Sears Exh. 5

UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF ILLINOIS
EASTERN DIVISION

EQUAL EMPLOYMENT OPPORTUNITY COMMISSION, Plaintiff,	)	
v.	)	Civil Action No. 79-C-4373
SEARS, ROEBUCK AND CO., Defendant.	)	Judge Nordberg

WRITTEN DIRECT TESTIMONY OF
DR. DAVID A. WISE REGARDING COMMISSION SALES

My name is David Wise. I am the Stambaugh Professor of Political Economy at the John F. Kennedy School of Government at Harvard University. My research deals with empirical analysis of human resource and public policy issues and associated methodological questions. Sears Exhibit 5-1 is a copy of my resume.

1. The Equal Employment Opportunity Commission's (EEOC's) Report on Sears, Roebuck and Co.'s Commission Sales Hiring and Promotion Practices (rev. Sept. 10, 1984) (EEOC Commission Sales Report) illustrates that if statistical techniques are applied and the results interpreted without regard to the situation that is being examined, the resulting conclusions are as likely to be fiction as fact. Three major characteristics of the report undermine its conclusions. First, important differences between the qualifications and preferences of men and women are ignored. Second, the report ignores the fact that relatively small unmeasured differences can lead to grossly inaccurate inferences given the low probability that any applicant will be

hired for a commission position. Third, in addition to making comparisons that are not meaningful, the statistical significance of the differences the EEOC reports in its hiring analysis is enormously exaggerated.

2. The EEOC Commission Sales Report is based on assumptions that defy common sense--the assumptions that men and women were equally interested in and qualified for commission sales positions at Sears during the years 1973-80. See, e.g., EEOC Commission Sales Report at 43, 74.

3. It is common knowledge that the labor force participation patterns of men and women are markedly different, reflecting their differing interests, educational choices, and family responsibilities.

4. The EEOC Commission Sales Report itself demonstrates that its underlying assumptions are false. For example, the report indicates that, on average, female applicants were younger, less educated, less likely to have commission sales experience, and less likely than male applicants to have prior work experience with products sold on commission at Sears. See EEOC Commission Sales Report at 34, 64; see Plaintiff's Exhibit 14; Plaintiff's Exhibits Siskin 36, 37. Data included in one of the EEOC's supplementary reports indicate that the overwhelming majority of persons who actually applied for commission sales positions were men. See Sears Exhibit 6-10.

5. In an attempt to overcome its false assumptions, the EEOC Commission Sales Report presents two analyses that control for several variables simultaneously -- "a logit analysis" and a

"multivariate cross-classification analysis." Both are statistical techniques designed to analyze the relationship between characteristics, such as prior employment experience, and outcomes, such as whether a person is hired.

6. The EEOC's logit and multivariate cross-classification analyses are quite dissimilar. The EEOC imposed restrictions on the logit that virtually assure that it will yield a smaller adjustment than the multivariate analysis. The multivariate allows the effect of each variable on the probability of hire to be different for each product line. In addition, for any one product line, the variables are allowed to "interact" with each other. For example, the multivariate allows the effect of product line experience to be greater if a person also has commission sales experience. Similar flexibility could have been incorporated in an appropriately defined logit model. It was not. Thus, the multivariate model is bound to fit the data better than the simplistic logit model.[1]

---

1. In addition, the sample of non-hired applicants used in the EEOC's logit analysis was inappropriately chosen on the basis of an "outcome" variable. Non-hired applicants were excluded from the logit analysis if they applied more than 90 days prior to the day when someone was hired into commission sales. Apparently the idea was that if someone applied when there were no commission sales openings, the applicant had no chance of being hired.

But, the fact that there were no hires does not mean that there were no openings. It may have been that a qualified and interested person simply did not come along. By defining commission sales openings in an after-the-fact way, the EEOC selected the non-hired sub-sample used in its logit analysis on the basis of the "outcome" it was trying to predict, i.e., commission sales hires.

This EEOC procedure is known as endogenous stratification or endogenous sampling. It biases downward the estimated effect of variables that are positively related to selection. It is an inappropriate procedure for selecting a sample.

7. In its discussion of its multivariate cross-classification analysis, the EEOC stresses what it terms "fragmentation bias" and "proxy bias." See EEOC Commission Sales Report, Appendix 3 at 24-29. In both cases, however, the arguments for the possibility of such biases rest on the EEOC's prior assumption that men and women with the same characteristics as measured in its analysis were the same in all other relevant respects. If, instead, differences in hiring probabilities for men and women reflect differences in qualifications and preferences that the EEOC did not measure, the adjustments based on the multivariate analysis are not large enough.[1]

8. The EEOC's multivariate cross-classification analysis is too crude to take into account the actual differences between the measurable qualifications of men and women applicants, let alone their different job preferences. The variables included in the analysis are defined in such a way that they mask or ignore relevant information.

