WORKING PAPERS FOR

THE CENTER FOR RESEARCH ON SOCIAL ORGANIZATION

UNIVERSITY OF MICHIGAN

DEPARTMENT OF SOCIOLOGY

Paper #78
August 1972

Copies available through:

The Center for Research
On Social Organization
University of Michigan
330 Packard #214
Ann Arbor, Michigan 48104

ESTIMATING AN EQUATION WITH MULTIPLICATIVE AND ADDITIVE TERMS, WITH AN APPLICATION TO ANALYSIS OF WAGE DIFFERENTIALS BETWEEN MEN AND WOMEN IN 1960.*

Ross M. Stolzenberg

University of Michigan

*This paper benefited from several discussions between the author and Paul M. Siegel and from comments by Robert Hauser on an earlier draft. Paula Hudis Snyder was kind enough to make available the data used herein. This research was supported by Manpower Dissertation Grant No. 91-26-72-24 from the U.S. Department of Labor. The author retains all responsibility for remaining shortcomings.

This paper presents a method of obtaining regression-like estimators for the parameters of equations of the form

$$Y = b_o \xi_1 \prod_i X_i^{bi} + \sum_j c_j Z_j + \xi_2$$

We will attempt to demonstrate that such equations are useful as causal models when independent variables are hypothesized to be both additive and non-additive in producing effects on a dependent variable. Our plan is to first present a brief summary of currently available methods for estimating non-additive stochastic models, to present our method, to test the method with computer-generated data of known distribution, and, finally, to apply the method to a very brief analysis of earnings differentials between men and women in 1960.

I. Currently available methods for analysis of data according to non-additive models

Probably the simplest way to handle non-additive relationships among explanatory variables is to postulate a model of the form

(1) $$Y = b_o \xi \prod_i X_i^{bi}$$

where the b_i are parameters and is a random error term. When logarithms of Y and the expression on the right side of the equation are taken, we obtain a new equation (1*) which can be estimated in a straightforward manner using ordinary least squares regression.

(1*) $$\text{Log } Y = \text{Log } b_o + \sum_i bi \text{ Log } Xi + \text{Log } \xi$$

Certainly there are situations in which equation (1) is a useful model, but some reflection on the theoretical implications of the model may make one hesitant to use it very often: the model implies that the effect of every causal variable on the dependent variable Y is a function of virtually every other causal variable in the model. (Mathematically, this is observed by noting that the partial derivative of Y with respect to any of the X_i is a function of all of the other X_i.) Theory does not often suggest models in which this sort of interdependence of causal effects occurs, since it amounts to requiring "interaction effects" among all variables.

An alternative to using models of the form of equation (1) is to define a new variable $Z=X_i \cdot X_j$, where X_i and X_j are two variables which are believed to have joint non-additive effects on Y. Equation (2) represents such a model, and one can clearly see that it is amenable to straightforward regression analysis.

$$Y = a + \Sigma b_i X_i + cZ + \xi \qquad (2)$$

Examples of equations of this form can be found in Lane (1968), Thurow (1967), and Blalock (1965). The problem with models represented by equation (2) is that they provide the analyst with no straightforward measure of the relative contribution of each of the "interacting" variables to their joint effect on Y. Further, it is possible for the product $Z= X_i \cdot X_j$ to be highly correlated with X_i and only barely correlated with X_j purely as an artifact of the difference between the ratio of the standard deviation of

X_i divided by the mean of X_i and the ratio of the standard deviation of X_j divided by the mean of X_j. To see this, consider the case where X_i and X_j are normally distributed, uncorrelated, and have unit standard deviations. But let X_i have a mean of 200 and let X_j have a mean of 2. Accordingly, we can see that a change in X_i from 199 to 201, a change of two standard deviations, will increase Z by a trifling amount -- less than one per cent, assuming that X_j has remained constant. But a change in X_j from one standard deviation below its mean to one standard deviation above produces a change in Z of 300 per cent, assuming that X_i remains constant. Z would be so highly correlated with X_j that it would add only insignificantly to a regression equation already containing X_j, simply as an artifact of the means and standard deviations of X_i and X_j. So, under certain circumstances there may be drawbacks to using a simple product of two variables to account for a joint, non-additive relationship between them in determining the value of a dependent variable.*

