereafter 1970 Census) yield data which allow us to identify a wide range of such job characteristics and to aggregate them to the broad wage and salary occupational classes we have defined.

The following two sections describe the specific job characteristics and the procedure used for aggregating the data to the appropriate level. In the final section the estimated values for the job characteristics are tabulated.

B. Description of Variables

The DOT data provide a comprehensive listing of occupational information which was designed to aid in job placement and career counseling, but which proved to be readily adaptable to our purposes. For each of the more than 12,000 occupational titles included on the DOT tape, there are variables on the nature of the work performed and the physical and mental demands of the work activities. Variables too specific to represent a broad category of occupations were eliminated. For example, each DOT occupation is assigned, from a list of one hundred possibilities, one to four work fields which describe specific methods of performing a given job, such as "loading-moving" or "hunting-fishing." Such a wide range of possible ratings precluded our use of these variables. However, some of the information they convey is captured by other variables. For example, variables concerning the physical demands associated with each job give us general information on the nature of the job. One would expect that a job assigned the work field "loading-moving" would require a considerable amount of strength on the part of the worker.

The job characteristic variables can be grouped under seven headings: worker functions, training time, aptitudes, temperaments, interests, physical demands, and environmental conditions. Each of these categories requires further explanation.[1]

1. Worker Functions

Worker functions refer to the relationship of the worker to data, people, and things and are coded in two ways. First, the functioning of the worker in relationship to each of these three fields is determined to be either significant or not significant. Secondly, the level of complexity of the worker's involvement with each is estimated. Table V-1 shows the structure of worker functions. It is a hierarchical arrangement, and the job is assigned that rating which "expresses the total level of complexity of the job-worker situation."[2] Lower values imply a higher level of complexity.

---

1. The U.S. Department of Labor Manpower Administration Handbook for Analyzing Jobs (Washington, D.C.: U.S. Government Printing Office, 1972) will provide the interested reader with a more detailed description of each of the job characteristics variables.

2. Ibid., p. 4.

Table V-1

Structure of Worker Functions

DATA	PEOPLE	THINGS
0 Synthesizing	0 Mentoring	0 Setting Up
1 Coordinating	1 Negotiating	1 Precision Working
2 Analyzing	2 Instructing	2 Operating-Controlling
3 Compiling	3 Supervising	3 Driving-Operating
4 Computing	4 Diverting	4 Manipulating
5 Copying	5 Persuading	5 Tending
6 Comparing	6 Speaking-Signaling	6 Feeding-Offbearing
	7 Serving	7 Handling
	8 Taking Instructions-Helping	

NOTE: The hyphenated factors Speaking-Signaling, Taking Instructions-Helping, Operating-Controlling, Driving-Operating, and Feeding-Offbearing are single functions.

Source: U.S. Department of Labor Manpower Administration, Handbook for Analyzing Jobs (Washington, D.C.: U.S. Government Printing Office, 1972), p. 5.

2. Training Time

Training time refers to both the general level of educational development and the specific vocational training that are required for average performance of a specific job. The general educational development scale is divided into three parts, representing independent estimates of average reasoning development, mathematical development, and language development. The specific ratings are described in Table V-2. On this scale, a higher value implies that a higher level of development is required.

Specific vocational training includes vocational education, apprenticeship training, in-plant training, on-the-job training, and essential experience in other jobs. This variable attempts to exclude the kind of education accounted for in the general educational development scale. Table V-3 shows how it has been specified.

3. Aptitudes

Aptitudes are defined as "the specific capacities or abilities required of an individual in order to facilitate the learning of some task or job duty."[3] Eleven aptitudes are identified for each job and assigned a rating from one to five. They reflect amounts of the aptitudes possessed in relation to the working population at large. That is, a rating of one implies that this group of workers possesses a very high degree of the aptitude (specifically, is in the top ten percent of the population), whereas a rating of five implies that this group possesses the aptitude to a very small degree (specifically, is in the bottom ten percent of the population). The aptitudes are listed in Table V-4.

---

3. Ibid., p. 233.

Table V-2

Scale of General Education Development (GED)

LEVEL	REASONING DEVELOPMENT	MATHEMATICAL DEVELOPMENT	LANGUAGE DEVELOPMENT
6	Apply principles of logical or scientific thinking to a wide range of intellectual and practical problems. Deal with nonverbal symbolism (formulas, scientific equations, graphs, musical notes, etc.) in its most difficult phases. Deal with a variety of abstract and concrete variables. Apprehend the most abstruse classes of concepts.	Advanced calculus: Work with limits, continuity, real number systems, mean value theorems, and implicit function theorems. *Modern algebra:* Apply fundamental concepts of theories of groups, rings, and fields. Work with differential equations, linear algebra, infinite series, advanced operations methods, and functions of real and complex variables. Statistics: Work with mathematical statistics, mathematical probability and applications, experimental design, statistical inference, and econometrics.	*Reading:* Read literature, book and play reviews, scientific and technical journals, abstracts, financial reports, and legal documents. *Writing:* Write novels, plays, editorials, journals, speeches, manuals, critiques, poetry, and songs. Speaking: Conversant in the theory, principles, and methods of effective and persuasive speaking, voice and diction, phonetics, and discussion and debate.
5	Apply principles of logical or scientific thinking to define problems, collect data, establish facts, and draw valid conclusions. Interpret an extensive variety of technical instructions in mathematical or diagrammatic *form*. Deal with several abstract and concrete variables.	Algebra: Work with exponents and logarithms, linear equations, quadratic equations, mathematical induction and binomial theorem, and permutations. Calculus: Apply concepts of analytic geometry, differentiations and integration of algebraic functions with applications. Statistics: Apply mathematical operations to frequency *distributions*, reliability and validity of tests, normal curve, analysis of variance, correlation techniques, chi-square application and sampling theory, and factor analysis.	Same as Level 6.
4	Apply principles of rational systems[1] to *solve practical problems* and deal with a variety of concrete variables in situations *where only limited standardization exists.* Interpret a variety of instructions furnished in written, oral, diagrammatic, or schedule form.	Algebra: Deal with system of real numbers; linear, quadratic, rational, exponential, logarithmic, angle and circular functions, and inverse functions; related algebraic solution of equations and inequalities; limits and continuity, and probability and statistical inference. Geometry: Deductive axiomatic geometry, plane and solid; and rectangular coordinates. Shop Math: Practical application of fractions, percentages, ratio and proportion, mensuration, logarithms, slide rule, practical algebra, geometric construction, and essentials of trigonometry.	Reading: Read novels, poems, newspapers, periodicals, journals, manuals, dictionaries, thesauruses, *and encyclopedias*. Writing: Prepare business letters, expositions, summaries, and reports, using prescribed format and conforming to all rules of punctuation, grammar, diction, and style. Speaking: Participate in panel discussions, dramatizations, and debates. Speak extemporaneously on a variety of subjects.

