

REPORT ON SEARS, ROEBUCK AND CO.'S
COMMISSION SALES HIRING AND PROMOTION PRACTICES
(Revised September 10, 1984)

Prepared for:
Equal Employment
Opportunity Commission

Prepared by:
Bernard R. Siskin, Ph.D

APPENDIX 3

WEIGHTED LOGIT AND MULTIVARIATE CROSS-CLASSIFICATION ANALYSES

The female proportion in the commission sales pool was adjusted in order to take into account various characteristics which Sears might consider in hiring commission salespersons or are related to factors which Sears might consider. The characteristics adjusted for and the methods used to perform the adjustments (weighted logit and multivariate-cross classification) are described below.

A. The Characteristics

Six characteristics generally were used for adjusting the commission sales pool.^{1/} The characteristics were divided into categories which were later collapsed for analysis purposes. The characteristics and categories follow.^{2/}

1. Job Applied For. The Job Applied For characteristic was grouped into eight mutually exclusive categories. Persons were assigned to the categories on the basis of the types of work they checked and/or wrote in on their applications (see Report, at 5).^{3/}

^{1/} For the logit analysis, additional variables were included and certain characteristics were redefined to conform to the logit model. These differences are described infra at 14.

^{2/} For the collapsed categories used for analysis, see infra at 15 and 25.

^{3/} The coding scheme used in assigning types of work to each of the categories is listed in Exhibit 1 attached to this Appendix.

(1) Commission Sales. This category includes persons (a)(i) who wrote in the words "commission sales" or "big ticket" or (a)(ii) who indicated an interest in sales positions by either checking sales, or writing in sales, and also indicated divisions or a product sold in a division designated for this purpose as a "commission sales division" and (b) did not indicate any other specific interest. For analysis of the full time commission sales pool, a division was designated as a commission sales division if in that division, during the period from 1973 through 1980, the proportion of full time sales hires who were commission sales hires exceeded the nationwide proportion of all full time sales hires who were commission sales hires (i.e., 51.9%).^{4/} For analysis of the part time commission sales pool, an analogous procedure was followed. The proportion of part time sales hires during the eight-

^{4/} The divisions designated as commission sales divisions for this purpose and the proportion of their full time sales hires that were commission are: Men's Dress Clothing (69.4%); Footwear (59.3%); Women's Shoes (64.1%); Men's Shoes (81.3%); Sewing Machines/Vacuum Cleaners (94.6); Stoves (94.7%); Washers/Dryers (93.9%); Refrigerators (95.5%); Freezers/Air Conditioners (90.9%); Hearing Aids (93.6%); Televisions (90.1%); Radios/Stereos (100%); Furniture (93.2%); Draperies (60.5%); Floor Covering (90.1%); Automotive Accessories (81.6%); Auto Tires (91.4%); Outlet Sales (71.6%); Fencing (82.2%); Plumbing/Heating (77.8%); Building Materials (81.8%); Kitchens/Dishwashers (88.9%).

(continued)

year period who were commission sales hires was 4.2%.^{5/}

(2) Commission Sales Plus Non-sales. This category includes all persons who would otherwise be included in category (1) above except that they also indicated "any" or a non-sales type of work.

4/ (continued)

Product lines which could not be matched unambiguously with a Sears division received a generic code (e.g., Appliances). For full time commission sales hires, these product lines were designated as sold in commission sales divisions if they were part of a product line grouping(s) (see infra at 7-10, for a complete description of the product line groupings, product lines, and generic codes) in which the proportion of full time sales hires who were commission sales hires exceeded 51.9%. The codes for these product lines are: 101; 106; 108; 110; 111; 113; 114; 115; and 116.