Other least squares techniques which permit joint, non-additive effects of predictor variables on the dependent variable are dummy variable regression analysis and its derivative, Multiple Classification Analysis (MCA). In both MCA and dummy variable re-

---

* This is not to say that inclusion of a product term is never appropriate or useful. We merely point out that under certain circumstances a product term may be a poor indicator of a non-additive relationship between two predictor variables.

gression, the distribution of two variables which are suspected of having non-additive effects, X_i and X_j, are partitioned into m and n intervals respectively. m·n dummy variables are defined such that dummy variable I_{pq} equals one for a given data case if that data case takes on values falling into the p^{th} interval of X_i and the q^{th} interval of X_j; the dummy variables are set equal to zero otherwise. The dependent variable is then regressed on all but one of the dummy variables, as well as other variables of interest. (MCA involves further machinations which are unimportant for present purposes.) The only real drawback to using MCA or dummy variable regression is that both techniques produce a fairly large number of coefficients. The large number of coefficients which supposedly make clear the joint effect of X_j and X_i on the dependent variable can obscure rather than reveal the pattern of causality, and the analyst often finds himself falling back on notions of variance explained rather than the pattern of causation.* In short, what MCA and dummy variable regression techniques lack is a means of summarizing the causal effects of "interacting" variables.

Finally, there are iterative least squares methods of estimating the parameters of stochastic models. These methods ("hill

---

* It is possible that X_i and X_j have non-additive effects in only one or a few regions of their joint distribution. In such cases, the analyst need not include the full complement of m·n-2 dummy variables, but can include X_i, X_j and a dummy for each of the regions in which the interaction effect is suspected. In such cases the number of coefficients may well be small enough to allow easy interpretation of the results.

climbing," steepest descent, etc.) can be used to fit the parameters of virtually any model to a set of data; they are quite impressive in this respect. But these methods are spectacularly expensive to use because each iteration requires a separate pass over the data. Further, parameter estimates do not always converge as the iterations proceed, and while convergence may obtain using one method, there is no guarantee that it will obtain with another method used to estimate the same model with the same data.

To sum up, there are a number of methods of handling non-additivity among causal variables in least squares statistical analysis. MCA and dummy variable regression analysis avoid some of the problems of certain other methods discussed here, but present too many coefficients to provide easy interpretation and no means of assessing the relative contribution of each "interacting" variable to their joint effect on the dependent variable.

II. The proposed method

Equation (3) is a model which allows for non-additive relations among some causal variables, the $\{X_i\}$, but which also allows other causal variables, the $\{Z_i\}$, to be additive in their effects on Y.

(3) $$Y = b_o \xi_1 \prod_{i \geq 1} X_i^{bi} + \sum_{i \geq 1} c_i Z_i + \xi_2$$

ξ_1 and ξ_2 are random error terms; the c_i and b_i are parameters. Notice that equation (3) provides a separate parameter for each of the "multiplicative" variables, plus an additional coefficient

b_o for their total joint effect; we will show how the presence of separate parameters for the multiplicative variables allows at least some assessment of their relative importance in determining their joint effect on Y. More will be said about the interpretation of these parameters and about the existence of two error terms, but we will first explicate a method by which to estimate the parameters of equation (3).

Estimating equation (3). The first step in estimating equation (3) is to follow the same procedure used to handle non-additive effects in MCA and dummy variable regression: the distributions of the members of $\{X_i\}$ are partitioned jointly and a set of dummy variables $\{I_l\}$ indicating "membership" in a given cell of the joint partition is defined. Next, Y is regressed on all but one of these dummy variables and the $\{Z_i\}$ according to equation (4).