Table V-2 (Continued)

LEVEL	REASONING DEVELOPMENT	MATHEMATICAL DEVELOPMENT	LANGUAGE DEVELOPMENT
3	Apply commonsense understanding to carry out instructions furnished in written, oral, or diagrammatic form. Deal with problems involving several concrete variables in or from standardized situations.	Compute discount, interest, profit, and loss; commission, markup, and selling price; ratio and proportion, and percentage. Calculate surfaces, volumes, weights, and measures. Algebra: Calculate variables and formulas; monomials and polynomials; ratio and proportion variables; and square roots and radicals. Geometry: Calculate plane and solid figures; circumference, area, and volume. Understand kinds of angles, and properties of pairs of angles.	Reading: Read a variety of novels, magazines, atlases, and encyclopedias. Read safety rules, instructions in the use and maintenance of shop tools and equipment, and methods and procedures in mechanical drawing and layout work. Writing: Write reports and essays with proper format, punctuation, spelling, and grammar, using all parts of speech. Speaking: Speak before an audience with poise, voice control, and confidence, using correct English and well-modulated voice.
2	Apply commonsense understanding to carry out detailed but uninvolved written or oral instructions. Deal with problems involving a few concrete variables in or from standardized situations.	Add, subtract, multiply, and divide all units of measure. Perform the four operations with like common and decimal fractions. Compute ratio, rate, and percent. Draw and interpret bar graphs. Perform arithmetic operations involving all American monetary units.	Reading: Passive vocabulary of 5,000-6,000 words. Read at rate of 190-215 words per minute. Read adventure stories and comic books, looking up unfamiliar words in dictionary for meaning, spelling, and pronunciation. Read instructions for assembling model cars and airplanes. Writing: Write compound and complex sentences, using cursive style, proper end punctuation, and employing adjectives and adverbs. Speaking: Speak clearly and distinctly with appropriate pauses and emphasis, correct pronunciation, variations in word order, using present, perfect, and future tenses.
1	Apply commonsense understanding to carry out simple one- or two-step instructions. Deal with standardized situations with occasional or no variables in or from these situations encountered on the job.	Add and subtract two digit numbers. Multiply and divide 10's and 100's by 2, 3, 4, 5. Perform the four basic arithmetic operations with coins as part of a dollar. Perform operations with units such as cup, pint, and quart; inch, foot, and yard; and ounce and pound.	Reading: Recognize meaning of 2,500 (two- or three-syllable) words. Read at rate of 95-120 words per minute. Compare similarities and differences between words and between series of numbers. Writing: Print simple sentences containing subject, verb, and object, and series of numbers, names, and addresses. Speaking: Speak simple sentences, using normal word order, and present and past tenses.

[1] Examples of rational systems are: bookkeeping, internal combustion engines, electric wiring systems, house building, nursing, farm management, and navigation.

Source: U.S. Department of Labor Manpower Administration, Handbook for Analyzing Jobs (Washington, D.C.: U.S. Government Printing Office, 1972), pp. 210-211.

Table V-3

Scale of Specific Vocational Preparation (SVP)

Level	
1	Short demonstration only
2	Anything beyond short demonstration up to and including 30 days
3	Over 30 days up to and including 3 months
4	Over 3 months up to and including 6 months
5	Over 6 months up to and including 1 year
6	Over 1 year up to and including 2 years
7	Over 2 years up to and including 4 years
8	Over 4 years up to and including 10 years
9	Over 10 years

Source: U.S. Department of Labor Manpower Administration, <u>Handbook for Analyzing Jobs</u> (Washington, D.C.: U.S. Government Printing Office, 1972), p. 220.

Table V-4

APTITUDES are the specific capacities or abilities required of an individual in order to facilitate the learning of some task or job duty. Following are the aptitudes included in this component:
G Intelligence
V Verbal
N Numerical
S Spatial
P Form Perception
Q Clerical Perception
K Motor Coordination
F Finger Dexterity
M Manual Dexterity
E Eye-Hand-Foot Coordination
C Color Discrimination

Source: U.S. Department of Labor Manpower Administration, Handbook for Analyzing Jobs (Washington, D.C.: U.S. Government Printing Office, 1972), p. 8.

The next two types of variables -- temperaments and interests -- differ from those described above in that they are binary in nature; i.e., the specific temperaments and interests are either associated with a job or not associated with it.

4. Temperaments

Temperaments are listed in Table V-5. They are generally described as adaptability requirements made on the worker in specific occupational situations and were incorporated in job analysis in order to differentiate jobs according to different personal traits that may be an indicator of success in the job.

5. Interests

There are five possible bipolar interest factors on which each occupation can receive a rating. Table V-6 specifies these factors. In establishing interest factors for a job, those pairs that were significant on the job were selected and then either (a) or (b) was chosen.

6. Physical Demands

Within the physical demands category, all occupations were rated on the basis of the amount of strength required by workers for average performance. At one extreme, sedentary work involves lifting of a maximum of ten pounds; whereas, at the other extreme, very heavy work entails lifting objects in excess of 100 pounds. In addition, the presence of any or all of five other physical demands was noted. This information is summarized in Table V-7.

Table V-5

TEMPERAMENTS for the purpose of collecting occupational data, are defined as "personal traits" required of a worker by specific job-worker situations. This component consists of the following 10 factors:

DCP	Adaptability to accepting responsibility for the direction, control, or planning of an activity.
FIF	Adaptability to situations involving the interpretation of feelings, ideas, or facts in terms of personal viewpoint.
INFLU	Adaptability to influencing people in their opinions, attitudes, or judgments about ideas or things.
SJC	Adaptability to making generalization, evaluations, or decisions based on sensory or judgmental criteria.
MVC	Adaptability to making generalizations, evaluations, or decisions based on measurable or verifiable criteria.
DEPL	Adaptability to dealing with people beyond giving and receiving instructions.
REPCON	Adaptability to performing repetitive work, or to performing continuously the same work, according to set procedures, sequence, or pace.
PUS	Adaptability to performing under stress when confronted with emergency, critical, unusual, or dangerous situations; or situations in which working speed and sustained attention are make-or-break aspects of the job.
STS	Adaptability to situations requiring the precise attainment of set limits, tolerances, or standards.
VARCH	Adaptability to performing a variety of duties, often changing from one task to another of a different nature without loss of efficiency or composure.