5/ For analysis of the part time commission sales pool, the divisions designated as commission sales divisions and the proportion of their part time hires who were commission sales hires are: Dresses (13.4%); Women's Nightwear (6.5%); Men's Sportswear (8.3%); Men's Dress Clothing (39.8%); Footwear (24.3%); Women's Shoes (26.9%); Men's Shoes (62.8%); Sewing Machines/Vacuum Cleaners (85.2%); Stoves (60.3%); Washers/Dryers (80.4%); Refrigerators (87.5%); Freezers/Air Conditioners (71.4%); Hearing Aids (6.3%); Televisions (25.4%); Radios/Stereos (17.3%); Furniture (51.2%); Draperies (4.5%); Floor Covering (29.7%); Automotive Accessories (25.5%); Truck Tires (100.0%); Auto Tires (62.6%); Outlet Sales (13.0%); Bargain Basement (8.7%); Fencing (22.6%); Plumbing/Heating (16.1%); Building Materials (13.3%); Kitchen/Dishwashers (29.7%).

For part time commission sales hires, a product was designated as sold in a commission sales division if the product was part of a product line grouping(s) in which the proportion of part time sales hires who were commission sales hires exceeded 4.2%. The codes for these product lines are: 101; 102; 106; 107 (males only); 108; 110; 111; 113; 114; 115; and 116.

(3) Noncommission Sales. This category includes all persons (a)(i) who wrote in "salesclerk," "sales and cashier," or "cashier"^{6/} or (a)(ii) who indicated an interest in sales positions and also indicated divisions or a product sold in a division designated for this purpose as a "noncommission sales division" and (b) did not indicate an interest in any other specific type of work. For this purpose, all divisions which were not designated as commission sales divisions in category (1) above were designated as noncommission sales divisions.

(4) Noncommission Sales Plus Non-sales. This category includes all persons who would otherwise be included in category (3) above except that they also indicated on their applications "any" or a type of work which does not meet the criteria for assignment to that category.

(5) Sales. This category includes all persons who indicated an interest in sales, which interest could not be categorized as either commission or noncommission sales under the above criteria, and did not indicate an interest in any other specific type of work.

^{6/} The words "noncommission," "noncommission sales," or "small ticket sales" were not written on any sample applications.

(6) Sales Plus Nonsales. This category includes all persons who otherwise would be included in category (5) above except that they also indicated an interest in a non-sales or "any" position.

(7) Non-sales. This category includes persons who indicated an interest in non-sales positions either by checking office, mechanical, or warehouse, or by writing in a non-sales position.

(8) Preference Not Indicated. This category includes all persons whose preference for type of work was not indicated on the application or was illegible or indeterminate.

2. Age. The Age characteristic was grouped into seven categories: (1) 16-17; (2) 18-19; (3) 20-24; (4) 25-34; (5) 35-44; (6) 45-64; and (7) 65 and over.

3. Education. The Education characteristic was grouped into four categories: (1) less than 12 years; (2) 12 years; (3) 13-15 years; and (4) 16 or more years.

4. Job Type Experience. The Job Type Experience characteristic was grouped into six categories. Persons usually listed more than one job type experience on the application. Therefore, the categories were ranked in a hierarchy in the order set out below. Each person was assigned to a category based on the highest ranked experience indicated on his application. The codes applicable

to written descriptions of prior job experience^{7/} for each of these categories are also indicated.

(1) Commission Sales. This category includes persons whose applications specifically indicate commission selling experience (code 6).

(2) Sales-High. This category includes persons whose applications indicate prior sales experience which was not identified as a commission sales position or as a lower level sales position such as "cashier," "sales and cashier," or "sales clerk" or as a sales position that included non-sales related tasks associated with lower level sales positions (codes 2 or 3).

(3) Sales-Low. This category includes persons whose applications indicate a prior sales position other than one meeting the criteria for categories (1) or (2) above (codes 1 or 5).

(4) Unspecified. This category includes persons whose prior experience was illegible^{8/} or indeterminate as to type of work (code 7).

7/ See Report on Development of Coding System For Prior Work Experience and Job Applied For From Sears' Employment Applications by Loren Solnick, at 6-7, Part A (Solnick Report).

8/ Persons with illegible applications were inadvertently included in category (6), "No Experience." However, because only 68 of 20,133 persons listed Job Type Experience that was totally illegible (including 7 of 1828 hires), the misplacement of these persons has no effect on any of the analyses.