$$Y = do + \sum_{l=1}^{L} d_l I_l + \sum_{i=1}^{K} c_i Z_i + \xi_2 \qquad (4)$$

There are L + 1 cells in the joint partition of the $\{X_i\}$. This first regression analysis has three purposes: First, it provides estimates for the coefficients c_i net of the effects of the members of $\{Z_i\}$ on Y through their correlation with the members of $\{X_i\}$. Second, we can use this first regression analysis to test whether or not the variables in X_i add significantly to the variance in Y explained by the members of $\{Z_i\}$. The test is performed by first regressing Y on only the members of $\{Z_i\}$, and

then performing the regression analysis indicated by equation (4). Following Lane (1968) we note that

$$\frac{R_1^2 - R_2^2}{1 - R_1^2} \cdot \frac{N-L-K-1}{L}$$

is distributed as F with N-L-K-1 degrees of freedom in the denominator and L degrees of freedom in the numerator, where L is the number of dummy variables representing the joint partition of the $\{X_i\}$ which are entered into regression equation (4), N is the sample size, R_1^2 is the R^2 for the regression analysis including the dummy variables $\{I_\ell\}$, and R_2^2 is the R^2 for the regression analysis in which Y is regressed only on the $\{Z_i\}$. Of course, the test is informative of the explanatory power of the $\{X_i\}$ only insofar as the partition of the X_i captures the variance of its constituent variables. The third purpose of estimating equation (4) is that it provides a set of values for the adjusted mean of Y in each cell of the joint partition of the $\{X_i\}$. The cell means are adjusted in the sense that they are net of the contribution of the $\{Z_i\}$ to the value of Y. These adjusted means are obtained according to equation (5), where M_j represents the adjusted mean in the j^{th} cell of the joint partition of the $\{X_i\}$, and the $\{d_i\}$ are the coefficients in equation (4).

(5) $$M_j = d_o + d_j$$

Having obtained adjusted cell means for the partition of the X_i's, it is a straightforward matter to obtain estimates of the

parameters b_i for equation (3): Corresponding to each cell in the joint partition is a range of values for each of the Xi's. A mean for each range of values is obtained, so that for each cell we have a mean value for each member of $\{X_i\}$ and a mean value for Y net of the effects of the $\{Zi\}$. According to our model, equation (3),

$$(3^*) \qquad Y^* = b_o \xi_1 \prod_i X_i^{bi}$$

where Y* is the value of Y adjusted (in the sense we have been using the term) for the effects of the $\{Z_i\}$. Clearly, then, the situation has reduced to the common and easily-solved problem of estimating an equation of the form of equation (1) with grouped data. The logarithm of equation (3*) is taken, giving us

$$(3^{**}) \qquad \text{Log } Y^* = \text{Log } b_o + \sum_i b_i \text{ Log } X_i + \text{Log } \xi_i$$

which is estimated by a weighted ordinary least squares regression analysis in which the weights are the number of cases in each cell of the partition and the data points are defined by the set of adjusted cell means of Y and the corresponding means of the X_i's. Thus, we obtain the parameters $\{b_i, c_i\}$ in equation (3).

Strictly speaking, there are two points which should be mentioned: first, there is the problem of the additive error term [2] which has not been subtracted from equation (3) to produce (3*), but which we have nevertheless not included in (3*). The reason that we have not included this error term is that equation (3*) is estimated on grouped data using mean values of Y adjusted for

the effects of the X_i. As the number of cases in each group increases, the mean of $\mathcal{E}_2$ tends increasingly toward zero. So if the number of cases in the cells of the partition of the X_i is sufficiently large, we may disregard the error term $\mathcal{E}_2$. The second point which we should mention concerns the problem of estimating equations of the form of equation (1) using grouped data. Estimation procedures require knowledge of the means of the logs of the X_i and Y in the various intervals. The mean of the log of a variable in an interval cannot be obtained from the mean, except under exceptional circumstances, so we are faced with a problem which requires some simplifying assumptions for its solution. One such assumption is that the log of the mean of the variable in question is equal to the mean of its log. This is a common procedure and we have followed it in our computations, though some inaccuracy is introduced as a result.*

---

* One might also assume that the X_i have a certain distribution in the intervals. If this distribution has a density function f (X), and if the interval is (a,b), a theorem in probability gives the following result: the expectation of X (i.e. the probability limit of the mean of log X) is

$$E[X] = \int_a^b \text{Log } X \; f(X) \; dX$$

So, for example, if X is assumed to have a rectangular distribution in (a,b) (i.e. X is "evenly" distributed over the interval), the mean of the log of X when X is in (a,b) is

$$\int_a^b \text{Log } X \frac{1}{b-a} \, dX = \frac{1}{b-a} \left(X \text{ Log } X - X\right) \Big|_a^b$$
$$= \frac{1}{b-a} \left(b \text{ Log } b - b - a \text{ Log } a + a\right)$$

However, one is still left with the problem of obtaining the mean of log Y, and we can suggest no better method than to approximate the mean of the log by the log of the mean.