Source: U.S. Department of Labor Manpower Administration, Handbook for Analyzing Jobs (Washington, D.C.: U.S. Government Printing Office, 1972), pp. 8-9.

Table V-6

INTERESTS

1a.	A preference for activities dealing with things and objects.	*vs* 1b.	A preference for activities concerned with the communication of data.*
2a.	A preference for activities involving business contact with people.	*vs* 2b.	A preference for activities of a scientific and technical nature.
3a.	A preference for activities of a routine, concrete, organized nature.	*vs* 3b.	A preference for activities of an abstract and creative nature.
4a.	A preference for working for the presumed good of people.	*vs* 4b.	A preference for activities that are carried on in relation to processes, machines, and techniques.
5a.	A preference for activities resulting in prestige or the esteem of others.	*vs* 5b.	A preference for activities resulting in tangible, productive satisfaction.

* Involvement with live animals is to be identified with "people and the communication of data" wherever an animal is dealt with on an individual basis.

Source: U.S. Department of Labor Manpower Administration, Handbook for Analyzing Jobs (Washington, D.C.: U. S. Government Printing Office, 1972), p.9.

Table V-7

PHYSICAL DEMANDS

1. Strength (lifting, carrying, pushing, and or pulling)
   - Sedentary work
   - Light work
   - Medium work
   - Heavy work
   - Very heavy work
2. Climbing and/or balancing
3. Stooping, kneeling, crouching, and or crawling
4. Reaching, handling, fingering, and or feeling
5. Talking and/or hearing
6. Seeing

Source: U.S. Department of Labor Manpower Administration, <u>Handbook for Analyzing Jobs</u> (Washington, D.C.: U.S. Government Printing Office, 1972), p. 9.

7. <u>Environmental Conditions</u>

The environmental conditions variables were specified in a similar way. In each occupation, it was determined on average whether workers tended to spend most of their time inside, outside, or approximately half of their time inside and half outside. Then, the presence of any or all of six other physical surroundings variables was recorded. These are listed in Table V-8.

In summary, each occupational title included in the DOT received ratings on three worker functions (coded in two ways), three general educational development factors, a specific vocational training factor, eleven aptitudes, the amount of strength required, and the work location. In addition, certain temperaments, interests, other physical demands, and other environmental conditions were determined to be either significant or not significant for the occupation.

C. Aggregation of the Data

The final objective of this task is the development of occupational characteristic values for twelve of the major Census occupational categories that were used in the estimation of the wage equations.[4] The aggregation procedure involved two considerations. First, a simple average of all jobs listed in a specific occupation was deemed undesirable, since this would mean assignment of equal weight to all jobs in the category and neglect of the distribution of the population among the various jobs in the category. Secondly, we wanted to reflect the time trend in the values of job characteristics. We believe this is related to the age composition of

4. The Dictionary of Occupational Titles covers only civilian occupations and so occupational characteristic data cannot be compiled for military jobs or for education.

Table V-8

ENVIRONMENTAL CONDITIONS

1. Work location (inside, outside, or both)
2. Extreme cold with or without temperature changes
3. Extreme heat with or without temperature changes
4. Wet and/or humid
5. Noise and/or vibration
6. Hazards
7. Atmospheric conditions

Source: U.S. Department of Labor Manpower Administration, Handbook for Analyzing Jobs (Washington, D.C.: U.S. Government Printing Office, 1972), p. 10.

a given job. Thus, if there were two occupations within a major occupational category in which the number of persons was the same, the values of the job characteristics for each would not receive equal weight in the aggregation unless the age distributions of the two occupations were also identical. In general, it would be desirable to give greater weight to the one in which the age distribution was more skewed to the young.

These considerations can be formalized as follows: Let Q_{ijk} be the value of the ith job characteristic in occupation j within occupational category k, C_{jmk} be the number of persons in age group m and occupation j within the kth category, and W_m be the weight assigned to the mth age group. Then, we would estimate the value of the ith job characteristic in the kth occupational class, JC_{ik}, to be:

$$(1) \qquad JC_{ik} = \frac{\sum_{j=1}^{n_k} \sum_{m=1}^{n} Q_{ijk} C_{jmk} W_m}{\sum_{j=1}^{n_k} \sum_{m=1}^{n} C_{jmk} W_m}$$

where n_k is the number of occupations in the category and n is the number of age groups into which the population has been divided.

There were two difficulties in implementing this approach. First, we had no data on the distribution of the population among the more than 12,000 DOT occupational titles. Thus, our only course was to arrange them according to the 3-digit Census codes in which they belong (this information is available from the DOT), and then to obtain average values for the characteristics within each 3-digit code. Secondly, while the 1970

Census provides information on the number of persons in each of five age groups within each 3-digit occupation, the assignment of weights was necessarily arbitrary. Guided by the intuition that the weights should decline as age increases, the following scheme was devised:

Age Group (m)	Weight (W_m)
Age $\leq$ 30 years	1/2
31 years $\leq$ Age $\leq$ 40 years	1/4
41 years $\leq$ Age $\leq$ 50 years	1/8
51 years $\leq$ Age $\leq$ 60 years	1/16
Age $\geq$ 61 years	1/16

The final problem was to account for those few cases where data on job characteristics were not available. In such instances, the missing values were assigned the values of the characteristics obtained by calculation of equation (1) with occupations with missing data deleted. The fraction of cases in which this occurred is given in Table V-9.

