

**CHAPTER 4
 THE EEO-1 DATABASE**

CHAPTER 4 THE EEO-1 DATABASE.....29
 §1. Accounting for Establishments That Failed to File EEO-1 Reports.....31
 §2. Reliability of the EEO-1 Statistics34
 §3. The Accuracy of EEO-1 Statistics in Understanding the Extent of Intentional Job Discrimination35
 §4. Applying Standard Statistical Principles to the Data to Identify Intentional Discrimination Under Supreme Court Decisions in Employment Discrimination Matters.....38
 §5. Endnotes.....39

The EEO-1 data base contains reports concerning a substantial part of the labor force of large and medium sized establishments in the United States. Establishments are single units, such as an office or a plant, where business is carried on. These reports identify the industry and the metropolitan area in which the establishment is located. For 1999, the EEO-1 reports cover about thirty three percent of the nation’s workforce.

Total full time workforce in 1999 was 133,488,000.⁴¹ About 25% of workers were in establishments of fewer than 20 employees, and 54% in establishments of fewer than 100 employees.⁴² Therefore, the EEO-1 reports could cover at most some 50% of the work force. Since we focus this study on establishments that are in Metropolitan Statistical Areas, we exclude about 20% of establishments that are not in these areas.

Table 1. The 1999 Census of the labor force and the EEO-1 population

THE 1999 LABOR POPULATION			
	Total Labor Force	Minorities	Women
Census Bureau*	133,488,000	43,958,000	62,042,000
EEO-1 Reports**	37,360,708	11,160,305	17,657,992
*US Census Bureau, Statistical Abstract of the United States:2000, Table #669, p.416.			
** All Establishments reporting on EEO-1 forms			

The universe for this study consists of the establishments within Metropolitan Areas that have 50 or more employees. The basic information in the EEO-1 report divides the workforce into nine occupational categories: officials and managers, professionals, technical workers, sales, office and clerical, craft workers, operatives, labor, and service workers. These categories are subdivided by sex, and by minority status, meaning Black, Hispanic, Asian Pacific, and Native

American employees.⁴³ Native Americans are not further examined here for technical reasons.⁴⁴

For 1999, we have analyzed the 160,297 reports filed by establishments with 50 or more employees that operate in Metropolitan Statistical Areas (MSA).⁴⁵ The MSA's are defined by the Census Bureau as "a core area containing a large population nucleus, together with adjacent communities that have a high degree of social and economic integration with the core."⁴⁶ We treat each MSA as a labor market because of this social and economic integration. The MSA's are established, and are generally used, for governmental and private research purposes unrelated to the study of discrimination.

We compared only establishments in MSA's that employ at least 50 workers to ensure reliability of the data. We required that an establishment have at least 20 employees in the occupational category examined; that there be two other establishments with at least 20 employees in that occupation; that there be at least 120 employees in the occupation in the MSA; and that no establishment have more than 80% of the employees in order to have sufficient employment to assure that there was a labor market for such workers, and that no single establishment dominated the market.⁴⁷ This data set is identified as the EEO-1 Labor Force.

Within the EEO-1 Labor Force, we examine each industry separately, because different industries have different technologies and different employment needs. Each establishment describes its principal product or activity on its EEO-1 form. Establishments are then classified by industry in accordance with the 1987 *Standard Industrial Classification (SIC) Manual*, Office of Management and Budget. This is a classification structure for the national economy. It provides data according to the level of detail, from the general to the quite specific. For example, manufacturing is a major industrial division. Food and kindred products (Code 20) is one of its major groups. This group is further divided into meat products (Code 201) and meat packing plants (Code 2011).⁴⁸ The major industrial divisions are identified by 1-digit codes, major groups by 2 digits, and further subdivisions by 3 and 4 digits. The major divisions in the private sector are: Agriculture, forestry and fishing; Mining; Construction; Manufacturing; Transportation, Communications, Electric, Gas and Sanitary Services; Wholesale Trade; Retail Trade; Finance, Insurance and Real Estate; and Services. The SIC numbers in the following tables refer to that classification system. Appendix B contains a list of SIC codes including the 1, 2, and 3 digits used in this report. In the nationwide report, we review industries by the two digit standard, while in the state reports, we use the more detailed three digit standard.⁴⁹

§1. ACCOUNTING FOR ESTABLISHMENTS THAT FAILED TO FILE EEO-1 REPORTS.

