USING EEO-1 DATA TO EXAMINE GENDER-BASED OCCUPATION SEGREGATION BY INDUSTRY

Bliss Cartwright, U.S. Equal Employment Opportunity Commission
Patrick Ronald Edwards, U.S. Equal Employment Opportunity Commission[1]

Key Words: EEO-1, equal employment opportunity; North American Industrial Classification System (NAICS),occupational segregation, glass ceilings, odds ratios

INTRODUCTION

Statistics has played a key role in enforcing equal employment opportunity laws and regulations. They have become a critical part of nondiscrimination litigation especially that based on Title VII of the Civil Rights Act of 1964[2]. Meier, Sacks and Zabell trace the reliance on statistics in this arena to a Supreme Court case involving the Hazelwood School District.[3] Over the years these techniques have been refined,[4] but statistics remain a key component of enforcing nondiscrimination laws. They have been integrated into the enforcement process at the Equal Employment Opportunity Commission (EEOC). See for example, the Commission's compliance manual section on Compensation.[5]

Today, one of the most difficult issues to examine in equal employment opportunity is the glass ceiling problem. In 1995, the Federal Glass Ceiling Commission concluded that, "today's American labor force is gender and race segregated – white men fill most top management positions in corporations."[6] The issue has taken on particular significance as women and minorities have increased their occupational status. The term is generally used to refer to instances where women and minorities have progressed within a firm but find it difficult to make the movement into key higher level management positions, or management positions at all. The social disadvantage of these glass ceilings is the inability of the most qualified employees to move into the most important positions due to irrelevant criteria such as race or gender.

The successful elimination of glass ceilings, through the use of government intervention, requires the identification of glass ceilings and the implementation of an effective enforcement strategy. The purpose of this paper is to examine statistical techniques that can be used to identify industries and firms where there is a high potential for gender-based management barriers. It is critical to stress that this purpose is the identification of techniques to use as initial screens not techniques appropriate to the development of statistical evidence. Key components in developing such a technique are appropriate statistical techniques as well as the availability of useable data. The EEO-1 report collected by the EEOC provides such information.

WORK FORCE AND INDUSTRY DATA

The EEOC operates a data collection system that, in essence, collects data from all employers in the United

---

[1]The authors are employees of the U.S. Equal Employment Opportunity Commission; however, the views expressed in this article do not necessarily reflect the views of the agency or the U.S. government.

[2] Morris, F. C. (1978) *Current Trends in the Use (and Misuse) of Statistics in Employment Discrimination Litigation*, Equal Employment Advisory Council, Washington, D.C., p. 1.

[3] Meier, P. , J. Sacks, and S. Zabell (1986) "What Happened in *Hazelwood:* Statistics, Employment Discrimination and the 80 Percent Rule,"p. 1 in *Statistics and the Law*, M. Degroot, S. Fienberg and J. Kadane, editors.

[4] See for example, Gastwirth, J. L. (1992) "Methods for Assessing the Sensitivity of Statistical Comparisons Used in Title VII Cases to Omitted Variables", *Jurimetrics, Journal of Law, Science and Technology*, 33: 20-33. Gastwirth J. L. and Greenhouse S. (1987) "Estimating a Common Relative Risk: Application in Equal Employment," *Journal of the American Statistical Association*, 82(397) 38-45.

[5] See *http://www.eeoc.gov/docs/compensation.html* for a copy of this section.

[6] *Good for Business: Making Full Use of the Nation's Human Capital: the Environmental Scan*, Federal Glass Ceiling Commission, Washington, D.C., March 1995, p. 12.

States with more than 100 employees. The EEO-1 data collects information from private employers, the EEO-3 gathers data from referral unions, the EEO-4 report is provided by State and local governments, and the EEO-5 data is obtained from public primary and secondary school districts. Data from colleges and universities once collected on the EEO-6 is now obtained by the National Center for Educational Statistics on its Integrated Postsecondary Education Data System (IPEDS) form.