D. The Results

Tables V-10 through V-16 show the values assigned to the various job characteristics. There are two types of variables which must be distinguished. The first are continuous variables. The value assigned is the weighted average of the job characteristic values of the individual occupations. The second are discrete variables. The value assigned is the weighted proportion of occupations within the broad occupational category for which the particular job characteristic was considered important. Table V-10 illustrates both types. The values in the first row of the

Table V-9

Percentage of Cases Where Job Characteristic

Values Were Missing

	Age Group (Years)				
Occupation	30 or Less	31 to 40	41 to 50	51 to 60	61 or Over
1	9.38	7.27	6.01	6.01	6.52
2	0.0	0.0	0.0	0.0	0.0
3	4.76	7.34	5.43	6.36	7.18
4	14.26	11.07	10.65	8.17	10.59
5	15.23	9.16	8.46	6.95	10.75
6	8.22	7.73	5.87	5.04	8.02
7	0.0	0.0	0.0	0.0	0.0
8	14.88	16.44	16.45	17.23	15.62
9	22.05	8.18	6.37	8.56	9.31
10	15.57	9.87	8.90	7.16	9.35
12	2.93	0.69	0.43	0.28	0.41
13	0.0	0.0	0.0	0.0	0.0

table can be interpreted as the weighted average value assigned to the worker function variable "data" in each of the twelve occupational categories. The possible range for this variable is 0 to 6 (see Table V-1). The values in the second row can be interpreted as the weighted proportion of occupations in each category in which the worker function "data" was considered significant by the rater. All variables of the second type are noted in the tables by the letter "p" beside the variable name.

While it was the availability of data which dictated the actual choice of job characteristic variables, examination of these tables indicates that the data do cover a rather wide spectrum. In addition, the variation in the values of some of the variables is quite marked across occupational categories. For example, the variable specific vocational preparation (Table V-11, line 4) ranges from a high of 7.24 in occupation 1 (implying that from two to four years of training are required on average) to a low of 2.74 in occupation 8 (implying that an average training time is less than three months). Or for example, interest 2 (Table V-14, lines 3 and 4) was highly significant for occupation 3, with approximately 95 percent of the occupations in the sales category being rated as displaying a preference for activities involving contacts with people over actiities of a scientific or technical nature, whereas in only about 1 percent of the occupations in category 6 was this variable even considered significant. As a final example, we note the variation in the strength variables (Table V-15, lines 1 through 5). In occupations 1, 2, 3, 4, and 13 most jobs were rated as being "light" or "sedentary," whereas in all other categories most were rated as varying from "medium" to "very heavy." A more careful examination of the tables will reveal that no values are truly contrary to expectations and that most are actually quite close to *a priori* beliefs.

In the final estimation of the utility function, a subset of these characteristics will be selected as representing factors by which occupations can be distinguished.

Table V-10

Worker Functions

Occupation / Job Characteristic	OCC 1 Profess./ Tech. & Kindred	OCC 2 Mgrs. & Admin. Exc. Farm	OCC 3 Sales Workers	OCC 4 Clerical & Kindred Workers	OCC 5 Craftsmen & Kindred Workers	OCC 6 Operatives Except Transp.	OCC 7 Transport Equip. Operatives	OCC 8 Laborers Exc. Farm	OCC 9 Farm Related Workers	OCC 10 Service, Exc. Priv. Household	OCC 12 Repairmen & Maint.	OCC 13 Electrical Related Employmt.
DATA	1.27	1.12	2.72	3.43	3.12	5.34	5.16	5.67	3.66	4.57	3.03	2.52
DATA-SIGNIFICANT[P]	.99	1.00	.98	.93	.63	.16	.22	.15	.48	.30	.71	.86
PEOPLE	4.64	4.56	5.19	6.40	6.25	7.70	7.14	7.65	6.39	6.68	7.31	6.49
PEOPLE-SIGNIFICANT[P]	.56	.97	.95	.38	.30	.04	.17	.06	.10	.46	.04	.19
THINGS	4.57	6.78	6.60	5.35	2.24	4.53	3.65	5.73	3.78	5.74	1.68	.178
THINGS-SIGNIFICANT[P]	.37	.03	.16	.36	.95	.99	.93	.97	.89	.60	.99	.91

See Table V-1 for explanation of worker function variables.

Table V-11

Training Time

Occupation / Job Characteristic	OCC 1 Profess./ Tech. & Kindred	OCC 2 Mgrs. & Admin. Exc. Farm	OCC 3 Sales Workers	OCC 4 Clerical & Kindred Workers	OCC 5 Craftsmen & Kindred Workers	OCC 6 Operatives Except Transp.	OCC 7 Transport Equip. Operatives	OCC 8 Laborers Exc. Farm	OCC 9 Farm Related Workers	OCC 10 Service, Exc. Priv. Household	OCC 12 Repairmen & Maint.	OCC 13 Electrical Related Employmt.
GENERAL EDUCATIONAL DEVELOPMENT:												
REASONING	4.99	4.52	3.70	3.32	3.45	2.41	2.61	2.13	2.89	2.61	3.37	3.82
MATHEMATICAL	4.15	3.76	2.82	2.42	2.49	1.58	1.49	1.26	2.19	1.69	2.52	3.32
LANGUAGE	4.71	4.11	3.57	2.75	2.49	1.60	1.86	1.45	2.44	2.17	2.73	3.13
SPECIFIC VOCATIONAL TRAINING	7.24	7.17	4.58	4.26	6.05	3.37	3.23	2.74	4.42	3.53	5.78	6.80

See Tables V-2 and V-3 for explanation of training time variables.

Table V-12

Aptitudes

Job Characteristic \ Occupation	OCC 1 Profess./ Tech. & Kindred	OCC 2 Mgrs. & Admin. Exc. Farm	OCC 3 Sales Workers	OCC 4 Clerical & Kindred Workers	OCC 5 Craftsmen & Kindred Workers	OCC 6 Operatives Except Transp.	OCC 7 Transport Equip. Operatives	OCC 8 Laborers Exc. Farm	OCC 9 Farm Related Workers	OCC 10 Service, Exc. Priv. Household	OCC 12 Repairmen & Maint.	OCC 13 Electrical Related Employmt.
INTELLIGENCE	1.85	2.14	2.83	2.96	3.03	3.55	3.34	3.69	3.18	3.44	3.05	2.69
VERBAL	1.92	2.20	2.68	3.06	3.33	3.86	3.86	3.89	3.52	3.56	3.40	3.00
NUMERICAL	2.33	2.66	3.13	3.27	3.36	3.96	3.91	4.18	3.86	3.92	3.49	2.94
SPATIAL	2.88	3.47	3.68	3.86	3.03	3.64	3.21	3.78	3.51	3.79	2.89	2.67
FORM PERCEPTION	2.87	3.45	3.50	3.57	3.03	3.49	3.86	3.69	3.57	3.72	2.90	2.54
CLERICAL PERCEPTION	2.78	2.97	3.21	2.69	3.85	4.20	4.01	4.40	4.09	4.02	3.97	3.67
MOTOR COORDINATION	3.47	3.95	3.71	3.45	3.22	3.44	3.21	3.69	3.53	3.63	3.19	3.16
FINGER DEXTERITY	3.41	3.98	3.76	3.54	3.32	3.59	3.99	3.77	3.73	3.84	3.15	3.05
MANUAL DEXTERITY	3.50	3.96	3.77	3.60	2.98	3.12	3.16	3.26	3.19	3.32	2.77	2.74
EYE-HAND-FOOT COORDINATION	4.63	4.86	4.64	4.94	4.33	4.73	3.27	4.48	4.18	4.29	4.50	4.28
COLOR DISCRIMINATION	4.21	4.58	4.38	4.66	4.35	4.61	4.39	4.54	4.16	4.51	4.38	3.87