(5) Non-sales. This category includes persons not classified in categories (1), (2), (3) or (4) above and who had non-sales experience (code 4).

(6) No Experience. This category includes persons for whom no experience was indicated.

5. Product Line Experience. This characteristic indicates whether an applicant has prior experience in a product line grouping without regard to whether that experience was sales related or not. This characteristic was used to take into account the possibility that product line experience itself provided an advantage in selection for a position related to that product line. For analysis of the full time commission sales pool, this characteristic was grouped into 13 categories of product line groupings (12 and "All Others")(see Exhibit 2, Part B); for part time, there were 15 product line groupings (14 and "All Others")(see Exhibit 2, Part C)^{9/}.

Product line groupings were created for analysis purposes because it often was not possible to determine the specific Sears division for which an applicant had related work experience. The following procedure was used. First, all listed experience on each application was coded. Where listed experience with a product corresponded to a specific

^{9/} This characteristic was redefined for the logit analysis (see infra at 14).

Sears sales division (e.g., Sewing Machines), the applicant was credited with experience for that specific division. Listed experience with products that could not be matched unambiguously with a Sears division received a generic code (e.g., Appliances).^{10/} Product line groupings were then formed by combining related divisions and generic codes.^{11/}

The composition of each grouping is based on the view that work experience with products that are generally related might assist an applicant in getting hired for a commission sales position in a division which sells any of the products in the grouping. The divisions and generic codes comprising a product line were determined on the basis of my own and Dr. Solnick's experience in the area of labor economics; the Sears Retail Compensation Manual, which lists divisions under the headings Home Fashions, Home Appliances, and Home Improvements (¶601, Rev. Dec. 1, 1970; ¶602, Rev. Oct. 1977); and the Sears Retail Training Manual, which provides that employees in certain related divisions are to receive the same type of training in

^{10/} A full description of the product line coding system is contained in the Solnick Report at 6, Part B.

^{11/} For example, an applicant who listed appliance repair experience would be credited with experience in the product line grouping "Appliances" as would a person who specifically listed refrigerator sales experience.

specific products. Set out in Exhibit 2, Part A, to this Appendix are all product line groupings and the divisions and the generic codes included in each grouping.

6. Commission Product Sales Experience. This characteristic was developed in order to identify sales experience which was likely to have been commission sales experience but was not specifically indicated as such on the application. The method used for classifying prior sales experience as commission product sales experience was to determine if the experience was with product lines which were part of a product line grouping sold at Sears on a commission sales basis.

The procedure followed in determining the product line groupings for which sales experience would be considered commission product sales experience was similar to that described in A.1.(1) above, supra, at 2, except that instead of comparing the proportion of sales hires in a division who were commission sales hires with the proportion of all sales hires who were commission, the comparison was made on the basis of product line groupings. Additionally, since a person's status as full time or part time during the period he obtained commission product sales experience is not relevant to whether or not he possesses such experience (and generally cannot be determined from the listed experience), the designations of product line groupings were not made

separately based on full time and part time hires. However, using the above criteria, product line groupings would have been designated commission or noncommission whether they were based on full time or part time sales hires in all cases except one -- Men's Apparel. For that product line grouping, because most of the sales hires were part time, sales experience in that product line grouping was designated as commission product sales experience based on the part time hires. The designation for generic codes grouped in more than one product line was based on the proportion of sales hires who were commission sales hires in all product line groupings of which the generic codes were a part.^{12/}

^{12/} Sales experience in the following product line groupings and generic codes was designated as commission product sales experience: Shoes; Appliances; Electronics; Home Furnishings; Auto; Home Building Material; 101; 102; 107(males only); 108; 111; 113; 114; 115; and 116.

B. Methodology For Weighted Logit and Multivariate Cross-Classification Adjustments

1. The Weighted Logit Analysis

A logit analysis was used to adjust the female proportion of the commission sales pool by taking into account characteristics of men and women in the pool. A logit is a form of regression technique which measures the effects of certain variables (independent variables) on the probability that a variable one wishes to predict (dependent variable) will occur. Unlike regressions where the dependent variable is continuous such as salary, a logit predicts the probability of an outcome of a yes or no event such as the probability of whether or not an applicant is hired.