One third of the establishments in Metropolitan Statistical Areas with at least 100 employees failed to file EEO-1 reports in 1999, despite legal requirements that they do so.⁵⁰ This amounted to 46,585 establishments. The EEO-1 reports include 76% of the employees reported by the Census Bureau. Most of the failures to report – 44,328 out of 46,585 – employ between 100 and 500 employees. 83% of Establishments with more than 500 workers do file these reports.

We have compared the EEO-1 data against the Census data. The result of the comparison is in the following Table.

Table 2. EEO-1 and Census reports, by size of establishments

Number of Reporting Establishments in the EEO-1 Data Set (Population and Labor Force) and Number of Establishments in the United States, 1999			
Total Number of Establishments by Size			
	EEO-1 Population	US Census	Percentage in EEO-1 data set
<50	8,275	6,632,900	0%
50-100	75,252	205,306	37%
100-499 Employees	96,315	151,940	63%
500-999 Employees	9,614	11,644	83%
1000+ Employees	5,572	6,654	84%
Total Establishments	195,028	7,008,444	3%
Total Establishments > 50	186,760	375,544	50%
Total Establishments >100	111,508	170,238	65%
Estimated Employment in establishments of 50 or more employees	43,124,864	57,703,695	75%
Number of Establishments in MSAs with 50 or more Employees			
	EEO-1 Labor Force	US Census	Percentage in EEO-1 data set
50-100	65,038	172,027	38%
100-499 Employees	82,368	126,696	65%
500-999 Employees	7,945	9,346	85%
1000+ Employees	4,946	5,796	85%
Total Establishments in MSAs > 50	160,297	313,865	51%
Total Establishments in MSAs >100	95,259	141,838	67%
Estimated Employment in Establishments of 50 or more employees			
Derived from census datatable 4.xls	37,360,708	48,851,394	76%

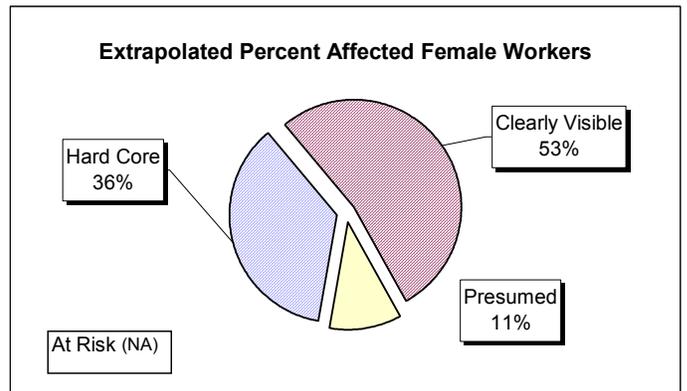
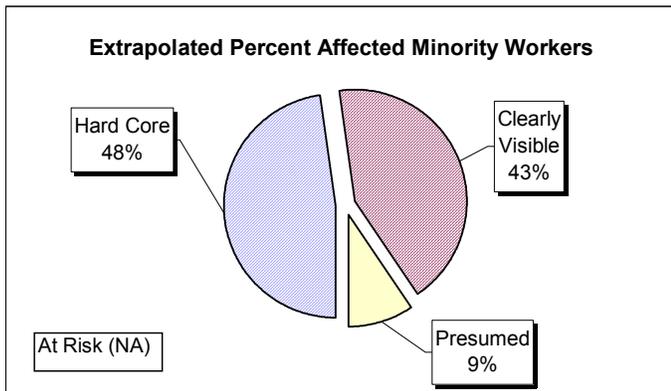
Those employers who do not file are in violation of their legal obligations under Title VII of the Civil Rights Act of 1964 and, if applicable, Executive Order 11246. Both the duty not to discriminate and the duty to report flow from the same Equal Employment Opportunity laws. Those who ignore the filing requirements may also ignore other obligations under the laws against discrimination. Since they ignored their duty to report, they may have violated the duty not to discriminate. Because they did not report, we cannot establish their industries, locations or the exact numbers of employees they had in each occupation by race, sex or national origin. Therefore we cannot directly apply the methodology used in this study. However, because these employers have violated the reporting provisions of the law, we have no reason to assume that they employed more minorities and women and at higher positions than the averages of those who did report. Accordingly, we apply the average of those who reported to those that failed to report.⁵¹