We will now test our method using computer-generated data of known distribution. We will explicate certain properties of the parameter estimates produced by the method in subsequent discussion.

Testing the method with computer-generated data. In order to provide at least a weak test of our method, we have used a computer to generate a data set composed of four normally distributed variables X_1, X_2, X_3, $\mathcal{E}$, and a variable Y which is defined by equation (6) below

$$Y = 0.5\, X_1^{0.3}\, X_2^{0.6} + 0.3\, X_3 + \xi \tag{6}$$

There are 1092 cases in the computer-generated dataset. Correlations means and standard deviations of X_1, X_2, and X_3 are given in table 1.; $\mathcal{E}$ has a mean of zero, a standard deviation of 0.2, and is uncorrelated with the other variables.

Table 1 -- Correlations, means and standard deviations of computer-generated variables

	X_1	X_2	X_3
X_2	.1249	1.0	.0385
X_3	.0087	.0385	1.0
mean	49.95	50.03	49.97
s.d.	1.02	0.89	0.95

X_1 and X_2 are each partitioned into 5 intervals, jointly defining 25 dummy variables, 24 of which were entered into the first regression analysis of Y on the dummies and X_3. The R^2 for this regression is 0.774. The analysis produces an estimate of 0.302

for the coefficient of X_3; the standard error for this estimate is 0.007. Since the estimate for this parameter is less than three-tenths of a standard error from the true value of 0.3, we feel safe in concluding that the method has been successful in recovering the coefficient of X_3. Performing the secondary regression on the results of the first analysis, we obtain the results presented in table 2 below. The R^2 for the secondary regression is 0.947.

Table 2 -- Secondary regression analysis parameter estimates compared to true parameter values

Parameter	True value	Estimated value	Standard error	True value-estimate standard error
$\text{Log}_e\ b_o$ [a]	-.693147	-.712869	.185943	.09
b_1	0.3	0.3162	0.0336	.49
b_2	0.6	0.5871	0.0336	.38

a) This is the $\log_e$ of the parameter; the parameter b_o has a true value of 0.5, the $\log_e$ of which is -.693147.

Note that the parameter estimates are all quite close to the true values. Indeed, if the difference between the estimated and true values of the parameters is divided by the standard errors for the estimates (see the far-right column of table 2), we see that the estimate for the log of b_o is less than a tenth of a standard error from the true value, and that b_1 and b_2 are less than a half and under two-fifths of a standard error from their respective true values. At even 25 per cent levels of confidence we could not re-

ject the null hypotheses that the parameter estimates are not different from their respective true values. So we conclude that our method has been successful in recovering the parameters of a model of the form given by equation (3). We will now proceed to apply the method to "natural" data.

III. An illustrative application of the method

In this section we very briefly apply our method to an analysis of earnings differentials between male and female workers in the United States in 1959-60. We will present theoretical reasons for choosing a model of the form of equation (3). The model will then be estimated separately for men and for women, and the results will be given to interpretation of the various coefficients.

The choice of an earnings model. Following Lester Thurow (1967), we hypothesize that the relationship between education and work experience in determining earnings is multiplicative, not additive. According to this hypothesis, employers not only offer higher starting wages to better-educated candidates for a job, but take years of schooling completed as an indicator of learning ability and make higher training investments in more-educated workers than in less-educated members of the labor force.* These higher training investments presumably make the experience (measured in time) gained by a better-educated worker more valuable

* This last point is supported by Doeringer and Piore's (1966) interview study of personnel practices in manufacturing firms.

than that gained by a less-educated worker, at least in terms of the increases in skill that come with on the job learning, resulting in a higher wage return to experience for more educated workers. So we wish to use an earnings model which captures this multiplicative relationship between education and experience.

Following Duncan (1969) we will include occupational attainment as a determinant of earnings by using the occupational prestige of the individual's 1969 U.S. Census detailed occupation category as an additive variable in the earnings model. The prestige score we use is Siegel's (1971).