a. The analysis failed to take into account the amount of experience an applicant may have had. An applicant with one day of experience was treated like an applicant with ten years of experience. Moreover, the effect of age was assumed to be the

---

1. EEOC expert Dr. Bernard R. Siskin argued that the extent of proxy bias might be loosely evaluated by comparing the adjustment based on the multivariate analysis with the adjustment based on the logit. But, the specifications of the two procedures are so different in important respects that the comparison as a measure of proxy bias has no meaning. As discussed above, restrictions in the logit compared to the multivariate are virtually bound to lead to a smaller adjustment in the logit than in the multivariate analysis.

same for both men and women. But, as the EEOC acknowledges elsewhere, "Generally women have less attachment to the labor force; hence, given a man and women [sic] of the same age, the woman is less likely to have the same amount of prior experience as a man." EEOC Report on Checklist Compensation Practices of Sears, Roebuck and Co., Appendix 5 at 1. If the amount of work experience goes unmeasured, the effect of age should be greater for men than for women. The EEOC Commission Sales Report ignores this fact.

b. Product line experience was treated too restrictively. Experience in a product line was allowed to affect the probability of hire only in that product line. But, this does not take into account the fact that a variety of experiences may be relevant to any particular job.

c. The analysis failed to consider the different types of work experience an applicant had. Instead, "[e]ach person was assigned to a category based on the highest ranked experience indicated on his application." EEOC Commission Sales Report, Appendix 3 at 5. In addition to employing a somewhat arbitrary ranking scheme, this procedure resulted in the loss of information on other types of job experience as well as information about the amount of this and other experience. For example, an applicant with one day of commission sales experience was treated the same as an applicant with one year of commission sales experience as well as one year of noncommission sales experience. More information was lost when the EEOC "collapsed" or "regrouped" variables.

9. Even if all available information on measured attributes of men and women were captured, unmeasured attributes and preferences would still surely play an important role in understanding hiring at Sears. The wide variety of products sold at Sears makes this apparent. Despite this, the EEOC assumed that men and women with the same characteristics as measured in its analysis were the same in all other relevant respects. For example, the conclusions in the report are based on the assumption that men and women with the same measured characteristics were equally likely to be interested in and qualified for selling products such as hardware, draperies, fencing, and automotive accessories. To maintain such an assumption requires a prolonged suspension of disbelief.

10. Individuals are likely to differ with respect to unmeasured variables. For example, men are much more likely to know about automobiles than are women and women are much more likely to know about women's clothes than are men. But, such obvious facts may be completely missed in the data used in an analysis. In any case, an absence of data does not demonstrate that everyone is the same. Inferences should not be made in a vacuum. Statistical techniques are designed to aid reason, not to fly in the face of it.

11. Even if differences between men and women due to unmeasured characteristics were relatively small on average, inferences about Sears' employment practices based on the assumption that there were no differences are likely to be grossly inaccurate.

a. According to the EEOC Commission Sales Report, the hiring of 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." EEOC Commission Sales Report, Appendix 3 at 18. "[F]or each full-time commission sales hire there were over 300 unsuccessful applicants." Id., Appendix 1 at 1. "[F]or each part-time commission sales hire there were over 250 unsuccessful applicants." Id., Appendix 1 at 2.

b. The very small probabilities of selection are important to the interpretation of the data. They magnify the effect of excluded variables and the significance of characteristics not easily measured. For example, suppose that the actual qualifications (unmeasured as well as measured attributes and preferences) of men and women to sell auto accessories are normally distributed, as in figure 1.

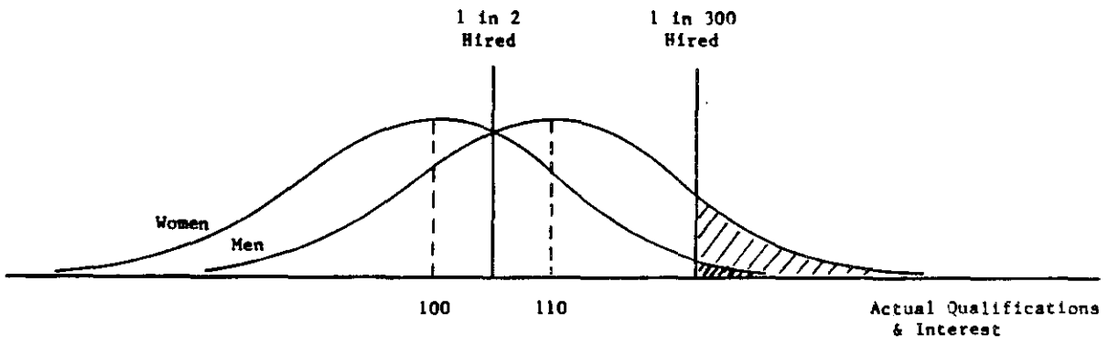

FIGURE 1

Men on average are better qualified than women, 110 versus 100. If this difference represents one standard deviation, 15.9 percent of the women would be more qualified than the average man. Assume that there are equal numbers of men and women applicants and that applicants are selected according to their qualifications, with the best selected first, then the next best, and so forth.