A logit allows one to estimate the effect of a given independent variable controlling for the effects of the other independent variables. For example, the logit allows one to measure the effect of a person's sex on the probability of selection for males and females similarly situated with respect to the other independent variables. The logit coefficient associated with each independent variable thus represents the estimated net effect of that variable upon selection after adjusting for the impact of all other independent variables considered.

Using the coefficient of the sex variable from the logit

analysis, an estimate of the female proportion of the full time and part time commission sales pool as adjusted by the variables controlled for in the logit was determined. The coefficient of the sex variable is statistically significant for both the full time and part time analysis.

The procedures followed in the logit analysis and in estimating the female proportion of the adjusted commission sales pool as well as in testing for statistical significance are described below. The variables used in the logit model are also explained.

a. The Logit Sample

The following steps were taken in constructing the sample used in the logit analysis. The sample included the entire commission sales hire sample, stratified into full time and part time hires, except (i) persons whose applications did not have a reverse side (for work history) and, consistent with the definition of the commission sales pool used in other analyses (see Report at 19-20), (ii) persons who applied for non-sales positions only. The resulting sample sizes were: 801 for full time and 975 for part time.

A non-hired sample was then drawn for both full time and part time applicants. An applicant was placed in the full time or part time pool if he or she indicated an interest, respectively, in full time or part time work by checking the appropriate box. If both full time and part time were check-

ed, the applicant was eligible for selection for both pools.

To control for the existence of vacancies at the time of application, any applicant who applied when there was no opening (hire) within 90 days after his or her application date at the store where he or she applied^{13/} was removed from the pool of applications eligible for selection into the sample. As with the hire sample, persons who applied for non-sales positions only were also excluded from this pool of applications.^{14/}

For each store, full time and part time non-hired samples were selected randomly on a 1:2 (hires to non-hires) basis. The resulting sample sizes^{15/} were: for full time, 801 hires and 1253 non-hires for a total of 2054; for part time, 975 hires and 1628 non-hires for a total of 2603. Each hire and non-hire sample was weighted in proportion to

^{13/} The 33 stores in the Non-Hired Sample were used. A commission sales opening was considered to exist for a 90-day period prior to a commission sales hire in the store where an applicant applied. This accorded with Sears' policy that an application remained active for 90 days after filing (see supra at 6). The Master Data Base file constructed from computer extract tapes Sears supplied to the EEOC was used to determine the dates of hires in each of the 33 stores.

^{14/} Persons who applied for non-sales positions only had virtually no chance of being hired for a commission sales position regardless of any other characteristics they may have possessed. Since the logit is an additive model, the inclusion of a variable for such applicants will cause an improper result.

^{15/} The exclusion of non-sales applications was done after the samples were drawn and was accounted for in the weighting process.

the actual number of persons in its respective population. Because the samples are weighted, the analysis is referred to as a weighted logit.

b. The Variables

The variables used in the logit analysis were: (1) Job Applied For (JABF); (2) Age; (3) Education; (4) Job Type Experience (HJEXP); (5) Product Line Experience (REXP); (6) Commission Product Sales Experience (SC); (7) Sex; and (8) Month of Application (M). Variables (1) through (6) are defined above (see supra at 1-7)^{16/}, except that the logit model required Product Line Experience to be redefined slightly for nonhires. The revised definition is: REXP = 1, if an opening existed for which an applicant had related product line experience at the store at which he or she applied within 90 days of his application date and = 0, otherwise.^{17/} The variable Month of Application was added; its definition is self-explanatory.^{18/} The logit analysis also included a constant term.

^{16/} The definition of Age differs to the extent that persons with unknown ages were randomly assigned to age categories on the basis of the known distribution of ages in those categories. If the space on the application for Education was not completed, that application was assigned to an "unknown" category.

^{17/} This model assumes that the value of having product line experience is the same for each product line.

^{18/} Footnote 18 appears on page 15.