The EEO-1 data collected by the EEOC provides a useful tool for examining glass ceilings. Carrington and Troske note that "These data [EEO-1] would be useful, but no one has (to our knowledge) used them to study gender segregation . . ." [7](p.510). These authors also note that EEO-1 data come closest to being suitable for the study of interfirm segregation that is truly representative of the national economy .[8]

EEO-1 reports are required by the Equal Employment Opportunity Commission as part of its mandate under Title VII of the Civil Rights Act of 1964, as amended. These annual reports indicate the composition of employer's workforces by sex and by race/ethnic category. In 2000, more than 39,000 employers submitted, as appropriate, individual establishment and headquarters reports for more than 190,000 reporting units with about 53 million employees.[9] The EEO-1 collects data on nine major job categories: (1) officials and managers, (2) professionals, (3) technicians, (4) sales workers, (5) office and clerical workers, (6) craft workers, (7) operatives, (8) laborers and (9) service workers.[10] Race/ethnic designations used in the year 2000 EEO-1 are White (not of Hispanic origin), Black (not of Hispanic origin), Hispanic, Asian or Pacific Islander, American Indian or Alaskan Native. In addition to the workforce data provided by the employer, information about each establishment is added to the database. This includes the establishment's North American Industrial Classification System code, the establishment's county and its metropolitan area code. It should be noted that EEO-1 data is confidential.[11]

A key factor in examining occupational segregation with a focus on glass ceilings is the ability to control for industry. Industrial classifications in many ways capture a firm's market in terms of both labor and products. It might also capture other differences in the manner in which individuals move within an organization, to higher status jobs. A new method of classifying a firm's industry may offer a more precise method of measuring industry for these purposes. From 1938 until 1997, the Standard Industrial Classification (SIC) was used The North American Industry Classification System (NAICS) was adopted by the Office of Management and Budget in 1997. By utilizing a more current and accurate classification system, research regarding equal employment opportunity benefits from classifying industries in a manner that is more likely to reflect a firm's labor market. Firms in the Information sector in the current NAICS system but difficult to categorize in the old SIC system can now be more accurately evaluated as having a different work force rather than having that evaluation based on an artifact of the reporting system's limitation.

ODDS RATIO

The calculation and testing of an odds ratio provides a useful technique for accessing the possible existence of a glass ceiling. (Of course, statistical evidence would just be one indication of a possible problem.) Fienberg, among others, discusses the value of using an odds ratio.[12] Finkelstein and Levin, provide a useful description of this statistic:

---

[7] Carrington, W. and Troske, K. (1995) "Gender Segregation in Small Firms," *The Journal of Human Resources*, 30: 503-533.

[8]Carrington, W. and Troske, K. (1998) "Sex Segregation in U.S. Manufacturing," *Industrial and Labor Relations Review* 51: 445-464.

[9] Private employer required to file are those with: (a) 100 or more employees, or (b) 50 or more employees and: (1) have a federal contract or first-tier subcontract worth $50,000 or more, or (2) act as depositories of federal funds or (3) act as issuing and paying agents for U.S. Savings Bonds and Notes. See *Job Patterns for Minorities and Women in Private Industry* (2000) U.S. Equal Employment Opportunity Commission.

[10] See "Section 5, Description of Job Categories" in the EEO-1 instruction booklet at, *http://www.eeoc.gov/stats/jobpat/e1instruct.html*

[11]EEOC obtains and maintains EEO-1 reports pursuant to its authority under section 709 of the Civil Rights Act of 1964, as amended, 42 U.S.C. 2000e-8. Paragraph (e) of that section prohibits the EEOC and its employees from disclosing EEO-1 reports to the public.

[12] Fienberg, S. (1977) *The Analysis of Cross-Classified Categorical Data*, The MIT Press, Cambridge, pp. 17-18. Also see Agresti, A. (1990) *Categorical Data Analysis* John Wiley and Sons, New York, pp. 14-15 and

The odds in favor of an event are defined as the probability that the event will occur divided by the probability that it will not. That is, if the probability of an event is p, the odds on the event are p/(1-p).