If 1 in 2 applicants were hired, 30.9 percent of the hires would be women. But, if only 1 in 300 applicants were hired, only 3.7 percent would be women. The lower the probability of selection, the smaller the proportion of women among the hires.

If men and women are erroneously assumed to be equally qualified and interested, the incorrect inference of different hiring standards for men and women is magnified by the small selection probabilities. If 1 in 2 applicants is hired, the ratio of the probability that a man is hired to the probability that a woman is hired in the example would be 2.24. A man would be 2.24 times as likely to be hired as a woman. The incorrect inference would be that the difference between 1 and 2.24 reflects different hiring standards given the false assumption that the ratio should be 1. If only 1 in 300 were hired, the ratio would be 26.4 and the difference between 1 and 26.4 would be assumed incorrectly to reflect different hiring standards.

c. That men are 26.4 times as likely as women to be hired when 1 in 300 applicants is hired does not mean that on average men are 26.4 times as qualified and interested as are women. Indeed, many women would be more qualified than the average man. It is just that among the one third of one percent of all

applicants that is most qualified and interested, men make up the large majority.

d. One might ask: if all members of a college class of 300 students applied to a Sears store, what is the likelihood that the most qualified to sell auto accessories would be a woman? If this happened at 100 Sears stores, in how many cases would the best qualified person in the applying class be a woman? Similarly, in how many of the applying classes would the person best qualified to sell women's clothing be a man?

e. Table 1 illustrates the importance of taking the small selection probabilities into account. Two alternative proportions of women in the applicant pool, 0.611 and 0.372, are shown. These proportions correspond, respectively, to the EEOC's estimates in its full-time hiring analysis of (a) the proportion of women in its unadjusted applicant pool and (b) the proportion of women in its applicant pool as adjusted by its multivariate cross-classification analysis. See EEOC Commission Sales Report, Table 5 at 21; Plaintiff's Exhibit Siskin 78 (rev. Jan. 24, 1985).[1] The difference between the average level of qualifications and interests of men and women is expressed in terms of standard deviations. An average for men 0.5 standard deviations above the average for women, for example, means that 30.9 percent of the women are still more qualified and interested

---

1. Tables 2 and 3 (attached) are analogous to Table 1. The proportions of women applicants shown in Table 2, 0.672 and 0.577, correspond to those in the EEOC's part-time hiring analysis for Sears' Eastern, Southern, Southwestern, and Pacific Coast Territories. The proportions of women applicants shown in Table 3, 0.653 and 0.560, correspond to those in the EEOC's part-time hiring analysis for Sears' Midwestern Territory.

than the average man.[1] Two alternative selection probabilities are shown, 1 in 100 and 1 in 300.

Table 1

ILLUSTRATION OF THE EFFECT OF DIFFERENCES BETWEEN MEN AND WOMEN ON THE PROPORTION OF WOMEN HIRES GIVEN LOW SELECTION PROBABILITIES

Mean level of qualifications and interest for men minus the mean level for women*	Proportion of Women Hires by Selection Probability and by Proportion of Women Applicants			
	1/100		1/300	
	P = .611	P = .372	P = .611	P = .372
+1.5	.017	.005	.009	.003
+1.0	.086	.030	.060	.021
+0.5	.291	.129	.255	.110
+0.25	.447	.231	.425	.216
0.00	.611	.372	.611	.372
-0.25	.755	.535	.771	.557
-0.5	.863	.693	.883	.730
-1.0	.968	.907	.978	.935
-1.5	.995	.982	.997	.990

* Expressed in standard deviations.

The values of P, the proportion of women in the applicant pool, correspond to the EEOC's estimates in its full-time hiring analysis of the proportion of women in its unadjusted and its adjusted (by its mulitvariate cross-classification analysis) applicant pools.

1. If the mean for men is greater than the mean for women by:
   a. 0.25 standard deviations, then 40.1 percent of the women are more qualified than the average man;
   b. 1.0 standard deviation, then 15.9 percent of the women are more qualified than the average man;
   c. 1.5 standard deviations, then 6.7 percent of the women are more qualified than the average man.