The characteristics Job Applied For, Age, and Job Type Experience, as used in the logit with certain categories collapsed for analysis purposes, are:

(1) Job Applied For. This characteristic, with Non-sales excluded (see supra at 13, note 14), was regrouped into five categories: (i) Commission Sales (Commission Sales and Commission Sales Plus Non-sales combined); (ii) Noncommission Sales (Noncommission Sales and Noncommission Sales Plus Non-sales combined); (iii) Sales (not combined); (iv) Sales Plus Nonsales (not combined); and (v) Preference Not Indicated (not combined).

(2) Age. This characteristic was regrouped into four categories: (i) 16 to 24; (ii) 25 to 34; (iii) 35 to 44; and (iv) 45 and over.

(3) Job Type Experience. This characteristic was regrouped into four categories: (i) Commission Sales and Sales-High; (ii) Sales-Low Level; (iii) Non-sales and No Experience; and (iv) Unspecified.

For full time and part time hires, respectively, a listing of the variables and the values of the coeffi-

18/ Footnote 18 (from previous page)

Persons with unknown Month of Application were placed in a reference group, which for this variable was the sixth month (M6). The reference group for each variable is listed in Exhibit 3, note 2.

cients of those variables from the logit analysis appear in Exhibits 3 and 4 attached to this Appendix.

c. Computation of Female Proportion of Commission Sales Pool as Adjusted

The estimate of the female proportion of the adjusted logit commission sales pool is computed from the following equation:

$$\text{Adjusted \% Female} = \frac{\#F \cdot P_f}{\#F \cdot P_f + \#M \cdot P_m}$$

where #F = the actual number of female applicants in the population; #M = the actual number of male applicants; and P_f and P_m are defined infra.

The equation representing the logit model by which the probability of hire is computed is:

$$\text{Prob of hire} = \frac{e^{\sum X_i \cdot B_i}}{1 + e^{\sum X_i \cdot B_i}}$$

where X_i = the i th variable in the model and B_i = the coefficient of the i th variable in the model. P_m and P_f , the probability of hire for men and women, respectively, are computed using the following equations:

$$P_m = \frac{e^{\sum \bar{X}_{im} \cdot B_i}}{1 + e^{\sum \bar{X}_{im} \cdot B_i}} \quad \text{and} \quad P_f = \frac{e^{\sum \bar{X}_{if} \cdot B_i}}{1 + e^{\sum \bar{X}_{if} \cdot B_i}}$$

where \bar{X}_{im} = the sample mean of the i th variable (excluding the sex variable) for males and \bar{X}_{if} = the sample mean of

the i th variable (excluding the sex variable) for females and, as indicated above, B_i = the coefficient of the i th variable in the model.

For, respectively, the full time and part time commission sales pools in the logit analysis as adjusted by the variables in that analysis, $\bar{X}_{im} \cdot B_i = -5.370, -5.127$ and $P_m = .0046, .0059$; $\bar{X}_{if} \cdot B_i = -5.703, -5.251$ and $P_f = .0033, .0052$. Applying the above formulae, the adjusted female percentage in the full time logit commission sales pool equals 53.4% and the adjusted female percentage in the part time logit commission sales pool equals 62.4%.

These percentages were calculated from logit pools which for the full time sample was 61.5% female and the part time sample was 65.3% female. To adjust the female proportions of the unadjusted commission sales pools in Tables 4 (full time) and 19 (part time) to take into account the effect of characteristics considered in the logit analysis, the unadjusted commission sales pools are multiplied by a factor that reduces them to the same extent that the characteristics reduced the female proportion of the unadjusted logit pool. This factor is used whenever the female proportion of the pool used in the analysis to control for the characteristics being considered differs from that of the original pool to be adjusted.