The ratio of two odds, unsurprisingly called the odds ratio, is another statistic frequently encountered in statistical and epidemiological work. For [a] test taken by men and women, the odds ratio is simply the odds on passing for men divided by the odds on passing for women.[13]

In examining glass ceilings, the ratio can be thought of as the odds of men being managers based on their employment in a promotion pool divided by the odds of women being managers based on their employment in the same promotion pool. The promotion pool can be assumed, in many cases, to be found in the EEO-1 job groups of professionals, technicians and sales workers, which can also be referred to as white collar jobs.[14] In examining an EEO-1 report, the odds for men would be the number of men reported as officials and managers divided by the sum of men reported as professionals, technicians and sales workers. The same calculation would be made for women, and the odds ratio would be the division of these two odds:

$$G = (F_m\ M_w) / (M_m\ F_w)$$

where

G = Odds ratio for a glass ceiling
M = Male
F = Female
$_m$ = Officials and Managers
$_w$ = Sum of professional + technician + sales worker job groups

The term $_w$ can be modified to reflect appropriate pools as necessary. An example of the odds ratio calculation to a glass ceiling issue is provided below. As the data is organized as in Table 1, the odds ratio of 0.431 indicates that men have higher odds of being a manager (0.7591) and women have lower odds (0.3270).

TABLE 1: APPLICATION OF ODDS RATIO TO GLASS CEILING

	WHITE COLLAR	MANAGERS
WOMEN	159	52
MEN	357	271
ODDS RATIO	0.431	
LOG OF ODDS RATIO	-0.842	

RESULTS

Glass ceiling-based odds ratios were computed for industries and for individual establishments using the year 2000 EEO-1 file and the four-digit North American Industrial Classification System. A number of steps were taken to examine these odds ratios by industry. The goal of these analyses was to identify those industries where firms were likely to have implemented practices that created glass ceiling issues. A number of preliminary steps were taken in this phase of the research. The odds ratio was transformed to the log of the odds ratio in order to produce a measure that was more symmetric. The industries were subjected to a series of screens to remove those that would have limited practical value

Breslow, N. and Day N. (1980) *Statistical Methods in Cancer Research*, IARC, Lyon, p. 125.

[13] Finkelstein, M. and Levin, B., *Statistics for Lawyers*, Springer-Verlag, New York, 1990, p. 2.

[14]Although this paper treats these three white collar job groups as a single unit, we recognize the need for future refinements that might account for differences in the relative importance of each group. We also acknowledged that white collar jobs may not be the correct feeder pool in some industries. For example, in retail grocery stores, managers often come from sales workers and from laborer job groups. This is another area for future research.

or were unusual in some sense. Specifically, the following screens were implemented. First, industries where the number of white collar workers was less than the total number of managers were removed because the white collar pool would be inappropriate. Second, industries with less than 100 establishments were eliminated to remove small sample size problems such as the ability of a single company to dominate results. Third and in a similar vein, industries characterized by small firms were eliminated. This was done by removing industries where the average number of employees fell into the smallest 10 percent of the distribution of firm size. Small firms are likely to be different from

TABLE 2: INDUSTRIES WITH THE LOWEST MEDIAN LOG OF ODDS RATIO

INDUSTRY	NUMBER OF ESTABLISHMENTS	MEDIAN ODDS RATIO (LOG)	FIVE PERCENTILE	NINETY-FIVE PERCENTILE
Fruit & Veg Preserving & Specialty Food Mfg	462	-1.32176	-3.40120	0.47000
Grain & Oilseed Milling	267	-1.19214	-2.78990	0.28768
Waste Treatment & Disposal	486	-1.11270	-2.77259	0.51083
Department Stores	9,988	-1.11211	-2.72867	0.29123
General Freight Trucking	1,527	-1.09861	-2.48491	0.80078
Pulp, Paper & Paperboard Mills	766	-1.09861	-2.59027	0.61310
Bakeries & Tortilla Mfg	700	-1.01832	-3.42860	1.24171
Remediation & Oth Waste Management Services	311	-0.99622	-2.93039	0.79462