Suppose that after controlling for some attributes, the adjusted pool is 37.2 percent women. Assume, however, that measured variables do not completely control for differences in the qualifications and preferences of men and women. If the actual qualifications of men are 0.25 standard deviations higher than those of women and only 1 in 100 applicants is selected, the proportion of women hires would be only 0.231, even though 40.1 percent of all the women applicants would be better qualified than the average male applicant. If 43.1 percent of the women were more qualified than the average man, the proportion of women hires would be 0.27, the proportion the EEOC reports as the actual figure in its full-time hiring analysis. Thus, even if on average, the differences are small, inferences based on the erroneous assumption that men and women are equally qualified and interested can be extremely misleading.

f. The table also makes it clear that the proportion of women hires would depend very greatly on the job to be filled. Auto parts and draperies are not the same thing. If women were 61.1 percent of the applicants, 1 in 100 applicants were hired, and, for a particular product, the average qualifications of men were one standard deviation higher than the average for women, women would make up only 8.6 percent of the hires. If for another product, the average for women were one standard deviation higher than the average for men, women would make up 96.8 percent of the hires.

If Sears were comprised of only two such divisions, the proportion of women hired would depend on the number hired in each division. In general, the number of hires in divisions for

which men are better qualified compared to the number of hires in divisions for which women are better qualified will have a substantial effect on the aggregate proportion of women hired.

12. A framework within which the actual hiring results can be interpreted may be developed by expanding the model to account for different levels of measured attributes among applicants.

a. With reference to figure 2, suppose that the measured attributes of women are positively related to their actual average qualifications according to the upwardly sloping line labelled women.

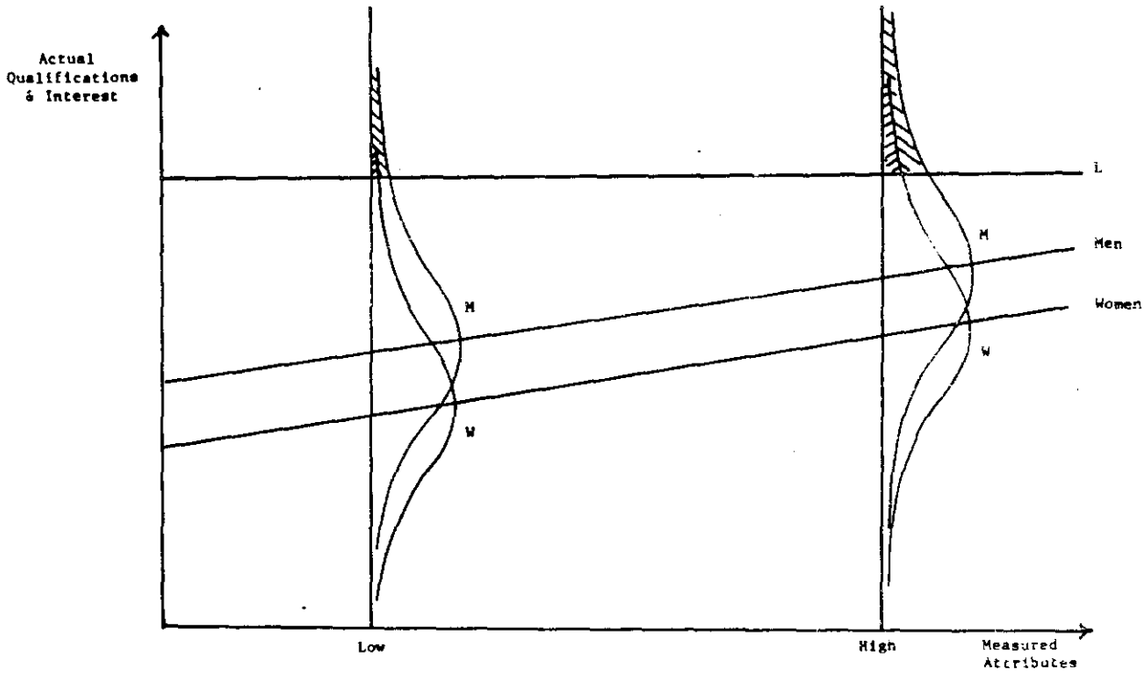

FIGURE 2

On average, women with a higher level of measured attributes are more qualified for and interested in commission sales. But, all women with the same level of measured attributes do not have the

same actual qualifications. Because of unmeasured attributes, some women with a low level of measured attributes will be much better at selling, and some will be much worse. This variation in actual qualifications is represented by the bell-shaped curve centered over the average level of actual qualifications for women with a low level of measured attributes. A similar distribution for women with a high level of the measured attributes is also shown.

b. A like representation is shown for men. But, the average actual qualifications of men are assumed to be somewhat higher than those for women. This difference could reflect attributes that could be measured but are not, as well as attributes such as interests and preferences that are more difficult to measure. At each level of the measured attributes, the distribution of actual qualifications among men is higher than the distribution among women.

c. Now suppose that to be hired, actual qualifications must exceed the level L. (In practice, this level would depend on the qualifications of the other applicants at the store to which a person has applied.) Suppose again that only a very small proportion of applicants is hired.

d. These assumptions lead to the following conclusions.

° First, at each level of measured attributes, men are considerably more likely than women to be hired because men are on average somewhat better qualified than women and because only a small proportion of applicants is hired.