It is well-known that the level of money earnings (though not necessarily the level of real earnings) is lower in the South and higher in the West than in the rest of the United States. We are hesitant to use a model in which a variable indicating region of residence multiplies the experience and education variables; we doubt that the rate of return to educational attainment and experience differs across regions of the country once we have held constant occupational attainment. Multiplying education and experience by a region variable would imply such a difference in these rates of return.

Because earnings are distributed log normally over about 65 per cent of the U.S. experienced civilian labor force in 1960,*

* This distribution is suggested by Cramer (1971,68). The figure of 65 per cent was obtained by the author after plotting the log earnings distribution on normal probability graph paper and observing the region in which the plotted points conformed to the expected straight line pattern.

we use $\log_e$ earnings of individuals in 1960 as our dependent variable. The use of log earnings rather than actual earnings means that absolute differences in the dependent variable are indicative of proportional differences in actual earnings.

Equation (7) presents a model which allows for a multiplicative relationship between years of experience in the labor force and educational attainment in determining an individual's log earnings, while allowing only for additive effects of occupational achievement and residence in different regions of the nation. \$ represents $\log_e$ earnings; P = the occupational prestige of the individual's 1960 Census detailed occupation; N = 1 if the individual resided in the northeast or north-central regions of the nation and N = 0 otherwise; W = 1 if the individual resided in the West and W = 0 otherwise; Ed = the individual's years of schooling completed; Ex = the individual's potential years of experience in the labor force. Ex is computed by subtracting the individual's years of schooling and the number 5 from his or her age. Caution is needed in interpretation of the experience variable, especially in making comparisons between men and women; we will discuss this issue later on. $\mathcal{E}_1$ and $\mathcal{E}_2$ are error terms.

(7) $$\$ = b_o \, Ed^{b1} \, Ex^{b2} \, \xi_1 + b_3 \, P + b_4 \, N + b_5 \, W + \xi_2$$

The data set used to perform the analysis is a subsample of the 1960 U.S. Census of Population 0.1 per cent public use sample. The subsample was drawn from persons aged 25 to 64 years old, who

were employed during the "census week" in 1960 and who reported earnings in excess of one dollar in 1959. The primary (or dummy variable) regression analysis was performed after partioning age and education into four and five intervals respectively, jointly defining 20 mutually exclusive categories. Table 4 shows the classification. Since experience (as presently defined) is a function of age and education, the partition into age-education categories also provides a set of mutually exclusive experience-education categories. Mean education and age within a given interval are taken to be the midpoint of the interval represented, except for the open-ended education interval, "college, four years or more," for which 16 years was used as the mean.

Dummy variables representing 19 of these categories were entered into separate regression analyses for women and men, according to education (7*), where the D_i are the dummy variables, K is a constant term, and all other variables and their parameters are identical to those in equation (7).

(7*) $$\$ = K + \sum_{i=1}^{19} c_i D_i + b_3 P + b_4 N + b_5 W$$

Results of the regression analyses are shown in table 3 below. We add the constant term K to each of the D_1's to obtain a value for the mean of \$ in the corresponding education-experience category adjusted for the effects of region residence and occupational prestige. The adjusted mean for the category corresponding to the dummy variable left out of the regression equation (7*) is just

the constant term. Table 4 presents the adjusted means for each category, the number of individuals in the category and the corresponding mean education and experience, by sex.

Table 3 -- Results of Primary Regression Analyses for Men and Women

Variable	Females		Males	
	Coefficient	Std Error	Coefficient	Std. Error
a) Constant	8.5830	3.8323	8.9950	2.6054
D 1	-.67437	.65466	-.69892	.21437
D 2	-.54973	.25658	-.12430	.15830
D 3	-.29012	.26694	-.0494	.16259
D 4	.06131	.33384	-.21103	.20523
D 5	-.18722	.33891	-.14724	.19451
D 6	-.79095	.38112	-.59846	.17782
D 7	-.42538	.22786	-.15343	.14956
D 8	-.41410	.22989	.13358	.15782
D 9	.18199	.32957	.00510	.20792
D 10	.29259	.32456	.44583	.18687
D 11	-.49953	.28900	-.35342	.19077
D 12	-.04249	.24130	-.12753	.14953
D 14	.13986	.27022	.15851	.23904
D 15	.43071	.43842	.37381	.21175
D 16	-.36864	.43700	-.42213	.18888
D 17	-.08650	.25147	-.27967	.16213
D 18	.27039	.37502	.45904	.22600
D 19	.0072	.37826	-.11383	.24681
D 20	.52947	.48151	.28181	.28163
Prestige	.01265	.00398	.01281	.00223
North	.48756	.12406	.16169	.06753
West	.33254	.16328	.16699	.09186
	R^2 for males: 0.234		R^2 for females: 0.223	