others in a number of important criteria such as organizational specialization.[15] Fourth, the ratio of managers to all employers was examined. The rationale here is that for a glass ceiling issue to be relevant, there has to be a certain proportion of managers. Industries with a small cadre of managers might be unique, and a glass ceiling issue might be irrelevant or at least difficult to measure. This screen was applied by removing those firms whose ratio fell into the smallest 10 percent of this ratio's distribution. Finally, four industries were removed because the white collar pool appears to incorrect for other reasons. In two of these industries Automobile Dealers and Other Motor Vehicle Dealers, status of managers appears to be lower than that of white collar workers, more specifically, sales workers. Residential and Building Construction was eliminated because a large portion of managers are likely to come from craft workers rather than the white collar pool. Finally in this group, Gasoline Stations were removed because the role of sales workers (cashiers) are represent a large portion of the work force but are not likely to be on the management tract.

The EEO-1 statistical file allows the aggregation of reports into 302 different industries based on the four- digit NAICS code. After applying these screens, the number of analyzed firms drops to 164. The mean of the log of odds ratio for these 164 industries is -0.476978, and the median is -0.49433. The standard deviation is 0.31923. The log of the odds ratio for each industry was examined. The median logs of the odds ratios were ranked, and those within the lowest five percentile of industries are displayed in Table 2. The lowest five percentile of the total median log of the odds ratio was -0.9920465. These industries can be thought of as industries with a high potential for gender-based management barriers.

Those industries in the highest five percentile of the median of the log of the odds ratio are displayed in Table 3. The ninety-five percentile of the total median was -0.0441746. These industries can be thought of as the least likely to have glass ceiling problems or the problem is less severe.

Another way of examining the odds ratio results is external to the statistic itself. Industries can be ranked on the basis of employee earnings. This approach provides insights into, not necessarily where the glass ceiling problem is the most severe in terms of employment opportunities, but where glass ceiling problems might result in the most harm with respect to lost earnings.

---

[15] Blau, P. and Schoenherr, R., (1971) *The Structure of Organizations*, Basic Books, New York, p. 56.

TABLE 3: INDUSTRIES WITH THE HIGHEST MEDIAN LOG OF ODDS RATIO

INDUSTRY	NUMBER OF ESTABLISHMENTS	MEDIAN ODDS RATIO (LOG)	FIVE PERCENTILE	NINETY-FIVE PERCENTILE
Scrty & Comdty Contracts Intermed & Brokerage	1,070	0.63426	-1.08619	2.48491
Electronic & Precision Equipment R&M	93	0.40547	-1.30277	2.26176
Cable Networks & Program Distribution	623	0.35767	-0.95885	2.06369
Employment Services	480	0.07033	-1.52424	1.90156
Commercial/Industrial Equip Rental & Leasing	225	0.00000	-1.74297	2.01490
Motion Picture & Video Industries	130	0.00000	-1.25462	1.81529
Professional & Commercial Equip & Supp Whsle	890	0.00000	-1.64866	1.79176
Radio & Television Broadcasting	994	-0.01303	-1.29928	1.02450

In order to rank firms in this manner, a summary database was created that combines the statistical results from the industry odds ratios using the EEO-1 database with basic information from the EEO-1 statistical file and information, particularly salary information, from the County Business Pattern data of 1999.[16] Per capita earnings were computed by dividing annual payroll by the number of employees. Table 4 shows the five industries with the highest per capita earnings. Table 4 shows that those firms with the highest rankings for salary are not ranked the same with respect to women managers, women white collar employees or on the computed odds ratios. However, the rankings for the log of the odds ratio in these high-paying industries do not indicate that they are ranked particularly low with respect to a glass ceiling measure. In fact, these ratios are both positive and negative.