- Second, the ratio of the probability that a man is hired to the probability that a woman is hired increases as the level of the measured attributes decreases. It is a rarer event for someone with a low level of measured attributes to have actual qualifications above L than for someone with a high level of the measured attributes. The lower the level of the measured attributes, the lower the probability of selection, and the lower the proportion of women among the hires.

- Third, the average sales performance of men hires will be above the average for women. Among hires -- those applicants whose actual qualifications exceed L -- men are on average better qualified than women at each level of measured attributes. Although measured attributes may not take into account many personal characteristics that predict sales ability, actual sales performance is based on all attributes of hires that relate to sales ability including those that are not accounted for or are not easily quantified in the empirical analysis.

- Fourth, the average sales performance of the least qualified men who are hired, those just barely

above L, will be equal to the average sales performance of the least qualified women.

e. This last implication would not hold with some forms of affirmative action. Assume that in order to meet affirmative action goals, women were hired at a level of qualifications lower than the level required for men. In this case, Figure 2 would be altered as shown in Figure 3, with the required level of qualification lower for women than that for men. Consequently, the sales performance of the least qualified women hires would tend to be lower than the performance of the least qualified men.

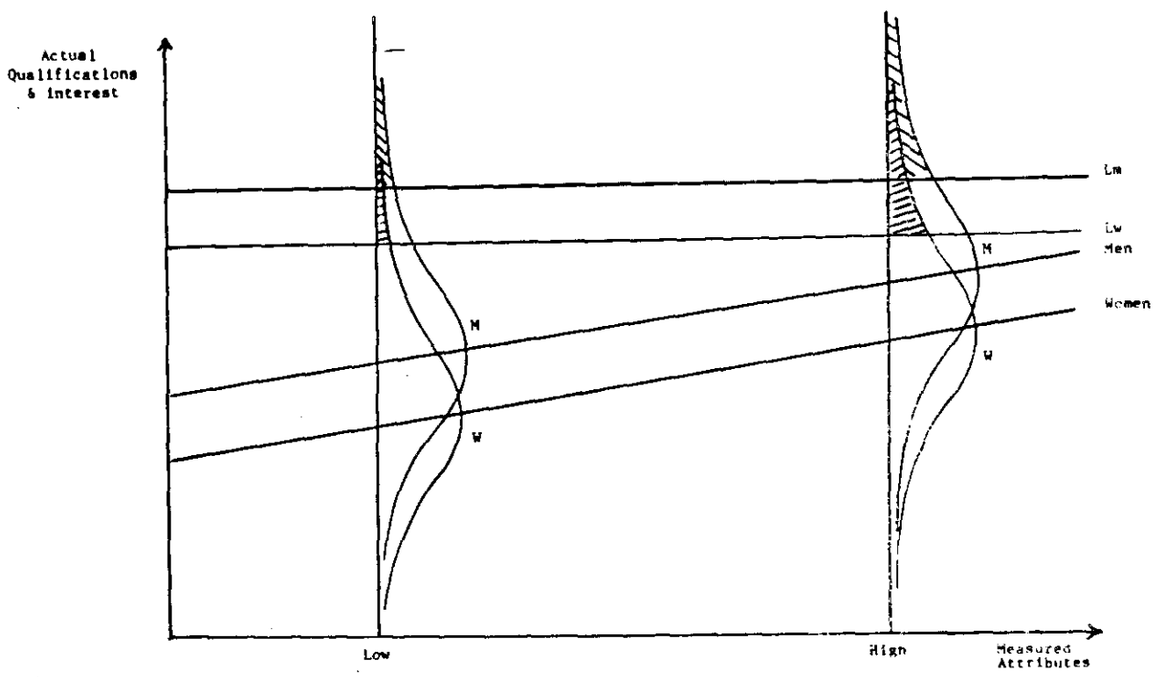

FIGURE 3

13. The EEOC Commission Sales Report contends that there are statistically significant disparities between its "expected"

numbers and the actual numbers of female commission sales hires. But, the key question is not by how many "standard deviations" or "Z-statistics" two numbers are different. Rather, the basic question is whether proper numbers are being compared. In the EEOC Commission Sales Report, they are not. Only if equally qualified and interested groups are compared are the Z-statistics meaningful. It serves no purpose to demonstrate that obviously different things are different.