Table V-13

Temperaments

Occupation / Job Characteristic	OCC 1 Profess./ Tech. & Kindred	OCC 2 Mgrs. & Admin. Exc. Farm	OCC 3 Sales Workers	OCC 4 Clerical & Kindred Workers	OCC 5 Craftsmen & Kindred Workers	OCC 6 Operatives Except Transp.	OCC 7 Transport Equip. Operatives	OCC 8 Laborers Exc. Farm	OCC 9 Farm Related Workers	OCC 10 Service, Exc. Priv. Household	OCC 12 Repairmen & Maint.	OCC 13 Electrical Related Employmt.
DCP^P	.524	.851	.074	.135	.291	.009	.038	.028	.368	.145	.046	.210
FIF^P	.082	.026	.016	.000	.004	.000	.0	.0	.010	.025	.004	.0
$INFLU^P$	.242	.154	.836	.022	.003	.0	.030	.0	.005	.011	.0	.0
SJC^P	.494	.536	.631	.107	.095	.062	.100	.040	.396	.281	.135	.170
MVC^P	.651	.448	.202	.307	.677	.338	.149	.080	.364	.114	.778	.928
$DEPL^P$	.533	.896	.939	.407	.288	.038	.191	.032	.061	.449	.061	.201
$REPCON^P$	.018	.020	.012	.411	.205	.650	.690	.810	.333	.464	.168	.083
PUS^P	.022	.009	.0	.019	.018	.004	.061	.008	.010	.121	.002	.037
STS^P	.436	.130	.073	.567	.814	.647	.147	.369	.250	.161	.880	.871
$VARCH^P$	.382	.520	.060	.245	.386	.059	.057	.133	.516	.338	.434	.657

See Table V-5 for explanation of temperament variables.

Table V-14

Interests

Occupation / Job Characteristic	OCC 1 Profess./ Tech. & Kindred	OCC 2 Mgrs. & Admin. Exc. Farm	OCC 3 Sales Workers	OCC 4 Clerical & Kindred Workers	OCC 5 Craftsmen & Kindred Workers	OCC 6 Operatives Except Transp.	OCC 7 Transport Equip. Operatives	OCC 8 Laborers Exc. Farm	OCC 9 Farm Related Workers	OCC 10 Service, Exc. Priv. Household	OCC 12 Repairmen & Maint.	OCC 13 Electrical Related Employmt.
$1A^P$	.199	.044	.187	.497	.640	.947	.738	.915	.779	.541	.727	.692
$1B^P$	.454	.348	.758	.220	.005	.001	.015	.009	.032	.114	.012	.017
$2A^P$	.183	.890	.947	.437	.200	.005	.212	.028	.082	.381	.069	.100
$2B^P$	.507	.046	.025	.016	.007	.006	.0	.015	.141	.011	.024	.216
$3A^P$	.018	.028	.192	.744	.222	.745	.801	.860	.513	.591	.184	.055
$3B^P$	.180	.029	.021	.002	.011	.002	.0	.001	.0	.020	.004	.017
$4A^P$	.282	.063	.0	.022	.002	.000	.004	.006	.010	.057	.0	.0
$4B^P$	.357	.155	.048	.254	.934	.754	.714	.369	.542	.217	.931	.919
$5A^P$	.321	.815	.041	.121	.277	.008	.037	.016	.083	.130	.057	.192
$5B^P$	.044	.013	.0	.006	.239	.035	.004	.019	.261	.082	.509	.428

See Table V-6 for explanation of interest factors.

Table V-15

Physical Demands

Occupation / Job Characteristic	OCC 1 Profess./ Tech. & Kindred	OCC 2 Mgrs. & Admin. Exc. Farm	OCC 3 Sales Workers	OCC 4 Clerical & Kindred Workers	OCC 5 Craftsmen & Kindred Workers	OCC 6 Operatives Except Transp.	OCC 7 Transport Equip. Operatives	OCC 8 Laborers Exc. Farm	OCC 9 Farm Related Workers	OCC 10 Service, Exc. Priv. Household	OCC 12 Repairmen & Maint.	OCC 13 Electrical Related Employmt.
STRENGTH[P]:												
VERY HEAVY	.002	.0	.0	.001	.015	.009	.009	.051	.010	.014	.003	.0
HEAVY	.005	.015	.010	.015	.134	.118	.119	.309	.257	.089	.180	.055
MEDIUM	.037	.016	.083	.156	.450	.384	.601	.344	.448	.504	.508	.393
LIGHT	.485	.435	.723	.427	.383	.436	.250	.293	.260	.359	.303	.435
SEDENTARY	.472	.534	.183	.401	.017	.053	.021	.003	.026	.035	.007	.118
CLIMBING AND/ OR BALANCING[P]	.021	.020	.001	.022	.331	.059	.184	.239	.291	.139	.176	.376
STOOPING, KNEELING, ETC.[P]	.025	.031	.067	.071	.454	.268	.339	.469	.584	.344	.597	.559
REACHING, HANDLING, ETC.[P]	.565	.278	.550	.829	.950	.996	.971	1.000	.958	.884	.989	.900
TALKING AND/ OR HEARING[P]	.721	.961	.953	.447	.342	.078	.175	.046	.209	.499	.308	.380
SEEING[P]	.527	.262	.189	.564	.726	.530	.607	.411	.598	.371	.699	.681