a) Variables D1, D2, . . ., D20 are the dummy variables representing cells in the joint partition of Education and Experience. Variable Di represents the ith cell.

Table 4 -- Adjusted means of $ for categories of the joint partition of Ed and Ex

Dummy Variable Index	Education Interval (Years)	Age Interval	Mean Age	Mean Ex	Mean Ed	Number in Category		Adjusted mean	
						Men	Women	Men	Women
1	0-7	25-34	30	20	5	17	2	7.199	6.299
2	8-11	25-34	30	15	10	53	22	7.778	6.424
3	12	25-34	30	13	12	46	19	7.853	6.683
4	13-15	25-34	30	11	14	19	10	7.691	7.035
5	16	25-34	30	9	16	23	7	7.755	6.786
6	0-7	35-44	40	30	5	34	36	7.304	6.182
7	8-11	35-44	40	25	10	73	36	7.749	6.548
8	12	35-44	40	23	12	53	36	8.036	6.559
9	13-15	35-44	40	21	14	18	10	7.907	7.155
10	16	35-44	40	19	16	27	11	8.348	7.266
11	0-7	45-54	50	40	5	25	15	7.549	6.931
12	8-11	45-54	50	35	10	74	28	7.775	6.931
13	12	45-54	50	33	12	32	25	7.902	6.973
14	13-15	45-54	50	31	14	12	19	8.061	7.113
15	16	45-54	50	29	16	18	5	8.276	7.404
16	0-7	55-64	60	50	5	25	5	7.480	6.605
17	8-11	55-64	60	45	10	46	24	7.623	6.887
18	12	55-64	60	43	12	14	7	8.359	7.244
19	13-15	55-64	60	41	14	11	7	7.789	6.966
20	16	55-64	60	39	16	8	4	8.184	7.503

* Dummy variable 13 was excluded from the regression analyses.

Logs of mean education, mean experience, and the adjusted mean of log earnings are taken and the secondary regression analysis defined by equation (7**) is performed using the number of individuals in each category of the joint partition as weights.

(7**) $$\widehat{\text{Log } \$} = \text{Log } b_0 + b_i \text{ Log Ed} + b_2 \text{ Log Ex}$$

The results of the secondary regression analyses are as follows; standard errors for coefficients are given in parentheses.

For males $$\$ = e^{1.796}\ \underset{(.0136)}{Ed^{.0871}}\ \underset{(.0105)}{Ex^{.0171}} \qquad R^2 = .711$$

For females $\$ = e^{1.449}\ Ed^{.1235}_{(.0218)}\ Ex^{.0529}_{(.0148)} \quad R^2 = .669$

Thus, we have estimates of all the parameters for equation (7); the estimated form of the earnings model, given separately by sex, is as follows:

For men:

(8) $\$ = 6.025\ Ed^{.0871}\ Ex^{.0171} + .0128\ P + .1617\ N + .1670\ @$

For women:

(9) $\$ = 4.258\ Ed^{.1235}\ Ex^{.0529} + .0126\ P + .4876\ N + .3325\ W$

In order to obtain some check on whether or not these results are reasonable, we substitute the mean values for Ed, Ex, N, W, and P for men into equation (8) and compare the values \$ obtained to the arithmetic mean of \$. We note that while ordinary additive regression analysis is constrained to reproduce the arithmetic mean of the dependent variable when the arithmetic means of the independent variables are "plugged into" the estimated regression equation, regression estimation of an equation of the form of equation (1) by taking the logs of all variables is constrained to reproduce the geometric mean of the dependent variable when the geometric means of independent variables are "plugged in" -- there is no guarantee that our method will reproduce either the geometric or the arithmetic mean when appropriate values are plugged in. The observed arithmetic mean of \$ for men is 8.425 for men; the value of \$ obtained from equation (8) and the arithmetic means of relevant variables is 8.468. So the results of

our analysis seem reasonable in terms of the ability of the model to approximate the observed arithmetic mean of $ from the observed arithmetic means of the independent variables.