IMPLICATIONS FOR ENFORCEMENT

Those industries with a high potential for gender-based management barriers, as identified in Table 2 might actually be the most difficult to investigate successfully and bring to litigation in order to remove such barriers. Firms in this group might argue that they are no worse than the rest of the industry and there is some relevant labor market factor creating the problem. These industries identified as having a high potential for gender-based management barriers, with the lowest odds ratios, may be better candidates for voluntary assistance efforts, as more wholesale changes are in order. The limited scope required in investigations and litigation might simply be too time consuming to change those industries. The range of the log of odds ratios in this group of industries is such that it includes firms where the log of the odds ratio is greater than zero .[17] (This can be gleaned from the column showing the log of the odds ratios at the ninety-five percentile point in the distribution of these measures.) The firms with high odds ratios in industries identified as having a high potential for gender-based management barriers may be more likely to have instituted personnel policies that are genuinely unlike their competitors and that successfully provide opportunities for career advancement regardless of gender. They may be valuable sources for identifying "best practices." Such practices in industries identified as having a low

---

[16] The County Business Pattern data was downloaded from *http:// www.census.gov/ pub/epcd/cbp.*

[17] While a screen was implemented to only include industries with 100 or more establishments filing EEO-1 reports, the number of establishments reported can drop below 100, as some observations were dropped due to missing values thus reducing the number of establishments in the final analyses.

potential for gender-based management barriers might have been successful due to the industry's labor market rather than the actual technique.

TABLE 4: RANK OF INDUSTRIES BASED ON PER CAPITA EARNINGS

INDUSTRY	RANK FOR PER CAPITA EARNINGS	RANK FOR WOMEN % OF OFF.&MGR	RANK FOR WOMEN % OF WHITE COLLAR	WOMEN % OF WHITE COLLAR	WOMEN % OF OFF.&MGR	RANK FOR ODDS RATIO (LOG)	ODDS RATIO (LOG)
Scrty & Comdty Contracts Intermed & Brokerage	1	103	156	30.1	32.1	20	0.09633
Information Services	2	86	98	42.9	36.0	70	-0.28697
Other Financial Investment Activities	3	76	111	39.8	37.5	40	-0.09722
Independent Artists, Writers & Performers	4	80	127	36.5	36.9	25	0.01760
Software Publishers	5	108	140	33.6	30.9	44	-0.12748

In contrast, industries with a low potential for gender-based management barriers, (see Table 3) include firms with very low odds ratios, much less than zero. (This can be gleaned from the column in Table 3 showing the log of the odds ratios at the five percentile point in the distribution of these measures.) From an enforcement point of view, these firms might be the most effective subjects of investigations and litigation, as the poor performers within the industry will be less able to defend their practices by pointing to labor market competitors or relying on external labor market explanations. The existence of firms with potential glass ceilings in these industries can be demonstrated.

The industry ranked as having the highest median log ratio (see Table 3) and as the highest paying industry (see Table 4) was Security and Commodity Contracts Intermediation and Brokerage (NAICS Code of 5231). This industry was examined to determine the feasibility of finding important outliers. This would be a key ingredient in determining if enforcement would be effective. The odds ratio when computed for individual establishments within this industry has a mean of 3.8 but a standard deviation of 6.8. The log of the odds ratio was 0.67575, and the standard deviation was 1.128. (As expected, using the log transformation reduces the variance.) There were 73 establishments with a statistically significant odds ratio using a Fisher exact test to determine if the odds ratio was significantly different than 1. These 73 establishments actually represent just 40 unique companies. However, the size of the deficits for these firms totals more than 1,500. A limitation to this approach is that deficits can be a function of the number of relevant employees. The firm with the large deficits of women managers might also employ a large number of women managers. Nevertheless, these results suggest that investigative and litigation activities might be successful among these firms.

CONCLUSION

The analysis of EEO-1 data with an odds ratio provides a useful technique for identifying both industries and firms where the employment of women as managers and as white collar employees suggests the possibility of a glass ceiling issue. Further efforts should be taken to develop measures capable of identifying industries that warrant closer examination. In addition, care must be taken to make certain that the pool of potential managers is defined in a manner that reflects the practices of the firm and/or industry. Further research into establishing relevant criteria for identifying such pools would be beneficial. The application of these techniques to race/ethnic groups that might also be confronted by occupational barriers needs to be examined as well. Such applications might have to deal with changes in the relevant pools based on the geographic location of employers.