Table V-16

Environmental Conditions

Occupation / Job Characteristic	OCC 1 Profess./ Tech. & Kindred	OCC 2 Mgrs. & Admin. Exc. Farm	OCC 3 Sales Workers	OCC 4 Clerical & Kindred Workers	OCC 5 Craftsmen & Kindred Workers	OCC 6 Operatives Except Transp.	OCC 7 Transport Equip. Operatives	OCC 8 Laborers Exc. Farm	OCC 9 Farm Related Workers	OCC 10 Service, Exc. Priv. Household	OCC 12 Repairmen & Maint.	OCC 13 Electrical Related Employmt.
WORK LOCATION[P]:												
INSIDE	.875	.860	.817	.913	.617	.864	.247	.558	.235	.763	.754	.681
OUTSIDE	.020	.010	.115	.013	.152	.035	.197	.301	.491	.027	.037	.037
EXTREME COLD[P]	.003	.003	.0	.000	.003	.017	.000	.028	.000	.006	.004	.0
EXTREME HEAT[P]	.004	.007	.0	.001	.035	.058	.001	.012	.0	.098	.034	.0
WET AND/OR HUMID[P]	.010	.009	.010	.008	.053	.069	.000	.074	.215	.140	.035	.0
NOISE AND/OR VIBRATION[P]	.028	.036	.011	.018	.409	.334	.299	.304	.301	.050	.344	.185
HAZARDS[P]	.036	.022	.0	.021	.306	.214	.168	.229	.260	.219	.193	.476
ATMOSPHERIC CONDITIONS[P]	.020	.008	.010	.007	.132	.146	.078	.068	.361	.037	.126	.0

VI. ESTIMATION OF THE OCCUPATIONAL CHOICE MODEL

In this chapter we describe our estimates of the occupational choice model. The chapter is divided into three sections; Section A describes the final formulation of the choice model, Section B presents the estimated parameters, and Section C describes the interpretation of these estimates.

A. The Model

The original specification of the choice model had the individual maximizing a utility function, U, over the set of available occupational choices. Conceptually we saw the individual considering each choice, calculating the satisfaction to be obtained from each choice and choosing the one presenting him with the greatest satisfaction. Furthermore, we assumed that the individual's satisfaction was a function of present and future income, the attributes of the potential jobs and their own tastes and preferences. We specified the following function as a reasonable approximation to this utility function.

$$U = Y^{\alpha} W^{1-\alpha} \quad ;$$

where

$$Y = Y_f + \lambda_1 Y_p + \lambda_2 \cdot \text{MIL} + \lambda_3 \cdot \text{EDUC} + \lambda_4 \cdot \text{NAV}$$

$$W = \sum_j J_j B_j$$

and

$$\alpha = \underline{AX} + e$$

where e is a random variable distributed $N(0,S^2)$.

1. Some Modifications

Two modifications to this specification have been required during the estimation phase of this work.

First, we originally specified fifteen alternative choices. This proved to be an unmanageable number of alternatives for computational purposes and we have been forced to reduce this to only five alternatives. This decision required us to modify the specification of W so that instead of W being a linear function of job characteristics, W is now a series of indices of job quality. This means that we are unable to identify the impact of specific job characteristics as we had hoped to.

The second modification deals with the specification of the distribution of e, the stochastic taste parameter. Our original assumption was that this was normally distributed with mean zero and a variance, s^2, to be estimated. This proved to be unsuitable for the problem, however, since the normal distribution places no upper and lower bounds on the distribution of e and therefore α is not bounded. Values of α outside of the 0,1 range are unacceptable for theoretical reasons, however. In particular, a value of α outside of this range implies that an individual would prefer a low paying job with poor working conditions to a high paying job with superior working conditions, an unacceptable conclusion. We therefore chose the truncated normal distribution. This distribution was used by Burtless and Hausman [1] (1977) in a similar estimation procedure and has three important properties for this analysis. First, it allows us to constrain α to the 0,1 range required by economic theory. Second, it allows for a fairly flexible distribution of the e's. Finally, the truncated normal distribution is easy to compute. This is especially important since computational costs place a significant constraint on this work.

2. Detailed Specification of the Model

In order to fully specify the model we must further define the construction of Y, W, $\underline{AX}$, and the distribution of e.

The measure of lifetime income, Y, was constructed as follows:

$$Y = Y_f + \lambda_1 Y_p + \lambda_2 \cdot MIL + \lambda_3 \cdot EDUC + \lambda_4 \cdot NAV$$

where:

Y_f is future lifetime income measured in millions of dollars as described in Chapter IV.

Y_p is current income as measured by the young worker regression results reported in Chapter IV.

MIL is a dummy variable equal to one if the choice is the military other than the Navy.

EDUC is a dummy variable equal to one if the choice is education.

NAV is a dummy variable equal to one if the choice is the Navy.

---

1. Burtless, Gary, and Hausman, Jerry. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." Mimeographed. Massachusetts Institute of Technology, Fall 1977.

The job characteristic measures, BJ, are a series of indices relating the job quality of Blue Collar, Military, Education, and Navy to White Collar work.[2]

$$W = (B_1, B_2, B_3, B_4, B_5)$$

where $B_1 = 10$, and B_2 through B_5 are parameters to be estimated.

The B_i's provide us with a set of relative job quality indices. These are defined relative to B_1 which is defined to have the benchmark value of 10.

The taste indicators in the model are represented by AX, where the A_i's are parameters to be estimated and the X's are taken from the NLS data and are defined as follows:

X_1 = age at graduation from high school.

X_2 = dummy variable equal to one if steady work is very important to the individual.

X_3 = dummy variable equal to one if freedom from supervision is very important.

X_4 = dummy variable equal to one if the individual lives in the South.

X_5 = dummy variable equal to one if the individual considers himself unlikely to be able to finish college.

X_6 = dummy variable equal to one if the individual has a physical limitation.

X_7 = dummy variable equal to one if the individual's father attended college.

X_8 = parents annual income measured in $1,000's.

X_9 = scaled vocabulary score (see NLS manual for a description of the test).

2. We have defined Blue Collar as occupations 5 through 10, 12 and 13 in Table IV-1, page IV-2. White Collar is composed as occupations 1 through 4 on that table.

X_{10} = scaled math score (see NLS manual for a description of the test).