As a result of this extrapolation, 75,793 establishments appear to engage in intentional discrimination against minorities in one or more of the nine occupational categories. This conclusion is based on the actual count of 41,020 establishments, coupled with 34,773 extrapolated. Similarly, 60,425 establishments appear to engage in intentional discrimination against women, based on an actual count of 34,309 establishments coupled with 26,116 extrapolated.

For the reasons outlined above, we cannot separate the non-reporting establishments into specific minority components, or industries or occupations. Therefore, we apply the extrapolation only to the general national statistics. For all other statistics, we report only the actual numbers, with the reminder that we may be undercounting intentional discrimination to a substantial extent because of the failure of establishments to report. The following table explains the extrapolation.

Table 3. Extrapolations to take account of non-reporting establishments, in numbers of establishments and affected workers.

Extrapolations to take account of non-reporting establishments					
ESTABLISHMENTS	Hard Core*	Clearly Visible*	Presumed*	At Risk*	All
Minority, Actual	12,739	15,906	6,782	5,593	41,020
Minorities, Total (incl. Extrapolated)	22,269	29,656	13,099	10,768	75,793
Female, Actual	8,222	14,801	5,696	5,590	34,309
Females, Total (incl. Extrapolated)	13,173	26,177	10,534	10,541	60,425
Minorities just Extrapolated numbers	9,530	13,750	6,317	5,175	34,773
Females just Extrapolated numbers	4,951	11,376	4,838	4,951	26,116
AFFECTED WORKERS	Hard Core*	Clearly Visible*	Presumed*	At Risk*	All
Minority, Actual	432,959	359,219	74,087	na	866,265
Minorities, Total (incl. Extrapolated)	649,267	584,467	127,349	na	1,361,083
Female, Actual	240,908	324,924	62,563	na	628,395
Females, Total (incl. Extrapolated)	343,398	504,513	104,221	na	952,131
Minorities just Extrapolated numbers	216,309	225,247	53,262	na	494,818
Females just Extrapolated numbers	102,489	179,589	41,658	na	323,736
COMPARISONS					
106,758 comparisons for minorities	205,215 establishments with extrapolations, minorities				
108,130 comparisons for females	208,393 establishments with extrapolations, female				
Notes: Establishments are classified into the "worst" discrimination category they fall into for any occupation. Extrapolations are based on 1999 County Business Patterns data. * Terms explained in Chapter 7					



§2. RELIABILITY OF THE EEO-1 STATISTICS

The statistics in the EEO-1 reports contain assurances of reliability. They are prepared by employers, pursuant to a legal requirement to do so. The signing official does so under risk of criminal prosecution for making willfully false statements.

The judgments about race, ethnicity or sex made by the employer are usually based on the employer's impression, rather than by asking the employee about his or her sex, race or ethnic background. For purposes of addressing discrimination, this is a very important consideration. Employers will apply their understanding of shared community impressions of color, sex, and language. These are the bases on which discrimination is founded, not on the internal and personal sense of individual workers about themselves.

The employer also decides to assign each employee to one of the nine occupational categories. While the categories are broad (see Appendix A), they have been in use for more than thirty years with only occasional conflicts about the appropriateness of the decision to treat a worker in one category rather than another. Their breadth is limited by the industry in which they work.

The employer defines the industry in which it is primarily engaged, under the standards of the Standard Industrial Classification System established by the Office of Management and Budget for a variety of informational purposes not related to discrimination issues.

Metropolitan Statistical Areas are generally used to define labor markets in an integrated economic and social environment for governmental and private purposes also not focused on discrimination issues.