14. Aside from the meaninglessness of its comparisons, the EEOC Commission Sales Report's calculations of statistical significance in its hiring analysis are incorrect.[1]

a. The EEOC's analysis essentially ignores the fact that hiring at Sears is geographically and administratively decentralized. The analysis is conducted as though an applicant at one Sears store is simultaneously an applicant at every Sears store. This error is particularly egregious when it is implicitly incorporated in the calculation of Z-statistics. The EEOC calculates Z-statistics as if all hires were drawn from the same pool with a known proportion of women. In fact, hires are not drawn from a single pool. They are drawn from store pools with unknown proportions of women.

b. The EEOC aggregated estimates from individual stores in its non-hire sample to obtain an estimate of the proportion of women in the pool from which it assumes hires are drawn. The variance of this estimate must be accounted for if measures of the statistical significance of the difference between the EEOC

---

1. A more detailed explanation of this point is contained in Appendix 1.

"expected" and actual number of women hires are to be obtained. In the EEOC Report on Selection of Non-Hired Applicants Sample,[1] EEOC expert Dr. Eugene P. Ericksen estimated the variances for the nationwide estimates of the female proportion of the non-hired applicant pool; however, the EEOC Commission Sales Report did not incorporate the variances into its calculations of statistical significance. In addition, when the EEOC calculated the statistical significance of the differences it found on the basis of its logit and multivariate cross-classification analyses, it also ignored the variance of the adjustment factors it used.

c. Calculating the Z-statistics taking these variances into account yields measures of statistical significance far lower than those presented by the EEOC. See Sears Exhibit 5-2, 5-3, 5-4. For example, the actual Z-statistic associated with the difference between the EEOC "expected" hiring rate (as adjusted by its multivariate cross-classification analysis) and the actual full-time hiring rate for 1973-80 is approximately 4. Yet, the EEOC reports a Z-statistic of 27.1.

15. The EEOC Commission Sales Report virtually ignores differences between the qualifications and interests of noncommission salesmen and saleswomen in its promotion analysis.

16. In part, the EEOC attempts to justify this absence of investigation by stating that "seniority is not positively

---

1. Plaintiff's Exhibit Ericksen 3.

correlated with the probability of being promoted." EEOC Commission Sales Report at 78 n.30. But the lack of a positive, linear relationship does not imply that there is no relationship at all.[1]

17. If nondiscriminatory promotion practices have been followed at Sears, equally qualified men and women would have an equal chance of being promoted (assuming that they were equally interested in promotion). If women were subjected to discrimination when Sears promoted persons from noncommission to commission sales, one might expect that more qualified women were passed over in favor of less qualified men. The data, however, do not support this hypothesis. The sales performance of commission salesmen was typically higher than that of commission saleswomen. On average, the top commission salesmen outsold the top commission saleswomen, and the bottom commission salesmen outsold the bottom commission saleswomen. This is consistent with affirmative action, but not easily reconciled with discrimination.

18. An appropriate analysis of the hiring and promotion decisions at Sears must go beyond the analysis undertaken by the EEOC. Differences in the preferences of men and women must be

---

1. The EEOC also relies on a study conducted by a Sears employee. EEOC Commission Sales Report at 78 n.30. The study to which the EEOC refers cannot be used to draw conclusions about the characteristics of noncommission salespersons because it does not separate noncommission salespersons from commission salespersons. Furthermore, only a small proportion of noncommission salespersons is promoted to commission sales positions, and it is not the characteristics of the average salesman and saleswoman that are the issue.

addressed and alternative indicators of the proportion women in a realistic applicant pool must be considered. Such an approach has been followed by Dr. Joan G. Haworth. I have participated with her in discussion of commission sales hiring and promotion at Sears and concur in the approach she has taken.

David A. Wise

Subscribed to and sworn before me
this 16th day of May 1985.

My Commission Expires: Jan 31 1987

Virginia M. Mitchell

Table 2

ILLUSTRATION OF THE EFFECT OF DIFFERENCES BETWEEN MEN AND WOMEN ON THE PROPORTION OF WOMEN HIRES GIVEN LOW SELECTION PROBABILITIES

Mean level of qualifications and interest for men minus the mean level for women*	Proportion of Women Hires by Selection Probability and by Proportion of Women Applicants			
	1/100		1/250	
	P = .672	P = .577	P = .672	P = .577
+1.5	.024	.014	.014	.009
+1.0	.115	.074	.086	.055
+0.5	.353	.262	.319	.233
+0.25	.514	.412	.496	.394
0.00	.672	.577	.672	.577
-0.25	.802	.728	.813	.743
-0.5	.892	.844	.906	.863
-1.0	.976	.963	.982	.973
-1.5	.996	.994	.998	.996

* Expressed in standard deviations.

The values of P, the proportion of women in the applicant pool, correspond to the EEOC's estimates in its part-time hiring analysis for Sears Eastern, Southern, Southwestern, and Pacific Coast Territories of the proportion of women in its unadjusted and its adjusted (by its multivariate cross-classification analysis) applicant pools.