Interpretation of the results. We must bear in mind a number of limitations of the present analysis: First, we have used cross-sectional data to estimate the parameters of a time-related process. Second, we have too few women in our sample to be very comfortable with the results of the secondary regression analysis on females: seven of the dummy variables for women represent categories containing fewer than 10 members of the sample. And, finally, the experience variable is undoubtedly a much weaker indicator of labor force experience for women than it is for men. However, the purpose of this analysis is to illustrate the use of a technique, and our analyses can serve this purpose in spite of the limitations just discussed.

We will attempt to show how our analyses can shed light on two questions: 1) Is education more or less important than experience in determining their joint impact on log earnings? And 2) What are the differences between the way education and experience jointly affect the attainment of log earnings by men and women?

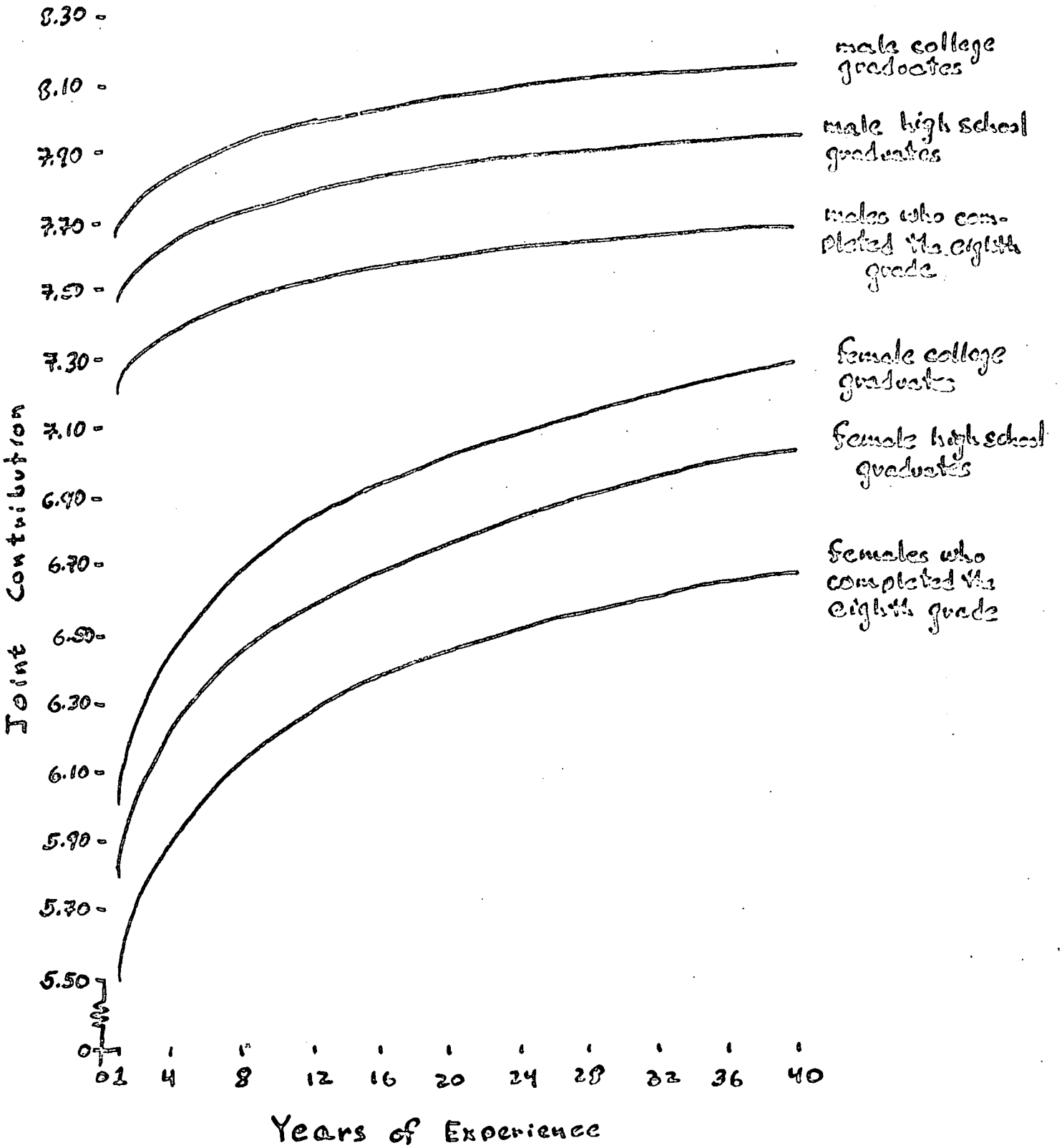
Since Ed and Ex are both in the same metric (years) we can compare their separate relative contributions to their joint effect on $ simply by comparing their exponents b_1 and b_2. Note