The Joint Reporting Committee, established by the EEOC and the Labor Department, receives the employer reports, has the information computerized, and supplies the information to the two agencies. For thirty-five years, EEOC has issued numerous surveys based on information derived from the employers. As far as we know, no one has challenged the accuracy of the process of computerizing the data, or the appropriateness of its use in describing the utilization of minorities and women in the workforce. It is appropriate to conclude that the information on which this study is based was accurately collected and computerized by the government. In short, the facts about the employment of persons by race, color and national origin, and occupations, industries and metropolitan areas contained in the EEO-1 reports are reliable.

§3. THE ACCURACY OF EEO-1 STATISTICS IN UNDERSTANDING THE EXTENT OF INTENTIONAL JOB DISCRIMINATION

The presumption that intentional discrimination is present when an establishment is more than two standard deviations below the average among its peers is an evidentiary principle designed by the Supreme Court to flush out “clandestine and covert” intentional racial discrimination against minorities.⁵² This presumption enables those challenging an establishment’s practices to require the employer to show that the statistics are inaccurate, or that it had only a non-discriminatory reason for its actions.

As the Technical Appendix demonstrates, questions can be raised about the accuracy of the MSA statistics arising from the fact that we do not know the exact location of establishments in the MSA. The Government, of course, has this information. We chose to examine only the MSA data to avoid EEOC pre publication review of our study.⁵³ Three parameters are used to identify intentional discrimination in this study: Metropolitan Statistical Areas, Industries and Occupations. All three have long been used in law and social sciences. Though singly each is quite broad, together each serves to sharpen the focus of the others to provide a practical framework for identification of intentional job discrimination. The content of each occupational category is limited by the technological requirements of the industry to which it is applied. Professionals in Scheduled Air Lines will include many pilots; while in Legal Services, the term will include many lawyers. The same is true of other occupational categories and the industries to which they relate.⁵⁴

Similarly, by comparing establishments operating in the same Metropolitan Statistical Area, this study ensures that establishments are subject to similar geographic, sociologic and economic realities. The MSAs areas are defined by Office of Management and Budget and applied to Census data, and used in business and social science research. The general concept is that of a core area containing a large population nucleus – usually at least one central city with at least 50,000 population and a total metropolitan population of at least 100,000 – having a “high degree of economic and social integration.... [O]utlying counties are included in the MSA if they meet specified requirements of commuting to the central counties and other selected requirements of metropolitan character (such as population density and percent urban.)”⁵⁵ The recognition of commuting patterns in the metropolitan area along with requirement of a “high degree of economic and social integration” combine to identify a labor market in which job information is shared through informal and formal sources.

Patterns of residential segregation may give rise to the issue of whether the MSA is an appropriate labor market for use of the two standard deviation analysis in jobs where establishments hire from nearby neighborhoods. If most establishments are located near each other, this problem is less significant, because they may all draw from the same neighborhood pools. But if establishments are widely scattered through different residential areas, there is the possibility that the neighborhood in which the establishment is located and supplies labor for some of the jobs will have a different racial profile from other neighborhoods, which is lower than the MSA profile. This might tend to overstate the extent of intentional job discrimination. In analyzing this issue, of course it would be helpful to have exact location information of the establishment in question.⁵⁶ In the absence of that information, it is necessary to consider the possibility on broader grounds.

The first consideration is that our methodology may understate the extent of intentional job discrimination in the MSA labor force. If a substantial part of the industry in the MSA is engaged in discrimination against minorities or women, their average utilization will be relatively low, and few intentional discriminatory will be recognized. But we cannot identify this situation. We accept the average as a fact that we do not question. An evaluation of the average must come from some other source of information or other analysis of it.

The second consideration is that minorities may be scattered through the MSA outside of heavily segregated areas so that they are in reasonable commuting distance from establishments and that many establishments may be located in areas where they may draw customers or consumers from racially different areas. In either case, the fact that there are segregated residential areas may not explain the low utilization of minorities. Our data base for grocery stores, for example, shows that there are 239 stand alone grocery stores out of a total of 7,925 stores reporting. The rest are stores that have a parent corporation. None of the stores have fewer than 50 employees. Many Grocery chains advertise by mail to their customer base which may extend across segregated residential areas, or have stores in residential areas with different racial patterns. In either case, provision of vacancy information may be nearly cost free.