Table 3

ILLUSTRATION OF THE EFFECT OF DIFFERENCES BETWEEN MEN AND WOMEN ON THE PROPORTION OF WOMEN HIRES GIVEN LOW SELECTION PROBABILITIES

Mean level of qualifications and interest for men minus the mean level for women*	Proportion of Women Hires by Selection Probability and by Proportion of Women Applicants			
	1/100		1/250	
	P = .653	P = .560	P = .653	P = .560
+1.5	.021	.013	.013	.008
+1.0	.105	.068	.078	.050
+0.5	.332	.248	.299	.220
+0.25	.493	.395	.474	.377
0.00	.653	.560	.653	.560
-0.25	.788	.714	.800	.729
-0.5	.883	.834	.898	.854
-1.0	.974	.960	.981	.970
-1.5	.996	.993	.997	.996

* Expressed in standard deviations.

The values of P, the proportion of women in the applicant pool, correspond to the EEOC's estimates in its part-time hiring analysis for Sears Midwestern Territory of the proportion of women in its unadjusted and its adjusted (by its multivariate cross-classification analysis) applicant pools.

STATISTICAL SIGNIFICANCE

The "Z-statistics" presented in the EEOC's hiring analyses do not reflect the likelihood that the differences the EEOC calculates could have occurred by chance. The "Z-statistics" that are presented greatly exaggerate the statistical significance of the differences.

EEOC UNADJUSTED ANALYSES

If, as the EEOC assumes, the hiring process can be viewed as though all hires are selected from the same pool, then a test of the null hypothesis that the predicted number of women hires is equal to the actual number is given by the statistic

$$Z = \frac{NP - NQ}{\sqrt{Var(NP - NQ)}} = \frac{N(P - Q)}{\sqrt{N^2Var(P) + N^2Var(Q)}} = \frac{P - Q}{\sqrt{Var(P) + Var(Q)}}$$

where N is the number of hires, P is an estimate of the proportion of women "sales" applicants in the non-hired pool, and Q is the proportion of women among the hires. Both P and Q are random variables.

EEOC expert Dr. Eugene P. Ericksen estimated P and gave considerable attention to estimates of its standard error.[1] Under the null hypothesis that the proportion of women among the non-hires is equal to the proportion among the hires, the

---

1. See Plaintiff's Exhibit Ericksen 3 (EEOC Report on Selection of Non-Hired Applicants Sample).

variance of Q may be estimated by P(1 - P)/N. An estimate of the variance of P, which I believe to be an underestimate,[1] has been calculated by Dr. Ericksen for each year. The variances were ignored in calculating the "Z-statistics" throughout the EEOC Commission Sales Report. The report treats estimates of P as though they were known population values which they clearly are not. In each instance, part of the observed difference between the predicted and actual proportions of women hires is due to the sample that happened to be drawn. Had a different sample of stores been selected, a different "difference" would have been found. The standard errors Dr. Ericksen estimated are indications of the range of possible values of P that would have been obtained had repeated samples been selected.

Because most of the uncertainty about the difference between P and Q is due to uncertainty about the true value of P, the variance of the estimated difference P - Q will be greatly underestimated and the "Z-statistic" will be greatly exaggerated if

---

1. There were two stages in the non-hired applicant sampling procedure. First, the stores were chosen. Second, a sample of applicants was chosen from each of the selected stores. Both of these two stages of sampling will make a contribution to the variance of the estimated P.

Dr. Ericksen's estimate of the variance of P is based on the variation among stores and does not fully capture the contribution to the variance made by having to estimate the value of P for each store. Therefore, the reported variance estimates are likely to understate the true variances.

As a check on his estimated variance, Dr. Ericksen also computed a Jackknife estimate of the variance of P. This check is subject to the same criticisms as the original variance estimate because it too is based on the variation among stores and not on the uncertainty from estimating P in each store. When many units are sampled, the Jackknife estimator used by Dr. Ericksen will be numerically equal to the variance estimator calculated using the formula for the variance of a ratio mean.

the variance of P is ignored. Thus, the statistical significance of the difference will be exaggerated.

EEOC ADJUSTED ANALYSES

The comparisons the EEOC makes as a result of its logit and multivariate cross-classification analyses are based on adjusted P values. The adjusted values are subject to further random variation.

When the adjustment is based on the logit model, the adjusted estimate of the proportion women in the non-hire pool has been derived (apparently) from the relationship

$$\frac{P}{1 - P} = r \frac{A}{1 - A}$$

where P is the adjusted proportion and A is the actual proportion women among the hires, and r is the ratio of the probability that a man will be hired to the probability that a women will be hired, r_m/r_w, after controlling for the variables used in the logit analysis. The ratio r has been approximated by

$$r \triangleq e^{-2B},$$

where B is the sex coefficient in the logit analysis. If B were zero, r would be equal to 1, and P would be equal to A -- in other words, the proportion women in the adjusted pool would be equal to the actual proportion among the hires. In EEOC expert Dr. Bernard R. Siskin's terms, the "expected" number of women hires would equal the actual number. Thus, a test of the hypothesis that the sex coefficient is zero is equivalent to the hypothesis that the "expected" number of women is equal to the

actual number. Put another way, given that the estimated coefficient on sex is presumed to reflect the difference between the hiring probabilities of men and women, a test of the hypothesis that there is no difference between them may be based on the estimated sex coefficient. The estimated coefficient is -.488 and its estimated standard error, as calculated by Dr. Ericksen, is .1378.[1] The coefficient divided by its standard error yields a Z-score of 3.54. Yet, when this estimate is used to adjust the estimated proportion of women in the non-hire pool for the nation, the "Z-statistic" is claimed to be 57.9.