X_{11} = dummy variable equal to one if the individual is married.

These variables determine the typical orientation of the individual's indifference curve. The particular variables used to estimate the model were chosen because of prior analysis and because of their intrinsic interest. For example, prior analysis had indicated that father's education was important to the choice of young men. On the other hand, it has often been indicated that individuals from the south are more likely to join the services than those from other geographic areas. We therefore included a variable (South) to determine if this effect could be measured. As it turned out, this variable was found to have very little effect on the probability of enlistment.

The distribution of the error term, e, determines how much these indifference curves vary across individuals. We indicated above that we assumed that e was distributed truncated normal. We will now describe explicitly the distribution used.

We define the truncated normal distribution to be

$$f(x) = \frac{\dfrac{e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}}{\displaystyle\int_{-\underline{AX}}^{1-\underline{AX}} \dfrac{e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}} \qquad \text{for } -\underline{AX} \le x \le 1-\underline{AX},$$

$$f(x) = 0 \qquad \text{for } -\underline{AX} > x > 1-\underline{AX}$$

where $\underline{AX}$ is the inner product of the parameters A_i and the X_i's described above.

B. Estimated Parameters

The model was estimated for three major ethnic groups: black, hispanic, and all others. The estimates were obtained using the maximum likelihood procedure described in Chapter II, and are given in Table VI-1. Table VI-2 shows the number of observations in each occupation for each ethnic group. As is readily apparent, the largest number of observations are available for the "white" group, but a substantial number exist for the other two groups as well.

The log likelihood values for these equations give us some idea of the accuracy of these estimates. In particular, we can calculate the "average" percent contribution to the log likelihood function as:

$$e^{\frac{\text{Log } \mathcal{L}}{N}} .$$

This gives us a measure of the average probability of estimating an individuals choice correctly. These values are shown in Table VI-1 as Ave%.

C. Interpretation of Estimates

There is one particularly surprising result that we have found in this work. Current income appears to be of particularly negligible importance in determining an individual's occupational choice.

While on first glance this result may seem odd, it does not seem unreasonable for an individual to choose his occupation on the basis of expected future income rather than immediate economic gains.

Table VI-1

Estimated Parameters

	White	Black	Hispanic
S	.2128	.2665	.2965
λ_1 Current	-.0005	-.0395	-.0065
λ_2 Mil	-.0005	.0236	.0027
λ_3 Educ	-.0010	-.0020	.0003
λ_4 Navy	.0131	.0015	-.0196
B_2 Blue Collar	11.80	11.90	12.99
B_3 Mil	11.08	10.83	11.48
B_4 Educ	9.688	9.907	10.03
B_5 Navy	10.61	11.54	12.65
Age	.01075	.01084	.02012
Steady Work	.0347	.0511	.0829
Free Sup.	-.0037	.0008	.0202
South	.0134	-.0003	.0379
AB College	-.0105	-.1049	-.0972
Disability	-.0217	-.0218	-.0183
Fath. Educ.	.1141	.1190	.1116
Par. Inc. ($1,000)	0.00334	0.00532	0.00446
Voc. Score (100)	.002556	.002558	.002719
Math Score (100)	.002871	.002879	.003021
Married	-.1748	-.1661	-.1597
N	5679	644	281
Log $\mathcal{L}$	-6976.58	-867.97	-387.32
Ave. %	29.26%	25.98%	25.20%

Table VI-2

Occupation	Ethnic Group		
	White	Black	Hispanic
White Collar	684	54	41
Blue Collar	1,963	216	100
Military	439	127	39
Education	2,427	220	95
Navy	166	27	6

In addition, because of the need to reduce the number of choices to five, we have reduced our ability to reliably identify the effects of compensation relative to job quality. Thus, we must consider the overall estimates rather than placing too much weight on the individual B's and λ's. With this understanding we turn to a discussion of how individual differences effect the probability that an individual will choose one of our five occupation classes.

The individual's probability of choosing a particular initial occupation will depend upon the value of $\underline{AX}$ and ethnic group for the individual. The value of $\underline{AX}$ will of course vary from individual to individual. Table VI-3 shows the probabilities of choosing each occupation for the three ethnic groups and a variety of values of $\underline{AX}$.

Table VI-3 indicates that an individual with an $\underline{AX} = .3$ who is white has a 6.3% chance of choosing a white collar job, a 78.3% chance of choosing a blue collar job, a 3.5% chance of joining a branch of the service other than the Navy, a 10.6% chance of continuing his education, and a 1.4% chance of joining the Navy. On the other hand, a white individual with an $\underline{AX}$ of .7 has a 9.4% chance of choosing a blue collar job, a 16.2% chance of choosing a blue collar job, a 3.2% chance of choosing the non-Navy military, a 69.7% chance of continuing his education, and a 1.6% chance of joining the Navy.

These results raise the question of what values of $\underline{AX}$ are typical and how they change for different types of individuals. To answer this question, let us begin by considering the typical range of the values of the X_i's as shown in Table VI-4.

We can now compute the maximum mean and minimum values for $\underline{AX}$ by multiplying the values for the X_i's by the values for the A_i's given in