Many of these stores use walk-ins or word of mouth recruiting techniques which favor residents who live in the neighborhood. Employers who have defended their low utilization of minorities on grounds of residential segregation have not fared well in the courts, particularly where they used no forms of recruitment other than walk-ins and word of mouth. The two standard deviation analysis has been used in two different ways in these cases. The analysis has been used to support a finding of intentional discrimination.⁵⁷ It has also been used to

support a finding of disparate impact discrimination.⁵⁸ Where the good faith of the employer is not at issue, localized hiring may be responsible for an unjustified disparate impact of its practices on minorities or women.⁵⁹ Either of these findings could be based on the two standard deviation analysis. We cannot tell which of these analyses would apply in any particular case. We include these establishments here, with the understanding that local issues can be raised to show the inadequacy of the statistics, or an exclusive non discriminatory reason for the low utilization.

The identification of relevant labor markets for employment discrimination law purposes has long been a source of disputes. The MSA definition may be as specific as research can get without being designed to favor one side or another in administrative or judicial proceedings. Employers tend to argue for areas that have the fewest qualified Minorities; Plaintiffs argue for areas that will provide higher Minority employment.⁶⁰

These issues – and others – concerning the definition of a labor market may arise in the specific situation of a given establishment which is alleged to be two or more standard deviations below the average.⁶¹ Knowledge of facts that are important in a determination of the appropriate labor market will most likely reside with employers through their day to day experiences. For this reason, once an establishment has been identified as two or more standard deviations below the MSA average, the burden rests on that establishment to show that a different labor market is more appropriate and will result in a lower average utilization level than that identified, so that the establishment does not belong in the 2 standard deviation category.

Decisions concerning these statistics in legal proceedings as well as use of this methodology for research, should consider the value of the MSA as labor market in order to further the Supreme Court's policy to flush out "clandestine and covert discrimination."⁶² These advantages include:

1. **Neutrality** — The MSA statistics were neither designed nor arrayed to support the claims of employers or workers in particular cases. This cannot be said of any statistical reports prepared for litigation.
 2. **Non-exclusiveness** — Admission of these statistics does not preclude any party from introducing alternate otherwise admissible statistical reports
 3. **Customary practice** — These statistics provide the first unbiased comparison of similarly situated establishment's utilization of minorities and women in the same labor market. They present a unique
-

insight into the practices of the industry in question in specified labor markets.

4. **Employer opportunities to plan** — Ability to address potential allegations of discrimination without awaiting costly and time consuming litigation is enhanced by the availability of two standard deviation information that may enable an employer to address existing practices without the pressures of the formal legal system. The likelihood that such planning will take place is enhanced if the legal system takes the two standard deviation rule seriously.⁶³
5. **An abstract challenge to the statistics** — Without presenting an alternative labor market analysis that shows that the establishment is not two standard deviations below the average, this will sound much like a general claim of “good faith” which was rejected as an insubstantial response to a “pattern or practice” claim of discrimination by the Supreme Court in 1977.⁶⁴

For the purposes of public understanding of the nature and extent of intentional job discrimination, the use of the MSA average appears to be the best approximation that is possible, and has a solidity that personal observation cannot achieve in our increasingly complex world.

§4. APPLYING STANDARD STATISTICAL PRINCIPLES TO THE DATA TO IDENTIFY INTENTIONAL DISCRIMINATION UNDER SUPREME COURT DECISIONS IN EMPLOYMENT DISCRIMINATION MATTERS.

Once the statistical facts about the employment patterns of workers have been established, they must be interpreted. The interpretation of statistics is not always easy. In this study, we have used standard statistical principles to analyze the data.⁶⁵ We then apply legal principles about the use of statistics in proving intentional job discrimination laid down by the Supreme Court in 1977, reaffirmed in 1984, and relied upon since then by judges, including particularly Supreme Court Justice O’Connor. The next chapter describes our application of these principles to the statistics.