The factor can be described as follows: the female

proportion of the analysis pool as adjusted divided by the female proportion of the analysis pool before adjustment. For example, since the female proportion of the full time analysis pool as adjusted was 53.4% and the female proportion before adjustment was 61.5%, the factor would be .87. To adjust the female proportion of the eight-year nationwide commission sales pool (61.1%) by the characteristics analyzed, the female proportion of that pool is multiplied by .87. More generally, to arrive at the female proportion of the full time commission sales pool as adjusted by either the weighted logit or the multivariate cross-classification analysis, each of the percentages in Table 4 (percent female of unadjusted commission sales pools) is multiplied by the ratio of the female proportion in the full time adjusted logit pool to the female proportion in the full time unadjusted logit pool. The same procedure applies for the adjustment of the part time commission sales pools in Table 19.

d. Variance Estimates of Sex Coefficient

The hiring of an applicant as a commission salesperson is a "rare event," since only a very small percentage of all commission sales applicants is ever selected for a commission sales position. Given this fact, the logit sample must be considered small for analysis purposes. Where these conditions are present, the assumption of normality necessary to test significance does not fit the logit model

well, since the logit model assumes asymptotic normality, but the model approaches normality very slowly at the tail where "rare events" occur. Consequently, estimates of variance based on normality are suspect. Hence, a Balanced Replicated Sampling technique was used to estimate the variance and test the significance of the sex coefficient.^{19/}

To compute the variances, 16 "half samples" were selected separately for the full time and part time analysis. The logit analysis was performed on each half sample to determine the sex coefficient of that half sample.

The variance of the entire sample can be expressed as follows:

$$\text{Var}(Z^*) = \frac{\sum (Z_i - Z^*)^2}{16},$$

where Z^* = the sex coefficient computed from the entire sample; and Z_i = the sex coefficient computed from the i th half sample. The variance of the sex coefficient for the full time analysis is 3.94 standard deviations from zero; for the part time analysis the variance is 3.43 standard deviations from zero. The data from which these figures were computed are listed for the full time and part time analysis in Exhibits 5 and 6, respectively.

^{19/} A Balanced Replicated Sampling technique does not require any distributional assumption in calculating the variance. For the theory and advantages of this technique, see Kish and Frankel, Balanced Repeated Replicates for Standard Errors, Journal of American Statistical Association, Vol. 65, no. 331, pp. 1071-94.

2. The Multivariate Cross-Classification Analysis

A multivariate cross-classification analysis was also used to adjust the female proportion of the commission sales pool. The method is described for one characteristic (univariate analysis) and more than one characteristic (multivariate analysis).

a. The Univariate Cross-Classification Adjustment

The percentage of commission sales hires falling within each category of a characteristic, based on the Hire Sample, was determined. Next, the female percentage in the commission sales pool (based on the Non-Hired Sample, see Report at 18-19) falling within that category was determined. The percentage of the commission sales hires within each category was multiplied by the female percentage of persons in the commission sales pool within the category and the results were summed.

The following example illustrates the above procedure. Assume that the Job Type Experience characteristic is divided into two categories -- sales experience and no sales experience -- and that 30% of the hires are within the first category and 70% within the second. Assume also that female applicants comprise 50% of the persons in the first category and 65% of the persons in the second category. The expected

female proportion of hires^{20/} as adjusted for the characteristic would be 60.5% $((.30 \times .50 = 15.0\%) + (.70 \times .65 = 45.5\%))$.

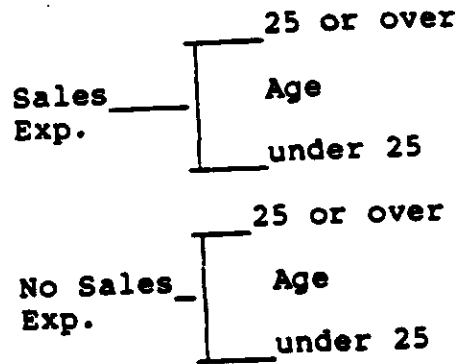
b. The Multivariate Cross Classification Adjustment

The same procedure is used to adjust the expected female proportion of hires for two or more characteristics (multivariate adjustment), except that the proportion of hires within each combination of categories is multiplied by the female proportion of persons in the commission sales pool within that combination of categories. Thus, in this type of multivariate adjustment the interaction of each category of one characteristic, or variable, with each category of the other characteristic(s) is considered.