Table VI-3

AX	Ethnic Group	OCCUPATIONAL GROUP: White Collar	Blue Collar	Mil	Educ	Navy
.05	W	1.28%	96.29%	0.94%	1.14%	0.35%
	B	6.60	80.29	8.77	0.56	3.79
	H	0.18	94.53	2.54	1.85	0.90
.10	W	1.87	94.48	1.29	1.87	0.49
	B	8.49	76.28	10.17	0.86	4.20
	H	0.29	92.66	3.30	2.61	1.13
.15	W	2.66	91.94	1.74	2.99	0.67
	B	10.79	71.71	11.63	1.29	4.58
	H	0.47	90.28	4.22	3.64	1.39
.20	W	3.66	88.49	2.26	4.69	0.89
	B	13.51	66.61	13.06	1.91	4.91
	H	0.74	87.30	5.29	5.00	1.68
.25	W	4.90	83.96	2.85	7.14	1.15
	B	16.63	61.03	14.41	2.77	5.16
	H	1.13	83.65	6.50	6.72	1.99
.30	W	6.30	78.25	3.46	10.56	1.43
	B	20.09	55.06	15.57	3.96	5.32
	H	1.71	79.29	7.82	8.87	2.31
.35	W	7.78	71.37	[illegible]	15.12	1.69
	B	23.80	48.84	16.47	5.53	5.36
	H	2.52	74.22	9.20	11.44	2.62
.40	W	9.20	63.47	4.50	20.91	1.93
	B	27.61	42.52	17.02	7.57	5.28
	H	3.63	68.50	10.55	14.42	2.90
.45	W	10.39	54.84	4.79	27.89	2.10
	B	31.34	36.31	17.16	10.12	5.07
	H	5.10	62.25	11.80	17.73	3.12
.50	W	11.18	45.93	4.86	35.86	2.17
	B	34.76	30.38	16.89	13.22	4.74
	H	6.98	55.64	12.84	21.27	3.27
.55	W	11.48	37.20	4.70	44.48	2.14
	B	37.69	24.90	16.21	16.38	4.33
	H	9.30	48.89	13.61	24.87	3.33
.60	W	11.23	29.13	4.33	53.30	2.01
	B	39.95	19.99	15.17	21.04	3.86
	H	12.09	42.20	14.04	28.37	3.31
.65	W	10.50	22.05	3.31	61.84	1.31
	B	41.43	15.72	13.86	25.64	3.35
	H	15.31	35.80	14.12	31.57	3.20
.70	W	9.39	16.15	3.21	69.70	1.55
	B	42.07	12.13	12.38	30.59	2.84
	H	18.94	29.86	13.85	34.33	3.02
.75	W	8.07	11.47	2.59	76.59	1.28
	B	41.89	9.19	10.81	35.75	2.36
	H	22.90	24.51	13.28	36.53	2.78
.80	W	6.68	7.92	2.02	82.37	1.02
	B	40.96	6.84	9.26	41.02	1.91
	H	27.12	19.82	12.46	38.09	2.50
.85	W	5.34	5.32	1.52	87.03	0.78
	B	39.40	5.02	7.78	46.27	1.33
	H	31.52	15.31	11.46	39.00	2.21
.90	W	4.15	3.30	1.11	90.65	0.58
	B	37.35	3.63	6.43	51.40	1.20
	H	35.99	12.46	10.55	39.28	1.92
.95	W	3.15	2.26	0.79	[illegible]	[illegible]
	B	34.93	2.59	5.23	56.33	[illegible]
	H	40.43	[illegible]	[illegible]	36.38	[illegible]

AD-A080 323 MATHTECH INC PRINCETON NJ F/G 5/9
ANALYSIS OF ENLISTMENT IN THE NAVY AS AN OCCUPATIONAL CHOICE: P--ETC(U)
JAN 80 D E WISE, R L HORST N00014-78-C-0488
UNCLASSIFIED 2033 NL

2 OF 2
AD A
080323

END
DATE
FILMED
3-80
DDC

Table VI-4

Typical Values for the X 's

	High	Mean	Low
Age	21	18	16
Steady Work	1	.29	0
Free from Supervision	1	.24	0
South AB College	1	.36	0
Disability	1	.06	0
Father's Education	1	.30	0
Parents Income	25	12	1.5
Vocabulary Score	75.5	50.2	25.5
Math Score	65.5	50	25.5
Married	1	.15	0

Table VI-1. In order to calculate the maximum $\underline{AX}$ we use maximum values of the X_i's if the corresponding A_i is positive and the minimum value if the A_i is negative. Table VI-5 gives the maximum mean and minimum values for each ethnic group.

Table VI-5

Maximum and Minimum Values of $\underline{AX}$

	Maximum	Mean	Minimum
White	0.85	.54	.10
Black	0.91	.57	.03
Hispanic	1.19	.79	.20

The values for $\underline{AX}$ differ substantially from individual to individual depending upon the individual's personal characteristics (X_i's). As an illustration, consider a 20-year-old ($X_1 = 20$) Hispanic who considers steady work important ($X_2 = 1$), does not consider freedom from supervision important ($X_3 = 0$), lives in the northeast ($X_4 = 0$), considers himself unable to complete college ($X_5 = 1$), has no physical disabilities ($X_6 = 0$), whose father did not attend college ($X_7 = 0$), whose parents' income is $12,000 per year ($X_8 = 12$), who scored 55 on the vocational exam ($X_9 = .55$), who scored 65 on the math exam ($X_{10} = .65$), and who is not married ($X_{11} = 0$). In this case, $\underline{AX}$ is found by multiplying the values of the X_i's by their respective A_i's and summing.

$$\underline{AX} = 0.02012 \cdot 20 + 0.0829 \cdot 1 + 0.0202 \cdot 0 + 0.0379 \cdot 0 + (-0.0196) \cdot 1 + (-0.0183) \cdot 0 + 00.00446 \cdot 12 + 0.002719 \cdot 0.55 + 0.003021 \cdot 0.65 + (-0.1597) \cdot 0$$

or $\underline{AX} = 0.52268$. If we now look at Table VI-3 we see that this individual's value for $\underline{AX}$ falls roughly half way between .5 and .55, the values given for $\underline{AX}$ in the table. The probability of this individual joining the Navy is, thus, between 3.27 and 3.33 percent, the corresponding probabilities shown in the table for Hispanics. since the individual's $\underline{AX}$ is about half way from .5 to .55, we interpolate his probability of enlistment as approximately 3.3 percent. Correspondingly, we would estimate 8.0, 52.6, 13.2, and 22.9 percent to be the probabilities of choosing white collar, blue collar, military, or education, respectively. Thus, it is a simple process to compute the probability of an individual choosing one of our occupational groups given the requisite data on the individual.

This leads us to the question of what type of individual is most likely to join the Navy. The values of $\underline{AX}$ that maximize the probability of joining the Navy are shown in Table VI-6.

Table VI-6

Values of $\underline{AX}$ that Maximize the Probability of Joining the Navy

Group	White	Black	Hispanic
$\underline{AX}$	0.51	0.34	0.56

These values are slightly below the mean value for these groups but are nowhere near the extreme values of AX. Thus, the person most likely to join the Navy would be a more or less average individual coming from a family with slightly lower than average income who scored slightly below average on math and vocabulary ability tests. Individuals coming from either very high or very low income families, individuals scoring either very high or low on math and vocabulary ability tests, or individuals who are married are less likely to join the Navy.

The variables which have the greatest impact on AX are age, parents income, vocabulary and math scores, and marital status. The remaining variables have small effects.