§5. ENDNOTES

41. US Census Bureau, Statistical Abstract of the United States:2000, Table No 643, p. 403, Table No. 868, p.544.
 42. *Id.* at p. 544, Table No. 866. Statistics are for 1997.
 43. The content of each of these categories and the description of the minorities in the report is described in Appendix A, the EEO-1 instruction booklet.
 44. The data on Native Americans does not include employment on reservations; and is too limited to permit the analysis of patterns of discrimination.
 45. A total of 195,028 establishments filed EEO-1 forms for 1999. Of these, 36,758 establishments with 5,927,702 workers including 1,080,094 minorities and 2,758,037 women were located outside Metropolitan statistical areas. We did not analyze them because we are unsure whether they were located in a labor market. We could not obtain information containing names or specific identifying addresses without submitting our work to a review by EEOC. To assure compliance with the confidentiality requirements, future research using these reports may be possible, but is likely to require supervision by EEOC. We have not attempted to extrapolate from MSAs to surrounding areas, as we did in the preliminary reports on the states of Washington and Georgia.
 46. US Census Bureau, Statistical Abstract of the United States:2000, p. 4, 908-917.
 47. Technical Appendix, §1.
 48. Statistical Abstract, 2000, p. 533-34.
 49. For explanation of the Standard Industrial Classification methodology, see Chapter 9, §4.
 50. See Table 2. Establishments with fewer than fifty employees are not required to report unless they are subsidiaries. Establishments with between 50 and 100 employees are required to file if they are federal contractors, or first tier subcontractors. It is impossible to estimate the percentage that is so required. For 1999, EEOC received 65,038 such reports while Census reported 172,027 establishments, for a ratio of 38%. Equivalent percentages were found in 1997 and 1998 with respect to establishments with 100 or more employees. Details of filing requirements appear in Appendix A.
 51. Explain the method of extrapolation
 52. Teamsters, Ch. 2, n. 16
 53. See note 5, *supra*.
 54. See technical appendix, §1.
 55. Statistical Abstract of the United States, 1999, p. 914.
 56. See Technical Appendix, § 2
 57. Barbara Lindemann and Paul Grossman, EMPLOYMENT DISCRIMINATION LAW, 3RD Edition, 1996, 697-713, esp. 700 at n. 13 (1996), Phillip J. Pfeiffer, EMPLOYMENT DISCRIMINATION LAW, 3RD Edition, 2000 Cumulative Supplement, .441-450 (2000)
 58. See Chapter 5, § 7 n 5.
-

59. *Watson v. Fort Worth Bank and Trust*, 487 US 977, 995 n.3 (1988)
 60. Barbara Lindemann and Paul Grossman, *EMPLOYMENT DISCRIMINATION LAW*, 3RD Edition, 1996, 1717-1737 (1996), Phillip J. Pfeiffer, *EMPLOYMENT DISCRIMINATION LAW*, 3RD Edition, 2000 Cumulative Supplement, .945-951 (2000)
 61. *Hazelwood School District v. United States*, 433 US 299 (1977)
 62. *Teamsters*, , supra Chapter 2, note 16. Admissibility decisions are governed by *Bazemore v. Friday*, 478 US 385 (1986). See Barbara Lindemann and Paul Grossman, *EMPLOYMENT DISCRIMINATION LAW*, 3rd Ed., 1728-1737. For cases upholding the use of the MSA as a proper measure of the geographic scope of statistics, see p. 1721, note 116. Use of the methodology by federal or state enforcement agencies is not as subject to the risk of being overbroad because the agencies, having the exact addresses of establishments can locate them specifically before commencing an investigation. An initial step in any investigation would be to obtain the employer's explanation for the statistics which can be evaluated to determine the utility of proceeding further. In private litigation, issues of the scope of the labor market may be addressed in summary judgment proceedings before trial.

The methodology and information in this study may illuminate—and be illuminated by—researchers in other disciplines concerning a wide range of issues other than the extent of intentional job discrimination. Researchers seeking more specific location information than is contained here may be able to obtain it from EEOC on condition that the agency reviews all such research prior to publication to ensure that the confidentiality provision of Title VII has been honored.
 63. *Faragher v. City of Boca Raton*, 118 S. Ct. 2275 (1998); *Burlington Industries, Inc. v. Ellerth*, 118 S.Ct. 2257 (1998)
 64. *Teamsters v. United States*, Chapter 2, n. 16.
 65. See Technical Appendix.
-