The procedure followed in calculating the multivariate cross-classification adjustment can be illustrated by adding to the above example a second two-category characteristic, Age, divided into 25 or over and under 25. For the example, the percentages of commission sales hires and applicants for each of the combinations of categories listed in the chart on the following page must be determined.

^{20/} Because the female proportion of the commission sales pool as adjusted equals the expected female proportion of commission sales hires, these terms are used interchangeably.

Combinations of Categories in Multivariate
Cross-Classification Adjustment of Two Category
Sales Experience and Age Characteristics



As seen in the chart, there are four cells or combinations for two characteristics of two categories each.^{21/}

For the above example, assume that 20% of full time commission sales hires have sales experience and are 25 or over; 10% have sales experience and are under 25; 50% have no sales experience and are 25 or older; and 20% have no sales experience and are under 25. Assume also that in the full time commission sales pool women comprise 46% of the persons with sales experience who are 25 or older and 54% of the persons with sales experience who are under 25; of the persons with no sales experience women comprise 55% of those 25 or older and 70% of those under 25. The percentages of full time commission sales hires and female applicants in the full time commission sales pool with the combinations of characteristics indicated in each cell above

^{21/} The number of cells or combinations equals the product of the number of categories of each characteristic being considered. In this example, the number equals 2 x 2 or 4.

is presented for this example in the following chart.

Percentages of Commission Sales Hires and Commission Sales Applicants with Combinations of Sales Experience and Age Characteristics

		Percent of Hires		Female Percent of Pool	
Sales Exp.	25 and over	20	x	46	= 9.2
	under 25	10	x	54	= 5.4
No Sales Exp.	25 and over	50	x	55	= 27.5
	under 25	20	x	70	= <u>14.0</u>
					56.1%

The expected female proportion of hires as adjusted for the two characteristics would thus be the sum of the products of each of the above cells or 56.1%.

c. Problems of Fragmentation Bias

If all six characteristics, divided into the categories indicated above, were included in the multivariate adjustment, there would have been 69,888 and 80,640 cells or combinations for the full time and part time analysis, respectively. Where there is a large number of cells relative to the number of applicants, the applicants will tend to be spread out among the large number of cells, thus increasing the number of cells in which there are only hires. In these circumstances, the sexual composition of the ex-

pected hires tends simply to become the sexual composition of the actual hires, thereby masking the size of any disparities on the basis of sex. This is referred to as "fragmentation bias."

In an effort to reduce fragmentation bias to the extent possible, the four characteristics Job Applied For, Age, Education, and Job Type Experience were regrouped.^{22/} Job Applied For was regrouped into six categories: the five included in the logit analysis (see supra at 15) and Non-sales. Age was regrouped into two categories: (i) 16 to 24 and (ii) 25 and over. Education also was regrouped into two categories: (i) 12 or less years and (ii) 13 or more years. Finally, Job Type Experience was regrouped into two categories: (i) Commission Sales and Sales-High and (ii) All others.

d. Underestimation Using Cross-Classification Adjustments

There are two factors which can cause the cross classification method to underestimate the expected female proportion of commission sales hires.^{23/} First, the analyses do

^{22/} Because the number of combinations of categories increases geometrically with the number of categories of each characteristic, a minor regrouping may reduce the number of cells substantially. For example, six characteristics of six categories each yields 46,656 cells, while six characteristics of five categories each yields 15,625 cells.

^{23/} By contrast, the results of the logit analysis are independent of the biases that cause this underestimation.

not reflect the extent to which the characteristics adjusted for serve as a proxy for sex, i.e., the fact that the apparent success associated with the characteristics, held disproportionately by males, is not related to their desirability but to the greater success of males. Second, these analyses incorporate fragmentation bias referred to above (see supra at 23), which, however, was largely eliminated by regrouping^{24/}. Thus, while the multivariate cross-classification analysis seeks to adjust a pool to reflect what the expected female proportion of hires would be if Sears treated persons with like characteristics equally, it is important to consider the extent to which its results overestimate the adjustment (i.e., underestimate the expected female proportion of hires as adjusted).^{25/}

If Sears were selecting men on the basis of sex, the categories of characteristics which were disproportionately male would receive a greater proportion of hires than would be warranted by the desirability of the category. This is the direct result of the method by which a univariate or multivariate cross-classification adjustment is calculated,

^{24/} After regrouping, the extent of fragmentation bias appeared to be minimal (less than 3% for full time and less than 1% for part time).