When based on the multivariate cross-classification analysis, the adjusted P values are of the form

$$P_a = (P)(Z_a/Z),$$

where Z_a, Z, and P are all random variables. Z is the proportion women in the multivariate cross-classification analysis sample and Z_a is the adjusted proportion in the sample.

The variance of P_a may be estimated by the delta method. If the vector $X = (P, Z_a, Z)$, $G(X) = P_a = (P)(Z_a/Z)$, and $dG(X)/dX = (Z_a/Z, P/Z, -PZ_a/Z^2)$, then

$$Var(P_a) = Var(G(X)) \doteq (dG(X)/dX)' \ (CovX) \ dG/dX,$$

where CovX is the covariance matrix of the elements of X, and the prime indicates the transpose of the vector dG(X)/dX.

The current Z_a is estimated to be .4063 and Z is taken to be

---

1. The standard error of the sex coefficient reflects sampling variation among both non-hires and hires.

.6674.[1] Dr. Ericksen estimated the standard error of the initial estimate of Z_a to be .023. The current Z_a is assumed to have the same standard error. The standard error of P is taken to be the standard error estimated by Dr. Ericksen. The standard error of Z is estimated by Var(Z) = Var(P)/b, where b is the proportion of the observations used to obtain the all years, national estimate of P that is also used to estimate Z.[2]

In addition to the variances of P, Za, and Z, the variance of P_a also depends on the correlations between the individual estimates P, Z_a, and Z. To calculate a nationwide standard error for P_a, the covariance between P and Z may be estimated by b Var(Z). If the correlation between P and Z_a is zero, $Cor(P,Z_a) = 0$, and $Cor(Z_a, Z) = 0$, the standard error of P_a is .025. With P_a estimated to be .372 for the nation as a whole, a test of the hypothesis $H:P_a = .27$ yields a Z-statistic of 4.00.[3] Yet, the EEOC reports the Z-statistic for all years combined to be 27.1. See Plaintiff's Siskin Exhibit 78 (rev. Jan. 24, 1985).

To calculate standard errors for the adjusted P_a values for

---

1. The original estimate of Z_a was .479. From prior estimates of the adjusted and unadjusted proportion of women non-hires, .450 and .627 respectively, and from the relationship .450 = (Z_a/Z) (.627), Z is equal to .6674.

2. Z is calculated from a subsample of the observations used to calculate the all years, national estimate of P. If a proportion b is used, Var(Z) ≐ Var(P)/b. Here b is taken to be .4 and Dr. Ericksen's standard error of the national estimate of P, .023, is used. The standard error of Z is then .0364. The covariance between P and Z is bVar(Z).

3. If both correlations are .4, the standard error is .0235 and the Z-statistic is 4.34. If Var(P) and Var(Z) were zero, the S.E. would be .021 and the Z-stat = 4.84. If $cor(P,Z_a) = .5$ and $cor(Z_a, Z) = 0$, the Z-stat = 3.36.

individual years, it is necessary to assume a covariance between P and Z. This covariance depends on the observations used to estimate Z that were also used to estimate P for the year. Assuming that this number is small, the covariance is assumed to be zero.

For the all years estimate, the elements that enter the computation are summarized as follows:

Variable	Estimate	S.E.	Cov with Z_a	Cov with Z
P	.611	.021	°	bVar(Z)
Z_a	.4063	.023	----	°
Z	.6674	$\sqrt{Var(P)/b}$	----	-----

A dot indicates that the corresponding covariance term must be assumed. For yearly estimates, the elements that enter the computation may be summarized by:

Variable	Estimate	S.E.[1]	Cov with Z_a	Cov with Z
P	Dr. Ericksen's estimate	Dr. Ericksen's estimate	°	°
Z_a	.4063	.023	----	°
Z	.6674	$.021/\sqrt{b}$	----	----

1. For the variable Z, the standard error, .021, is Dr. Ericksen's estimate for all years and b is taken to be .4.

Standard errors of the unadjusted and the adjusted P values for the nation and for each year are shown below, assuming that the covariances are zero.

Year	S.E. of P	S.E. of P_a
All years	.021	.025
1973	.035	.032
1974	.059	.049
1975	.024	.036
1976	.034	.036
1977	.044	.042
1978	.035	.041
1979	.029	.040
1980	.057	.053

To compute the variance of differences $P_a - Q$, the variance of Q is also accounted for. It is again estimated by $P_a(1 - P_a)/N$.