Graph 1 -- Fitted Joint Effect of Education and Experience on Log Annual Earnings

vs.

Years of Experience,

By Sex at Selected Educational Levels



that for both women and men, the exponent of Ed is larger than that of Ex.

In comparing the way education and experience affect the attainment of log earnings of men to the way they affect the log earnings of women, we might first note that the exponents for both education and for experience are larger for women than for men. This indicates that at a given level of education and experience, a marginal increase in the education or experience of a woman worker will increase her joint "return" to education and experience by a larger proportion than a similar increase in the education or experience of a man would increase his joint return to education and experience. Graph 1 presents the magnitude of the size of the joint experience-education contribution to $ at different levels of experience and education, according to equations (8) and (9). Note that while the joint effect of Ed and Ex for women is more sensitive to changes in Ex than it is for men, the size of the joint contribution is substantially less for females than it is for males. When we consider that the parameter b_o is almost 50 per cent larger for men than it is for women (6.025 vs. 4.258), this result is not surprising.

Finally, suppose that our experience variable is terribly inaccurate and that working women are, on the average, obtaining only half as much employment experience as men in the first 15 years after they complete their schooling. What would Graph 1

look like if we adjusted the women's experience accordingly? The slope of the curves for women would be twice as steep as they are presently over the range of one to 15 years of experience, since the X axis would be collapsed without commensurate change in the Y axis. In other words, our results would be changed in degree, but not in substance: the joint contribution of education and experience to log earnings would be even more sensitive to changes in experience for women than for men, and the size of the contribution would remain lower for women than for men.

Conclusions

We have presented a method for estimating equations which are useful in certain circumstances when an "interaction effect" or multiplicative relationship between two or more causal variables is hypothesized and purely additive effects of other causal variables are believed to occur. The method has been tested on computer-generated data and has been found to reproduce with adequate accuracy the parameters of the equation which was used to generate the data. The method has been applied to analysis of differences in the attainment of annual earnings by men and women, and the parameter estimates obtained appear to be quite reasonable. We certainly do not claim to have presented an answer to all or even many of the problems involved in handling multiplicative and additive relationships among causal variables simultaneously. However, we do believe that the method we have presented is useful in certain circumstances, and we have attempted to show very briefly some of the insights that can be obtained by using it.

REFERENCES

Blalock, H. M.
1965 Theory Building and the Concept of Interaction, ASR, 30, 1965, pp. 374-381

Cramer, J. S.
1971 Empirical Economics, N. Y.: American Elsevair Publishing Company

Doeringer, P. and M. J. Piore
1966 Internal labor markets, technological change, and labor force adjustment," Cambridge, Mass.: mimeographed paper.

Duncan, D. D.
1969 Inheritance of poverty or inheritance of race? Chapter 4 in D. Moynihan (ed.), On Understanding Poverty, N. Y. : Basic Books

Lane, A.
1968 Occupational mobility in six cities, ASR, 33, pp. 556-564.

Siegel, P. M.
1971 Prestige in the American occupational structure, unpublished doctoral dissertation, Department of Sociology, The University of Chicago.

Thurow, L.
1967 The occupational distribution of returns to education and experience for whites and Negroes, American Statistical Association, Proceedings of the Social Statistics Section, 1967.