^{25/} This procedure also does not consider whether or not a particular characteristic is job related nor whether the characteristic is actually used by Sears.

i.e., by multiplying the proportion of hires with a particular characteristic(s) and the female proportion of applicants with that characteristic(s).

Consider the following hypothetical example:

	<u>With Characteristic</u>		<u>Without Characteristic</u>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Hires	20	1	4	5
Applicants	100	20	20	100
Hire rate	20%	5%	20%	5%

Clearly the factor does not matter. Without regard to the characteristic, Sears would hire men at a 20% rate and women at a 5% rate. However, the univariate analysis would see that 70% of the hires had the factor while only 50% of the applicants had the factor and would incorrectly confuse the sex effect for the factor effect. The univariate analysis would adjust the expected female percent of hires from 50% down to 36.7% $((.7) (16.7) + (.3) (83.2))$.

Even if the factor is related to chance of selection, the female proportion of the commission sales pool would still be underestimated by an adjustment for a characteristic if men disproportionately possessed that characteristic and tended to be selected over women irrespective of the characteristic. This is due to the fact that the sex effect is confounded with the characteristic effect and the univariate and multivariate cross-classification technique

assigns that portion of the sex effect confounded with the characteristic to the characteristic.

To illustrate this consider the following hypothetical case:

	<u>With Characteristic</u>		<u>Without Characteristic</u>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Hires	200	50	50	50
Applicants	1000	500	500	1000
Hire rate	20%	10%	10%	5%

In this case the probability of selection for men is twice that of similarly situated (with respect to the characteristic or factor) women and the characteristic doubles an applicant's probability of selection.^{26/} Unadjusted the pool is 50% female. If we define P as the probability of selection without the factor if sex does not matter, then absent sex preference the expected number (E) of female (F) hires is

$$(2P) (500) + P(1000) = 2000P = E (F \text{ hires})$$

and the expected number of male (M) hires is

$$(2P) (1000) + P(500) = 2500P = E(M \text{ hires}).$$

The expected percent female absent discrimination is thus:

$$\frac{2000P}{2000P + 2500P} = \frac{2000P}{4500P} = 44.4\%.$$

^{26/} The data is generated by the equation Prob of Selection = Psex x Pfactor x P, where P = 7.78%; Pfactor = 1, if not present, 2 if present; and Psex = .643 if female, and 1.28 if male.

The univariate adjustment, however, overadjusts to

$$(.714) (.333) + (1.286) (.667) = 42.9\%.$$

This overadjustment by the cross-classification technique will always occur when there is a preference for men and men disproportionately fall within a category that is more successful than another. The greater the sexual preference, the greater is the overadjustment, and the greater the sexual disparity in the characteristic, the greater the overadjustment.

For the characteristics and categories used in the univariate and multivariate analyses, men in the apparently more successful categories generally were more successful than men in the other categories. This appeared to eliminate the possibility that the apparently greater success of persons in a category was entirely the result of the proportion of males in that category. However, this analysis indicated that in most cases women in the same category as the more successful men were not similarly successful and that, hence, some portion of the greater chance of success apparently associated with the category of the characteristic was actually a reflection of the fact that men were more successful than women regardless of the characteristic. Thus, there is an overadjustment for the characteristic. However, the univariate and multivariate cross classifica-

tion analyses do not provide any way of measuring the extent to which they underestimate the female proportion in the commission sales pool nor do they allow a determination of the extent to which the characteristic considered is a proxy for sex bias.

in the commission sales pool nor do they allow a determination of the extent to which the characteristic considered is a proxy for sex bias.